**Research Paper**

# A Cross-cultural Study on the Discussion Sections of Ph.D. Dissertations with a Focus on Academic Collocations

**Vida Fathi Bonabi[1], Nesa Nabifar[2*], Saeideh Ahangari[3]**

[1]Ph.D. Candidate, Department of English, Tabriz Branch, Islamic Azad University, Tabriz, Iran
[2*]Assistant Professor, Department of English, Tabriz Branch, Islamic Azad University, Tabriz, Iran
[3]Associate Professor, Department of English, Tabriz Branch, Islamic Azad University, Tabriz, Iran

**Abstract**

The present study aimed to examine the rhetorical structure of the discussion sections of English Ph.D. dissertations authored by Iranian students in an EFL context, Indian ESL students and native (English speaking) Ph.D. students at Applied Linguistics. To this end, 300 discussion sections were gathered from three contexts under study at the time frame of 2005 to 2020. Following Benson et al's (1986) model, the researchers analyzed the discussion sections of the dissertations for both lexical and grammatical collocations and their different sub-classifications. The findings showed that the natives overused collocations in a significantly greater number in developing the Ph.D. dissertations discussion section than those of EFL and ESL Ph.D. students' dissertations. Moreover, despite variations, there were similarities between EFL and ESL Ph.D. students' dissertations in terms of utilizing the sub-categories of collocations that can be considered as signs for standardization of academic writing by non-native speakers of English. The implications for researchers, teachers and students were discussed.

## INTRODUCTION

It is becoming evident that language is primarily formulaic, and that understanding formulaic sequences is an important part of speaking smoothly and spontaneously (Schmitt, 2004; Wray, 2012). It's also been suggested that formulaic language has a big impact on language learning (Durrant & Schmitt, 2009). In a variety of academic forms, corpus-based studies have progressively documented the importance of formulaic multiword sequences, such as collocations (Hyland, 2008; Martinez & Schmitt, 2012). Collocations are defined as words or collocates that have a particularly strong link with one another in generating their meaning (Wray, 2008). They were first proposed by

Firth, who stated, "You shall know a word by the company it keeps" (Firth, 1957, p. 14).

Collocations are prevalent in natural language use (e.g., Erman & Warren, 2000; Nagao, Makoto & Mori, 1994; Xiao, & McEnery, 2006), as corpus studies have showed (e.g., Erman & Warren, 2000; Nagao, Makoto & Mori, 1994; Xiao, & McEnery, 2006), that collocations play a significant role in distinguishing socially-s (Hyland, 2012). It is becoming clear that language is mainly formulaic in nature, and that mastering formulaic sequences is a key aspect of speaking fluently and naturally (Schmitt, 2004; Wray, 2012). They can be thought of as the foundations of academic writing, and learners and researchers must be familiar with how to use them in academic writing. Academic writing requires authors to

*Corresponding Author's Email:
nabifar@iaut.ac.ir

textualize their work as a contribution to the field, as well as creating themselves as credible representatives of the domain (Hyland, 2001). For most of the postgraduate students writing is a challenging and formidable task, they face more difficulty in writing the discussion of the findings after reporting the results. Many problems observed in this section are related to the writers' obscure perception of their variables on the one hand and underdeveloped writing skill and unfamiliarity with principles of academic writing on the other hand. In order to write successfully, researchers and students, particularly non-native postgraduate students, should be conversant with the development of formulaicity in academic discourse. Furthermore, in order to join a discourse community, students must be familiar with the rhetorical structures of research articles and their various parts, such as abstracts, introductions, methods, discussions, and so on; thus, students must learn the rhetorical structures and linguistic features of academic writings.

Studies of collocations and formulaic sequences are found in various disciplines and across cultures in research articles and academic writings such as (Abdollahpour & Gholami, 2018; Chen & Baker, 2010; Durrant & Schmitt, 2009; Doró, 2014; Pho PD, 2008; Samraj, 2005). However, to the best knowledge of the researchers, no study has considered the academic collocations and their sub-classifications based on a well-known framework that is Benson et al (1986) model in the discussion sections of Ph.D. dissertations in Applied Linguistics among three contexts of EFL, ESL, and native speakers of English. To fill this gap, this exploratory study aimed to explore the types, frequencies and differences of collocations in the discussion sections of dissertations written by native and non-native EFL Ph.D. students. This study is significant since through the findings of this study more contributions should be made to writing more clear and structured discussion sections in the theses and dissertations by non-native ELT students.

**LITERATURE REVIEW**

The current research focuses on collocations, which are classified as formulaic sequences.

Formulaic characteristics of language account for a significant amount of everyday language use (AlHassan & Wood, 2015). A formulaic sequence (FS) is conceptualized by Wray (2012) as a sequence, continuous or discontinuous, of words or other meaning elements, which is or appears to be, prefabricated, that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar. One type of formulaic sequence is collocations, which are arbitrarily limited pairs of lexical items (Granger, 2018). In the literature, collocations are defined in a variety of ways. In 1957, Firth is credited with coining the term collocation. Firth argued that you shall know a word by the company it keeps (1957, cited in Palmer, 1981). Collocation is a concept he uses to describe the pairing of terms in the same context, such that if you see one, you're likely to encounter another. Thornbury (2002, as cited in Mounya, 2010) notes that two words are collocates if they occur together with more than chance frequency, such that, when we see one, we can make a fairly safe bet that the other is in the neighborhood. Collocations are two or more words that are consistently coupled, implying that some sentences co-occur in natural language with higher than random frequency, regardless of how they are defined. In this approach, the notions of co-occurrence and frequency are stressed. Because the phrase "fast" 'co-locates' with the word "food," "fast food" is a "collocation. This means that when we see speed, we expect to see food. We also remark that we can't use quick or rapid in lieu of fast because it does not sound right. According to the Oxford Collocations Dictionary, collocation is the way words mix in a language to make natural-sounding speech and writing (2002).

According to Smadja (1990) collocational information, which might be difficult to obtain for second language learners, particularly non-native English speakers, is extremely useful in sentence construction. Non-native English language users have a limited knowledge of phraseological patterns utilized in the English language, according to previous study (Cotos, Huffman, & Link, 2015).

Nowadays, collocation has been generally considered as an important characteristic to distinguish native speakers from non-native speakers. Research on academic writing has also revealed that collocations are not only common but do have a particularly important discourse function within the academic community. Laufer and Waldman (2011), for example, argue that collocations are an integral part of academic writing. Collocations, according to some studies, are a feature of scholarly writing (e.g., Gledhill 2000, Li & Schmitt 2009, Paquot 2008).

Academic writing has long been regarded as a challenging undertaking for many EFL students at the graduate and postgraduate grades, especially for non-English speakers (Abdollahpour & Gholami, 2018). In order to be successful and engaging writers, EFL learners and beginner investigators need to understand some natural fragments and cycles of words. By adopting, they will be able to prepare effective scholarly articles and attract any discourse group to which they are aimed. To do so, they need use natural terms or collocations, which are the topic of phraseology in applied linguistics, to write fluently. The study of the structure, meaning, and use of collocations is referred to as phraseology (Cowie & Howarth, 1996). According to Cowie (2001), phraseology varies from lexicology in that the latter is concerned with the meaning of single words, while the former is concerned with "linguistic units," or collocations that have become particularly significant in language development. According to Glaser (1998), "phraseology as a sub discipline of the linguistic system is an expanding field of research and has attracted interest from many sides". (p.125)

Formulaic language, FSs in general and collocations in particular, has been studied by corpus linguistics and psycholinguistics (e.g. Biber, 2006; Cortes & Csomay, 2015; Ellis, 2012). They have also gained attention in academic writing, in particular, in case of non-native novice writers (e.g., Chen & Baker, 2010; Ebeling & Hasselgård, 2015; Bian & Wang, 2016). One of the most noteworthy findings in this respect is that non-native nov-

ice authors prefer to stick to a restricted number of formulaic terms that are overused (Hyland, 2008; Salazar, 2014). One explanation for the lack of concentration could be that collocations are said to make up the bulk of native speakers' mental lexicon, implying that they are more extensively preserved and remembered from memory than from a database (Wray & Perkins, 2000). Shin and Chon's (2019) study, on the other side, discovered the inverse. The use of lexical bundles (frequently repeated word sequences) was compared in L1 and L2 novice academic writing corpora that were closely resembled for register and writing prompts. The researchers searched the datasets for lexical bundles and studied their function and structure. The results indicate that the two groups employ bundles in a comparable way, hence the data are valuable in determining the extent to which native and nonnative students enter college with a working knowledge of formulaic language.

The current study is an effort to compare native and non-native English speakers in two contexts of EFL and EST. There are still few studies that are similar to this one; for example, Durrant and Schmitt (2009) looked at two groups of writers' argumentative essays prepared under timed settings by postgraduate students on pre-sessional EAP courses at a British university, and first-year undergraduates on in-sessional EAP courses at an English-medium university in Turkey. Non-native authors heavily rely on high-frequency collocations, but they underuse less frequent, strongly linked collocations, according to the findings (items which are probably highly salient for native speakers). These data support the notion that non-native writing is devoid of idiomatic phraseology. As stated, despite the presence of enough literature regarding collocations, there is a gap of research as to the frequency, types and the differences of academic collocations employed in the discussion sections of Ph.D.dissertations written by ELT students from three contexts of EFL (Iranian non-native speakers of English), ESL (Indian Ph.D. s), and native speakers of English in one study. Then, each study or research which can bridge this gap is of primary importance. The

research question addressed in this study is: Are there any significant differences in terms of the frequency and types of the academic collocations used in the discussion sections of Ph.D. dissertations written by native and non-native ELT students, namely, EFL students, ESL students, and native students?

## METHOD
### Corpus

The purpose of the current study was to investigate the frequency, type, and differences of academic collocations in the discussion sections of Ph.D. dissertations written by native and non-native ELT students from three countries. To fulfill this purpose, the researcher gathered the data from the discussion sections of 300 ELT post graduate students' dissertations (100 Iranian ELT students' Ph.D. dissertations, 100 Indian ELT students' Ph.D. dissertations, and 100 English ELT students' Ph.D. dissertations) within the time period from 2005 to 2020 amounting to 398,432 words in total. The reason for choosing Ph.D. s within a 15-year periods was the fact that the commencement date for the award of Ph.D. s in each country was different and sometimes the

graduation time is different in countries, for example, in Iran, the students can complete Ph.D. s even after 6 years (12 semesters) or more and suspend it until their papers extracted from their theses are published in one of the high value journals, but in some countries, the students can graduate just after completing the dissertation without the need to publish articles. The data was gathered from different websites of the universities in Iran (as an EFL context) such as Tabriz University, Isfahan University, Tehran University, Ahvaz University, Urmia University, and so on as well as from Irandoc.ir as the bank for Iranian theses and dissertations. For India (as an ESL context), the national university of India was the reference and database for the data collection, and regarding native speakers' dissertations, the UK and USA universities like Lancaster University, McGill University, the University of Birmingham, the University of Leeds and the University of Maryland and their websites were focused. In case of native authors, the researchers made sure that the name of the author was among English names as far as possible.

**Table 1**
*Grammatical collocations structure (adopted from Benson et al., (1986))*

| Type | Structure | Example |
|------|-----------|---------|
| G1 | noun + preposition (excluding noun + of, noun + by, noun + concerning; regarding; in regard to; with regard to) | The results to a different population |
| G2 | noun + to + infinitive | The students to accept |
| G3 | noun + that-clause | The truth that |
| G4 | preposition + noun | For a job |
| G5 | adjective + preposition | Significant in predicting |
| G6 | adjective + to + infinitive | Cell phones to access news |
| G7 | adjective + that-clause | More sensitive that |
| G8 | 19 English verb patterns | Long (1983) proposed a model to account for the relationship…. |

### Theoretical Framework

The design of the current study is descriptive-exploratory one. To fulfill the aim of this study, frequent academic collocations in the corpora were identified and compared in order to prepare a list of the collocations and their different types based on the Benson et al's (1986) classification of English collocations. Based on this classification, collocations fall

into two major groups: lexical collocations and grammatical collocations. Lexical collocations are further divided into seven types, whereas grammatical collocations are divided into eight. Lexical collocations contain nouns, verbs, adjectives, and adverbs. Lexical collocations may be verb+ noun, adjective+ noun, noun+ verb, adverb+ adjective and verb+ adverb. They may be verb+ noun/pronoun (or

prepositional phrase), (e.g. reach a verdict), verb+ noun (e.g. reject an appeal), adjective+ noun (e.g. strong tea), noun+ verb (e.g. blood circulates), noun+ of+ noun (e.g. a pack of dogs), adverb+ adjective (e.g. hopelessly addicted), and verb+ adverb (e.g. affect deeply). In contrast to lexical collocations, grammatical collocations not only contain a dominant word such as a noun, adjective, and verb, but also involve a preposition or grammatical structure such as infinitives and clauses. Grammatical collocations are delineated from G1 to G8 (Hung & Chin, 2018) as displayed in Table 1. It is worth noting that the examples were obtained from the corpora under study.

## Procedure

At the onset of the study, the dissertations were downloaded which required a strenuous effort on the part of the researcher. The library websites of the mentioned universities were searched in order to find the targeted ELT dissertations to include in the study. Some of ESL or native speakers' dissertations were obtained through the ProQuest Dissertations and Theses Portal. The British Library Electronic Digital Thesis Online Service (EThOS) was also helpful for identifying possible Ph.D. dissertation to include in the study. Early Ph.D. s (2005 to 2010), however, were especially difficult to locate and inquiries were sent to university librarians for help with this as well as for assistance in obtaining copies of the Ph.D. dissertations

Library records were, however, sometimes incomplete, and some of the earliest doctoral theses did not have discussion sections or their discussion sections were merged in the result sections. Hence, the researcher was obliged to exclude them from the process of the research. Evans (2003). Liu and Buckingham (2018), and Simpson (2009) reported having similar problems in their collection of doctoral theses as did Herman (2017) in a recent study of the development of doctoral education in South Africa. However, the difficulties faded as the research in corpus area and dissertation analysis were worth bearing the difficulties. The reason for choosing Ph.D. dissertations was the dramatic growth of Ph.D. dissertations across the world in recent decades (McCarthy & Wient, 2019; McIntyre, 2015).

Based on the researcher's initial survey, the average length of Iranian Ph.D. dissertations' discussions ranges from 800 to 1100 words that the second survey approved 90506 words in total. The discussion sections of Indian authors ranged from 1000 to 2500 words, with an average length of 118920 words. Finally, the discussion sections of native authors ranged from 1100 to 2700 words, with an average length of 121430 words. Then, the discussion sections were extracted and converted into Plain Format. To fulfill the aim of this study, the frequent academic collocations used in the corpora were identified and compared in order to prepare a list of the collocations and their different types based on the Benson et al's (1986) classification of English collocations. For the purpose of investigating the frequency of collocational expressions, this study utilized AntConc/kfngram software.

AntConc/kfngram software is a free downloadable software program which is user-friendly and can analyze a long text/ multiple texts, report frequency, and determine multi-text occurrence. The researcher began the analysis by skimming each discussion section to get the feel of the overall organization. A Chi Square data analysis was used to report the existence of any significant difference between two groups of authors in using collocations. To ensure the reliability of the analysis in the process of data categorization, 30% of the data was rechecked and reanalyzed independently for different types of collocations based on the framework under study by a second researcher who had taught discourse analysis for years at university and was familiar with the data analysis.

The second rater coded 30% of the data, taken randomly from the corpus and finally, the inter-rater reliability was estimated and reported. The inter-rater agreement, measured using Cohen's Kappa formula, was found to be Kappa = 0.818, $p = 0.000$.

## RESULTS

As already stated, collocations were classified into lexical and grammatical ones based on Benson et al.'s (1986) classification of English collocations. For the purpose of investigating

the frequency of collocational expressions, this study utilized AntConc/kfngram software to answer the research question. Table 2 shows macro-structures of lexical collocations in the discussion sections of Ph.D. dissertations written by the participant native and non-native ELT students, namely, EFL students, ESL students, and native students

**Table 2**

*Macro-structures of Lexical Collocations in the Corpus*

| Lexical collocations | EFL | ESL | Natives |
|---|---|---|---|
| *V* | 31065 (35.3%) | 35355 (29.5%) | 27297 (22.4%) |
| *N* | 40310 (44%) | 61262 (51.2%) | 72266 (59.6%) |
| *Adj* | 14185 (15.6%) | 17759 (14.8%) | 19251 (15.9%) |
| *Adv* | 4950 (5.1%) | 5444 (4.5%) | 2616 (2.1 %) |
| *Total* | 90505 | 119820 | 121430 |

Based on the results of data analysis, in the corpora, the function words such as "the, of, and, in, to, a, as, for" were used in high frequency. As Table 1 shows, concerning the discussions written by the Iranian EFL Ph.D. s, from the total word count of 90505 in the discussion sections and from four instances of structures, the frequency of noun-based lexical collocations as a highly frequent structure was 40310 (44%), verb-based collocations made up about 35.3% of the collocations and adverb-based lexical collocations with the rate of 5.1% were the least frequent collocations after adjective-based ones with the rate of 15.6%. Very similar to Iranian Ph.D. s at an EFL context, the Indian post-graduate students (ESL context) used more noun-based collocations (51.2 %) and less adverb-based lexical collocations with the rate of 4.5%. Finally, the native speakers of English with higher collocation types (121430 instances) used more noun-based collocations (59.6%) and less adverb-based lexical collocations with the rate of 2.1%. Figure 1 clarifies the results.
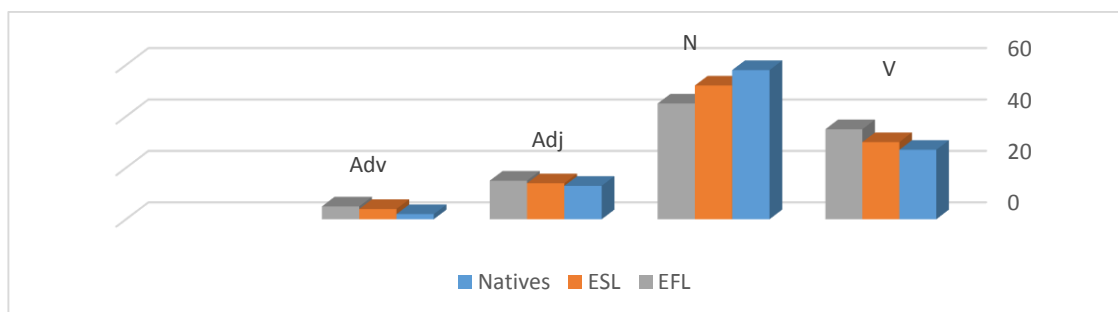


**Figure 1**

*Macro structure of Lexical Collocations by Groups*

The same procedure was conducted for obtaining the rate of grammatical collocations and their subcategories that are varied from G1 to G8. Table 3 indicates the results of the grammatical collocations used in the corpora.

**Table 3**

*Micro-structures of Grammatical Collocations in the Corpus*

| Type | EFL | ESL | Natives |
|---|---|---|---|
| G1 noun + preposition | 17195 (19%) | 17973 (15%) | 21857 (18%) |
| G2 noun + to + infinitive | 12670 (14%) | 14378 (12%) | 27297 (15%) |
| G3 noun + that-clause | 10860 (12%) | 16774 (14%) | 15758 (13%) |
| G4 preposition + noun | 5430 (6%) | 10783 (9%) | 13357 (11%) |
| G5 adjective + preposition | 6335 (7%) | 5991 (5%) | 3642 (3%) |
| G6 adjective + to + infinitive | 8145 (9%) | 8387 (7%) | 12143 (10%) |
| G7 adjective + that-clause | 1810 (2%) | 3594 (3%) | 6071 (5%) |
| G8 19 English verb patterns | 28060 (31%) | 41940 (35%) | 27376 (25%) |
| Total | 90505 | 119820 | 121430 |

As Table 3 indicates, concerning the discussions written by Iranian EFL Ph.D. s, from the total word count of 90505 in the discussion sections and from eight instances of the grammatical collocations, the frequency of G8 was 28060 (31%) and the frequency of G1 that took the second place was 17195 with the rate of 19%. Also, the rates of G4 to G7 were less than 10%, with the G2 (14%), which took the third place after G1, and G3 in the fourth place. Surprisingly, the same patterns generated from the analysis of two corpora of Indian Ph.D. dissertations and native ones with the high frequency of G8 and the low frequencies of G4 to G7, but there was a difference in the order of the use of G3 and G2 between the EFL and ESL contexts in that in ESL context G3 (19%) took the third place. The results concerning the natives' dissertations revealed that only the rates of G5 and G7 were less than 10%. However, all these results led us to this general conclusion that three groups obeyed the standards of academic writing in terms of grammatical collocation use. Figure 2 visualizes the results of the grammatical collocations' sub-categories
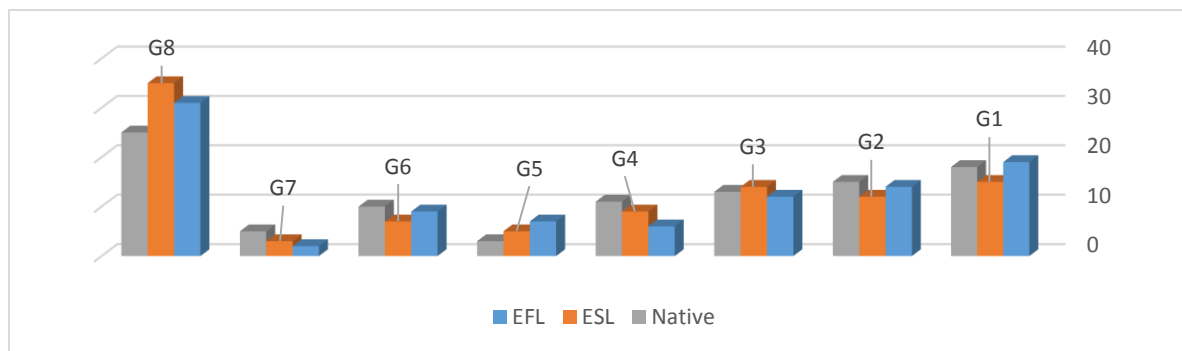


**Figure 2**
*Micro structure of Grammatical Collocations by Group*

Figure 2 shows that Ph.D. students in different contexts had the same patterns in using grammatical collocations' subcategories, in total. For example, all students, regardless of the context where they grew up, used more G8 and less G5s and G7s. In other words, Ph.D. students' frequencies of grammatical collocations' subcategories did not vary on the basis of the students' context or first language. Furthermore, based on the results of the data analysis, all three groups did not use verb+ adverb in their writings and this sub-category was absent in the data.

In order to investigate the existence of any significant differences among three groups of Ph.D. students in using lexical and grammatical sub-categories, Chi-square tests were run. The first chi-square analysis in terms of lexical sub-classifications shows that there is a significant difference among groups. $X2(2, N= 331,755) = 630.0858, P< .00001$. The second Chi-square test was run to investigate the existence of any statistical significant differences among three groups of Ph.D. students in using grammatical sub-categories the results of which revealed significant differences among three groups, $X2(2, N= 331,755) = 637.7686, P< .00001$.

As the rates and distributions of collocations in general were the highest among the native speakers' Ph.D. dissertations (121430 instances) in comparison to Indian ESL Ph.D. dissertation (119820 instances), and those of Iranian EFL Ph.D. dissertations (90505 instances), it can be concluded that natives acted differently in using grammatical collocations in comparison to the other two groups.

**DISCUSSION**

The research presented in this article focused on the use of academic collocations in the discussion sections of the Ph.D. dissertations written by EFL, ESL, and native students. The results of the data analyses revealed that the total use of collocations by the EFL Ph.D. s was 90505, the ESL Ph.D. s used 119820 collocations as a whole, and the native speakers of English had the highest frequency in using academic collocations that was 121430. Based on the results concerning lexical collocations,

the use of noun-based lexical collocations by the EFL Ph.D. s had the highest frequency, and the adverb-based ones were the least frequent collocations after the adjective-based and verb-based ones. Very similar to Iranian Ph.D. s at an EFL context, Indian postgraduate students (ESL context) and native speakers of English used more noun-based collocations and less adverb-based ones, while the rates of using all these sub-categories of lexical collocations were different. The rates related to the EFL context were 44% for noun-based, 35.3% for verb-based, 15.6% for adjective-based-, and 5.1% for adverb-based collocations. In the ESL context, the rates were 51.2% for noun-based-, 29.5% for verb-based, 14.8% for adjective-based, and 4.5% for adverb-based collocations. The natives' rates were 59.6% for noun-based collocations, 22.4% for verb-based collocations, 15.9% for adjective-based collocations, and 2.1% for adverb-based collocations. In comparison to the other forms of lexical collocations, these data demonstrated a high utilization of the noun sub-category. The plethora of the use of noun phrases in this study could be due to the fact that, in general, written texts have longer, more complex words and phrases, more nominalizations, more noun-based phrases, and more lexical variations. Writers use more expanded noun phrases to make their writing more descriptive. They aid in the reader's mental picture of what the writer is attempting to describe, and academic writing, in particular, is informational in nature, necessitating the use of more noun phrases (Halliday & Hasan, 1989; Liu et al., 2018). This conclusion is in line with published rates of formulaic sequences such as collocations in full-length research publications, where they were found to make up at least half of the text (Biber et al., 1999; Biber & Conrad, 1999; Cortes, 2004; Hyland, 2008; Salazar, 2014). The results concerning grammatical collocations showed that all three groups (EFL, ESL, and natives) had the highest frequency of using 19 English verb patterns (G8), and that the use of noun+ preposition (G1) took the second place, and the use of noun+ to+ infinitive (G2) took the third place except for the ESL context in which G2 took the fourth place after G3 (noun+ that-

clause). The rates of using other sub-categories were less than 12%. All these results revealed that there was a statistically significant difference among three groups of writers in using both lexical and grammatical academic collocations and that native speakers of English used more academic collocations in comparison with ESL students in India where English is their formal and instructional communication language. Additionally, ESL students were observed to use more collocations than Iranian EFL students. In total, there was a statistically significant difference in the use of collocations among three groups, with native speakers having the most instances with 121430 collocations. These results are in line with Conklin and Schmitt's (2012) conclusion that formulaic sequences make up at least 30-50% of language overall. In general, the results showed the importance of the formulaic sequences such as collocations, lexical bundles and idioms in academic texts such as research articles and dissertations. There is also the other face to the coin in which Cobb (2000) claims that all collocations are of an arbitrary nature and there is no logic underlying them. His viewpoint contradicts the findings of the current study, as he believes that the arbitrary nature of collocation poses a challenge for non-native English speakers. Lewis (2000) agrees with this viewpoint, emphasizing that in English, the unacceptability of particular combinations is based on convention rather than consistency in individual item meanings. Learners who aren't aware of these rules may come up with undesirable combinations. Subjects' unfamiliarity with English collocations as a result of insufficient exposure is another source of difficulty. The more a learner encounters a particular form of collocation, the better they will understand and apply it. Tajalli (1994) maintains that exposure or lack of exposure to a certain type of collocation influences the learning of that kind of collocation. The first language transfer can also be a factor. Misuse of a target language sequence, overuse, underuse, and use of learner-idiosyncratic combinations, according to De Cock (2004), are four forms of L1 transfer, notably among non-native English speakers.

There is also the possibility of another component, namely culture. The sense of culture is presented in bulk of studies in the area of contrastive rhetoric where the rhetorical choices established by authors will be affected by cultural norms, values, and belief systems used in particular sociocultural contexts (e.g., Doró, 2014). This means that the type of culture and context in which the author raised affect the way of writing. All these can provide some justifications for why Iranian EFL learners used collocations the least as compared to Indian and native English authors.

Natives utilized more collocations in the discussion section of their Ph.D. dissertations, according to the current study. ESL and EFL students, on the other hand, used collocations at nearly the same rate. The probable explanation for the extensive use of collocations by natives stands for their high familiarity with the phrases, formulaic sequences, and collocation, which is a positive aspect for natives The claim regarding the familiarity of natives with collocations is in harmony with Hyland's (2002) statement that asserted "academic writing is not a single undifferentiated mass, but a variety of subject-specific literacies" (p.352).

Although native students used academic collocations the most, the same rate and frequency of employing particular sub-classifications of academic collocations in both EFL and ESL students in compared to native speakers can be significant evidence for standard academic writing in both EFL and ESL students. By standard academic writing, we mean the familiarity of non-natives to the rules of academic writing in comparison with the natives. However; based on the findings of this study it is suggested that particular attention be paid to the teaching of all types of collocations because of the learners' general weaknesses in using different types of collocations in the early stages of language learning.

**CONCLUSION**

The primary purpose of this research was to examine the rates and frequencies of using academic collocations and their different sub-classifications in the discussion section of the Ph.D. dissertations written by natives, Indian ESL, and Iranian EFL Ph.D. students in the field of English Language Teaching within the period from 2005 to 2015. The results showed that factors like being in an EFL context, being native speaker of a language like English, and being in an ESL context affect the use of academic collocations in the discussion sections of Ph.D. dissertations. Furthermore, the factors like the arbitrary nature of collocation, the unfamiliarity of subjects with English collocations due to insufficient exposure, first language transfer, and culture may change the result of this study. The findings could provide some insight into the conventions of academic collocations and how they are actually utilized by post-graduate students as future researchers. They could also heighten the general awareness on the proper use of academic collocations and more importantly on the various possible sub-categories to increase their attention toward rhetorical aspect of academic writing. Using appropriate academic collocations can help students publish their papers in prestigious journals. In sum, it is suggested that teachers devise exercises for EFL students that boost the involvement of learners in the process of recognition and production of grammatical collocations, especially in the academic writing.

Despite its limitations which include a small sample size due to the difficulties of data collecting this study could be extremely useful in improving student writing and research behaviors when it comes to making their own voices known through discussion of the findings in relation to other studies this is especially true. Teachers are recommended to teach academic collocations and their use in academic writing during MA/MS programs, and make the students aware of and familiar with different purposes and rhetorical functions of academic collocations. The students can be given writing assignments, enough practice, and experience in employing academic collocations for different text types and occasions with different purposes.

## References

Abdollahpour, Z., & Gholami, J. (2018). Building blocks of medical abstracts: frequency, functions, and structures of lexical bundles. *Asian ESP Journal*, *14*(1), 82-110.

AlHassan, L., & Wood, D. (2015). The effectiveness of focused instruction of formulaic sequences in augmenting L2 learners' academic writing skills: A quantitative research study. *Journal of English for Academic Purposes*, *17*, 51-62.

Benson, M., Benson, E., & Ilson, R. (1986). Lexicographic description of English (studies in language companion, v. 14). *Amsterdam: J. Benjamins Publishing Company.*

Bian, X., & Wang, X. (2016). Chinese EFL Undergraduate's Academic Writing: Rhetorical Difficulties and Suggestions. *Indonesian Journal of Applied Linguistics*, *6*(1), 20-29.

Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, *5*(2), 97-116.

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

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Longman grammar of written and spoken English. *Harlow: Longman.*

Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language learning & technology*, *14*(2), 30-49.

Cobb, E. (2000). Testing EFL vocabulary. *Oxford, UK: Pergamon Press.*

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, *32*, 45.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: examples from history and biology. *English for Specific Purposes, 23*, 391-423.

Cortes, V., & Csomay, E. (2015). Corpus-based research in applied linguistics: Studies in honor of Doug Biber (Vol. 66). *Amsterdam: John Benjamins Publishing Company*

Cotos, E., Huffman, S., & Link, S. (2015). Furthering and applying move/step constructs: Technology-driven marshalling of Swalesian genre theory for EAP pedagogy. *Journal of English for Academic Purposes*, *19*, 52-72.

Cowie, A. P., & Howarth, P. (1996). Phraseological competence and written proficiency. *British studies in applied linguistics*, *11*, 80-93.

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures (BELL)*, *2*(1), 225-246.

Doró, K. (2014). Citation practices in EFL undergraduate theses: a focus on reporting verbs. *UPRT 2013: Empirical Studies in English Applied Linguistics*, 32.

Durrant, P. & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International review of applied linguistics 47*, 157-77.

Ebeling, S. O., & Hasselgård, H. (2015). Learners' and native speakers' use of recurrent word-combinations across disciplines. *Bergen Language and Linguistics Studies*, *Vol. 6.* 87-106

Ellis, R. (2012). Language teaching research and language pedagogy. *The UK: John Wiley & Sons, Inc.*

Ellis, N., & Simpson-Vlach, R. (2009). Formulaic language in native speakers: Triangulating psycholinguistics, corpus linguistics, and education. *Walter de Gruyter GmbH & Co. KG 5(1),* 61-78.

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text & Talk. 20(1).* 29-62.

Evans, V., & Tyler A. (2003). The semantics of English prepositions: Spatial scenes, embodied meaning, and cognition. *Cambridge University Press.*

Firth, J. R. (1957). A synopsis of linguistic theory, 1930-1955. *Studies in linguistic analysis.*

Gholami, L., & Gholami, J. (2020). Uptake in incidental focus-on-form episodes concerning formulaic language in advanced adult EFL classes. *Language Teaching Research*, *24*(2), 189-219.

Gläser, R. (1998). The stylistic potential of phraseological units in the light of genre analysis. *Phraseology: Theory, analysis, and applications*, Oxford: OUP. 125-143.

Halliday, M.A.K. & Hasan, R. (1989). Language, Context, and Text: Aspects of Language in a Social-Semiotic Perspective. *Oxford: Oxford University Press.*

Herman, C. (2017). Looking back at doctoral education in South Africa. *Studies in Higher Education.* 42(8), 1437-1454.

Hung, M. & Chin, Y. (2018). The use of corpus and a collocation framework in the comparison of the English tests of two major college rntrance examinations in Taiwan. *Asian Journal of Language, Literature and Culture Studies,* 1-11.

Hyland, K. (2000). Disciplinary discourses: Social interactions in academic writing. *London: Longman.*

Hyland, K. (2001). Bringing in the reader: Addressee features in academic articles. *Written communication*, *18*(4), 549-574.

Hyland, K. (2002). Directives: Argument and engagement in academic writing. *Applied Linguistics, Volume 23, Issue 2.* Pages 215- 239

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for specific purposes*, *27*(1), 4-21.

Hyland, K. (2012). Disciplinary identities: Individuality and community in academic discourse. *Cambridge: Cambridge university press. Ernst Klett Sprachen.*

Lewis, M. (2000). Teaching collocation: Further developments in the lexical approach. *Hampshire: Heinle Cengage Learning.*

Liu, Y. & Buckingham, L. (2018). The schematic structure of discussion sections in applied linguistics and the distribution of meta-discourse markers. *Journal of English for Academic Purposes, 34,* 97- 109

Liu, H., Huang, J., Pan, Y., & Zhang, J. (2018). Barycentric interpolation collocation methods for solving linear and nonlinear high-dimensional Fredholm integral equations. *Journal of Computational and Applied mathematics*, *327*, 141-154.

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied linguistics*, *33* (3), 299-320.

McCarthy, P. & Wient, M. (2019). Who are the top Ph.D. employers. Advancing Australia's knowledge economy. *Melbourne: The University of Melbourne.*

McIntyre, D. (2015). Towards an integrated corpus stylistics. *Topics in Linguistics 16(1),* 59-68

Nagao, M., & Mori, S. (1994). A new method of n-gram statistics for large number of n and automatic extraction of words and phrases from large text data of Japanese. In *COLING 1994 Volume 1: The 15th International Conference on Computational Linguistics.*

Pho, P. D. (2008). Research article abstracts in applied linguistics and educational technology: A study of linguistic realizations of rhetorical structure and authorial stance. *Discourse studies*, *10*(2), 231-250.

Salazar, D. (2014). Lexical Bundles in Native and Non-native scientific writing. *Amsterdam/ Philadelphia: John Benjamins.*

Samraj, B. (2005). An exploration of a genre set: Research article abstracts and introductions in two disciplines. *English for specific purposes*, *24*(2), 141-156.

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

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

Shin, D., & Chon, Y. V. (2019). A multiword unit analysis: COCA multiword unit list 20 and collogram. *Journal of Asia TEFL*, *16*(2). 608.

Smadja, F. (1991). From n-grams to collocations an evaluation of Xtract. In *29th Annual Meeting of the Association for Computational Linguistics.* 279-284.

Tajalli, G. (1994). Translatability of English and Persian collocations. Paper presented at the second conference on translation. *Tabriz University, Tabriz.*

Wray, A. (2008). Formulaic language: Pushing the boundaries. *Oxford Applied Linguistics, Oxford: Oxford University Press.*

Wray, A. (2012). Formulaic sequences. *The encyclopedia of applied linguistics*. *1-5. Wiley Online Library.*

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication*, *20*(1), 1-28.

Xiao, R., & McEnery, T. (2006). Collocation, semantic prosody, and near synonymy: A cross-linguistic perspective. *Applied linguistics*, *27*(1), 103-129.

**Biodata**

**Vida Fathi Bonabi** is a Ph.D. candidate of TEFL at Islamic Azad University, Tabriz Branch, Iran. She is a faculty member at Islamic Azad University, Urmia Branch and head of the department too. She has been teaching most of the courses related to English teaching and translation majors. Her main areas of interest include discourse analysis, methods and techniques of language teaching, linguistics, materials design and syntax.
Email: *vidafathi90@gmail.com*

**Dr. Nesa Nabifar** is an assistant professor at Islamic Azad University, Tabriz Branch, Iran. She holds Ph.D. in Linguistics and applied linguistics. Her areas of research are second language acquisition, first language acquisition and discourse analysis. She teaches linguistics, academic writing and language teaching courses at graduate and post graduate levels. She has published a lot of papers and has attended different international and national conferences.
Email: *n.nabifar@iaut.ac.ir*

**Dr. Saeideh Ahangari** is an associate professor of TEFL at Islamic Azad University, Tabriz Branch, Iran. Her areas of interest include language teaching methodology, research, psycholinguistics and assessment. She has been teaching courses related to language teaching and testing at the graduate and postgraduate levels for 25 years. She has published and presented many papers during these years.
Email: *Ahangari@iaut.ac.ir*
*saeideh.ahangari@gmail.com*