Native and Non-native Use of Lexical Bundles in Discussion Section of Political Science Articles

Mohammad Meisam Safarzadeh
PhD Candidate, AllamehTabataba'i University
MM.Safarzadeh@yahoo.com

Abbas Monfared
PhD Candidate, Allameh Tabataba'i University
a_monfared85@yahoo.com

Mohamad Sarfeju
M.A, Payam Noor University
m.sarfeju@gmail.com

Abstract

The study of lexical bundles, among types of text analysis, is gaining importance over the others in the last century. The present study employed a frequency-based analysis approach to the use of lexical bundles. The discussion section of 60 political science articles, with corpora around 253,063 words were investigated in three aspects of structure, form, and function of lexical bundles. The present study selected its data pool out of scholarly articles from qualified journals in the field of political sciences. One part of the data pool was made up of 30 articles written by American native speakers. The second half of the data comprised the 30 articles written by Iranian scholars in political sciences. The findings showed that native and Persian-speaking writers employed the same forms of lexical bundles, and there were significant differences concerning the nativeness and functions. Bearing in mind the findings of the present study, material developers would think of the possibility of the addition of lexical clusters into the materials. It can similarly be valuable for the development of the second language writing strategies, for those who need to write in academic contexts especially political contexts.

Keywords: Discourse Analysis, Lexical Bundles, ESP, Nativism, Non-nativism

Received: March 2012; Accepted: December 2012

1. Introduction

The literature of applied linguistic over the last two decades has seen numerous pieces of study around discoursal characteristics of texts. This group studies might directly or indirectly contribute to the field of ESP (e.g., Grabe & Kaplan, 1999; cited in Biber, Conrad & Cortex, 2004). That group of studies which made use of corpus data somewhat paved the way for the tenable recognition of lexical chunks (Altenberg, 1998; Wray, 2002; cited in Chen & Baker, 2010).

To deal with the complex nature and to put emphasis on the significance of lexical units, different approaches, methods, and criteria have been employed (Biber et al., 2004). Moreover, scholars are attaching more importance to the study of lexical units over other types of discourse analyses (Barber, 1962; Dudley-Evans & ST John, 1988; Ewer & Huges-Davies, 1971). While the number of studies demonstrates the significance of multi-word units, there is little consensus on the definition and description of the characteristics, identification methodologies, and even fixed expressions for calling them (Biber et al., 2004). Thus, there seems to be no single comprehensive approach which can comprehensively capture the whole.

As mentioned above, the findings of these groups of studies can positively and effectively inform the fields of ESP in general and EAP in particular (Byrd & Coxhead, 2010). The main focus in ESP is on language in context, rather than on decontextualized grammar teaching and teaching of language mechanics (Fiorito, 2012). Academic writing in a particular field entails knowledge of the field and having control over the conventions on which the discourse of the field is established if one is willing to have voice and to be effective in capturing and manipulation of the available linguistic features based on rhetorical requirements of a particular discourse community. Bearing

in mind the fact that there is no consensus on what should be done to see the behavior of lexical bundles, this study aimed to investigate these units from form, structure and function points of view. To these ends, the present study was conducted to explore whether Persian-speaking writers use lexical bundles in the same sequential patterns (forms), frequency, and function as native speakers typically do in their written texts.

2. Review of Literature

2.1. Word Combinations Research

The study of fixed/multi-word expressions has received much attention on the part of the researchers for a long time. Different researchers apparently followed their own rubrics in naming these expressions, some names are more frequently used including 'lexical phrases', 'formulas', 'routines', 'fixed expressions', 'prefabricated patterns', and 'lexical bundles' (Biber et al., 2004, p. 372).

These fixed expressions (Moon, 1998) have the potentiality to define a type of criteria to define a threshold level to distinguish native speakers from non-native ones as Haswell (1991, p. 236) argues "the absence of such clusters might reveal the lack of fluency of a novice or newcomers to that community". The more frequent use of lexical items signals the more competent language use within a register (Cortes, 2004). In effect, competent user's preferences for certain word clusters over others demand a kind of sensitivity on the part of a novice user to gain control of a new register (Hyland, 2008). Moreover, most of these expressions or word sequences can be patterned in the form of fixed expressions since 80% of natural language could be formulated in this way (Altenberg, 1998). It is also argued that "most everyday words do not have an independent meaning, or meanings, but are components of a rich repertoire of

multi-word patterns that make up a text" (Sinclair, 1991, p. 108). Pawley and Syder (1983) also emphasized the significance of fixed phrases and put forward the specific discourse functions they perform, which are thought to play a significant role in fluent language production, particularly spoken language. Besides, it is probably related to some degree of maturity or competency in production as writers increasingly rely on collocations or word sequences and the less use of these fixed expressions may be the characteristic behavior of novice writers (Haswell, 1991). It is interesting that there are researchers who consider that the contribution of the pragmatic use of a word to a sense of coherence can also be captured in the application of these lexical patterns; thus, lexical phrases can consist of one word to many words (Byrd & Coxhead, 2010).

2.2. Form and Structure

While some researchers are willing to use their own terms once they talk about the morphological characteristics of lexical bundles, one is increasingly expected to see the employment of form and structure criteria more than others. But, disagreement in the application of alternative terms for lexical bundles or phrases might indicate a certain degree of terminological misperception. Clusters, recurrent word combinations, lexical phrases, phrasicon, n-grams, bundles, and recurrent word strings are among the terms frequently and interchangeably applied by scholars in this field of study (Chen & Baker, 2008).

According to Biber et al. (1999), most studies on the frequent word combinations apply the structural classification of lexical expressions in the Longman Grammar of Spoken and Written English (Biber et al., 1999). They suggested fourteen groups of lexical bundles for conversation, and twelve

groups for academic prose; however, commonalities exist across groups. Chen and Baker (2010) differentiated three general structural categories of "NP-based", "PP-based", and "VP-based", the NP-based referring to any noun phrases containing post-modifier fragments such as 'the aim of', the PP-based referring to those beginning with preposition followed by a noun-phrase fragment like 'within the realm of', and VP-based be referring to any combinations containing a verb constituent such as 'have/has to do with'. Biber et al. (2004) argued that most lexical bundles do not act as a whole structural unit; instead, they connect two structural units although the manner they make connections among discourse contexts, genres, and registers is various, prevalent use of lexical bundles certainly is acknowledged in various studies (Biber & Barbieri, 2007; Wray, 2000; Wray & Perkins, 2000). In academic writing, most bundles are included in noun or prepositional phrases (Hyland, 2008a). Hyland (2008a) also argued that the noun phrase containing of-phrase constituents is the most common structure in academic genres as it is also noted by Byrd and Coxhead (2010) that academic prose is thought to be 'noun-centric'. Moreover, they acknowledged the coincidence of results with Hyland (2008b) in the fact that passive bundles are one of the characteristic features of scientific writing. Furthermore, Chen and Baker (2010) stressed the different structural properties of lexical bundles comparing conversations and the academic prose, the former being clausal and the latter being phrasal.

Biber et al. (2004) recognized three main structural categories of lexical bundles. Type one involves bundles containing verb phrase fragments.

Table 1. Structural Types of Lexical Bundles (Adopted from Biber et al., 2004, p. 381)

Lexical bundles incorporated in verb phrase fragments
1a. (connector +) 1st/2nd person pronoun + VP fragment:
e.g., you don't have to, I'm not going to, and well I don't know
1b. (connector +) 3rd person pronoun + VP fragment:
e.g., it's going to be, that's one of the, and this is a
1c. discourse marker + VP fragment:
e.g., I mean you know, you know it was, I mean I don't
1d. verb phrase with active verb:
e.g., is going to be, is one of the, have a lot of, take a look at
1e. Verb phrase with passive verb:
e.g., is based on the, can be used to, shown in figure N
1f. yes-no question fragments:
e.g., are you going to, do you want to, does that make sense
1g. WH-question fragments:
e.g., what do you think, how many of you, what does that mean

Type two comprises bundles which make use of dependent clause fragments along with simple verb fragments.

Table 2. Structural Types of Lexical Bundles (Adopted from Biber et al., 2004, p. 381)

Lexical bundles incorporated in dependent clause fragments
2a. 1st/2nd person pronoun + dependent clause fragment:
e.g., I want you to; I don't know if, you might want to
2b. WH-clause fragments:
e.g., what I want to, what's going to happen, when we get to
2c. If-clause fragments:
e.g., if you want to, if you have a, if we look at
2d. (verb/adjective+) to-clause fragment:
e.g., to be able to, to come up with, and want to do is
2e. That-clause fragments:
e.g., that there is a, that I want to, that this is a

Type three contains bundles which are phrasal in structure such as prepositional phrases.

Table 3. Structural Types of Lexical Bundles (Adopted from Biber et al., 2004, p. 381)

Lexical bundles incorporated in noun phrase and prepositional phrase fragments

3a. (connector +) Noun phrase with of-phrase fragment:
e.g., one of the things, the end of the, a little bit of
3b. Noun phrase with other post-modifier fragment:
e.g., a little bit about, those of you who, the way in which
3c. other noun phrase expressions:
e.g., a little bit more or something like that
3d. Prepositional phrase expressions:
e.g., of the things that, at the end of, at the same time
3e. Comparative expressions:
e.g., as far as the, greater than or equal, as well as the

Considering the form rubric, researchers try to describe bundles in terms of the length of bundle unit or the number of constituents, e.g., how many words should be counted as one bundle unit while most of the time bundles of a shorter length are subsumed under the longer ones. In order to make a corpora manageable, to avoid idiosyncrasies, and to provide a condition for a concordance tool to have precise checks, most researchers (Biber et al., 2004; Chen & Baker, 2010; Cortes, 2004; & Hyland, 2008a) agreed on a four-word unit of lexical bundle unit.

2.3. Function

Many of current studies have viewed and analyzed bundles by taking into account two criteria namely structure and function (i.e., Biber & Barbieri, 2007; Biber et al., 1999; Biber et al., 2004; Chen & Baker, 2010; Cortes, 2004;

Eisenmann, Wagner & Cortes, 2008; Hyland, 2004; Hyland, 2008). Functionally, particular significant features are accredited to the occurrence of lexical bundles. They typically have particular discourse functions concerning stance, discourse organization, or referential framing; and they are significantly more common in spoken discourse than in written discourse. In thespoken discourse, interlocutors have to meet the requirements of any discourse situation through performing varieties of functions, not essentially in simultaneity (Biber et al., 1999).

2.3.1. Lexical Bundles Functions

1. Classic Model:

Comprising larger corpora containing casual conversations, textbooks, course packs, service encounters, institutional texts, and so on, Biber's model (2004) has gained the attention of the most researchers. In this taxonomy, he made a distinction among three main categories: stance expressions, discourse organizers, and referential expressions.

I. *Stance expressions*

Stance expressions offer a formula for the writer to interpret next proposition so as to translate two types of meaning containing epistemic and attitude/modality. Moreover, it should be taken into account that stance bundles can be either personal or impersonal.

II. *Discourse organizers*

Discourse organizing bundles are used by speakers and writers in order to introduce a topic, put emphasis on a topic, and to elaborate or to make a topic clear.

III. *Referential bundles*

This group comprises identification of entities or single-out of particular attributes of an entity in order to have significance in comparison to the other entities or other attributes of an entity. This category is further divided into four subgroups of identification or focus, imprecision indicators, specification of attributes, and time, place or text reference.

2. Alternative models and approaches

Employing an inductive approach, Hyland (2008) revised the classic model to group bundles model, the one meeting the requirements of post-modern principles introduced by Kumaravadivelu (2001) since it is less fixed and welcomes varieties. He introduced three broad categories including research-oriented, text-oriented, and participant-oriented categories.

Byrd and Coxhead (2010) made an effort to possibly terminate the apparent terminological confusion by defining lexical bundles. They introduced three broad groups containing presentation of content, organization of discourse or text, and expressions of attitudes. They could not attain a comprehensive analysis of bundles and they made a case for it through the argument that limiting the analysis is an effort to provide a system that teachers probably find more directly appropriate and applicable to EAP instruction.

2.3.2. Structural and Functional Categories

Regardless of what the structure is involved in the formation of any bundle unit, within one unit numbers of functions are to be performed (Biber & Conrad, 2006). Accompanied by the claim that most academic functions are to be carried out in the form of bundles containing nouns and prepositional phrases, Hyland (2008a) detected the various uses of bundles in the discipline

of electrical engineering with 213 four-word clusters detected 20 times per million words while discipline of biology had the least range of application. Similarly, specialized readership, i.e., talking to narrow members of a discourse community, makes it strange for writers to use bundles in a different way of other disciplines.

Stance bundles are typically composed of dependent clause constituents, and referential bundles are made of noun phrase or prepositional phrase constituents; remarkably, discourse organizing bundles can nearly use all three structural categories (Biber et al., 2004).

The above instances establish the direct connection concerning structural categories and discourse functions. Certain discourse functions are distinguishing attributes of a particular discipline, while articulated in different structures overlapping across different disciplines (Cortes, 2004).

2.4. Lexical Bundles' Operationalization

In order to operationalize any research variables, the researcher should attempt to find the responses to two basic questions: first, how it is defined and second, how it is measured (Brown, 2003). To operationalize a group of lexical bundles, scholars first defined lexical bundles definitely and elaborately (e.g., Biber et al., 1999; Hyland, 2008; Cortes, 2004). In order to define one unit of lexical bundle, distinguishing attributes should be taken into consideration.

The first criterion to consider is the cut-off frequency, which determines how many bundle units should be incorporated for further analysis (Chen & Baker, 2010). The typically established frequency threshold for large written corpora ranges from 20 to 40 per million words (e.g., Biber et al., 2004; Cortes, 2004) while some argue that defining frequency threshold to 20 times per million words is to be conservative (Hyland, 2008a). Chen and Baker (2010)

used the frequency and occurrence threshold to four-word lexical bundles occurring 25 times per million words among at least three different texts. They thought that standardized frequency should be converted into raw frequency since standardized frequency is unable to find 'expected impartiality. The second criterion to consider is the recognition of a bundle unit across different texts. One lexical bundle unit should occur cross-textually in at least 3-5 different texts (Biber et al., 1999; Biber et al., 2004; Cortes, 2004; Chen & Baker, 2010) or at least in 10% of the whole corpora to avoid idiosyncrasies from different writers or speakers (Hyland, 2008a). The third and final criterion accounts for the length of the frequent word clusters ranging from 2 to 6 word units (e.g., Barber & Barbieri, 2007). Bearing in mind these criteria, the researcher should consider one unit of lexical bundle as a unit of word strings nearly from 2 to 6 word units with the frequency of occurrence of 20 times per million words across at least 3 different contexts or 10% of the whole texts.

3. Research Questions

As there is no consensus on what should be done to see the lexical bundles' behavior, these units should be studied from a different prospective. Moreover, form, structure and function are of much concern to the present study. Then, this study set out to investigate if Persian-speaking writers use lexical bundles in the same sequence, frequency, and function as native speakers (NSs) normally do in one specific section of Discussion section of Political Sciences (DPS) articles, and provide answers to the following questions:

1. To what extent Persian-speaking writers use lexical bundles in similar sequential patterns (Form) in DPS articles as NSs do?
2. Do Persian speaking writers use lexical bundles as frequently as NSs do in DPS articles?

4. Method

4.1. Corpus of the Study

The present study selected its data pool out of scholarly articles from qualified journals in the field of political sciences. One group of the articles of the corpora is made up of 30 ones written by American or British native speakers. Most articles were from *American Journal of Political Science* and *American Political Science Review*. The second half of the articles comprised of the 30 articles written by Iranian scholars in the field of political sciences, mostly from *International Studies Quarterly* and *Middle East Studies* journals. Table 4 illustrates the distribution of articles across the publication sources.

Table 4. Corpora Articles across Journals

Corpus	Journal	No. of articles
Native writers	American Journal of Political Science	24
	American Political Science Review	4
	The Journal of Politics	2
Persian speaking writers	Middle East Studies	17
	International Studies Quarterly	6
	Third World Quarterly	2
	Asian Survey	3
	Political Research Quarterly	2
	Total	60

As the present study incorporated discussion section of the selected papers, the convenience of adaptable texts was of crucial concern. Most officially published articles do not allow for texts manipulation; thus, it was of crucial concern to apply Google OCR in order to identify texts and make them modifiable. Table 5 shows the number of articles, concerning the nativity variable in corpora.

Table 5. Corpora Concerning Nativity

Corpus	No. of texts	No. of words
Native speaking writers	30	128,452
Persian speaking writers	30	124,611
Total	60	253.063

4.2. Lexical Bundles Identification

To recognize lexical bundles as rubrics of "recurrent word units", the frequent occurrences of a unit should be taken into consideration in order to reach the confidence to know one unit as a lexical bundle (Biber et al., 2004). To detect the lexical strings and to attain the frequency of occurrence of each one, the corpus was entered into the corpus analysis software AntConc. AntConc is a corpus analysis software application, which has met regular revisions, and the last version of this software is available for free for different operation systems. This latest edition of the software offers additional features which make it applicable for a variety of corpus-based analyses.

5. Results

The purposes of present study was to describe lexical bundles in terms of form, structure, and function, to identify these features, and to compare the use of lexical bundles across native and non-native writers in one particular section of a specific genre. Taking a frequency-based approach, lexical units were identified, and lists of lexical bundles were captured following the above mentioned features. Once the texts have been processed, the software generated lexical bundle units descending form the most to the least frequent. Native writers' corpus yielded 178 lexical clusters out of 128,452 words, and

Persian speaking writers' corpus yielded 132 lexical clusters out of 124,611words.

5.1. The Use of Lexical Bundles: Frequency

Once lists of bundles were captured, Chi-square analysis was conducted; the cells were recognized; the frequency of use for each was provided; the expected values were given; and the residual of these bundles were illustrated. Table 8 displays the frequency of occurrence of lexical bundles as the observed number, the expected number and the standard frequency which is recognized as residual.

Table 6. Lexical Bundles Frequency

	Observed N	Expected N	Residual
Native	178	155.0	22.0
Iranian	132	155.0	-22.0
Total	310		

Table 9 illustrations the results of Chi-square analysis showing no significant difference between the use of lexical bundles for two groups in terms of frequency (Chi-square=3.22, P=.072>.05). Thus, it can be concluded that Iranian EFL learners use lexical bundles as frequently as NSs do in the discussion section of political science (DPS) articles.

Table 7. Chi-square Results

	Frequency
Chi-Square	3.225a*
D.F.	1
*Asymp. Sig.	.072

a. 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 201.0.*

**Asymp stands for asymptotic, i.e., non-exact or approximate significance.*

5.2. The Use of Lexical Bundles: Form

According to some earlier studies, the most reliable form for lexical clusters to be accepted as one lexical unit is a four-word one, and longer clusters or shorter ones neither contribute to a novel cluster, nor attain the central constituents of one lexical unit, eliminating necessary fragments (Chen & Baker, 2010; Hyland, 2008). In other words, if one lexical bundle unit is shorter than its standard form (four-word unit), it might have some missing parts which are typically the core components. For instance, *"as can be seen"* changing into *"can be seen"* with constituent *"as"* missing, cannot be a bundle and instead it is called a verb phrase (VP). Moreover, longer bundles beyond their standard form, certainly would allow for the attachment of other word classes which might not be the authentic constituent, for instance, *"as can be seen"* changing into *"as can be seen a"* an article (*"a"*) which is not the core constituent of this bundle due to the fact that it can be replaced by another article based on the following word.

Yielding prior results, an analysis of Chi-square was conducted to see whether Persian speaking and native writers use lexical bundles in the same sequential patterns in DPS articles. Table 10 shows the frequencies, percentages and standardized residuals of the lexical bundles regarding their form in the two groups of articles. Since none of the standardized residuals were above the range of +/- 1.96, it was concluded that there was no significant difference concerning Persian and native speaking writers' lexical bundles use regarding their sequential patterns in DPS articles.

Table 8. Lexical Bundles Frequencies, Percentages and Residuals (Form)

	Form	
NATIVE	Count	178
PERSIAN	Count	132
Total	Count	310

Correcting for a two-by-two table, the Chi-square value of zero further pointed out insignificant differences between Persian and native speakers' use of lexical bundles concerning their sequential patterns in DPS articles (Chi-Square (1)=0, P=1>.05). Bearing in mind these results, it can be said that Persian speaking writers do not use lexical bundles in the same sequential patterns in DPS articles as native speakers do. Table 11 displays the cells for the Chi-square analysis.

Table 9. Chi-square Cells for the Recognition of Lexical Bundles (Form)

	Value	Df	Asymp. Sig. (2-sided)
Continuity Correction	.000	1	1.000

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 22.76.

b. Computed only for a 2x2 table

5.3. The Use of Lexical Bundles: Structure

Considering three main structural categories suggested by Biber et al. (2004), the captured lexical units were analyzed to realize how they fit into each category. The first category of structure in lexical bundle unit comprises constituents incorporating verbs such as *"you don't have to"*. The second category comprises fragments incorporating dependent clauses such as *"I want you to"*, and the last category includes constituents incorporating noun phrases and prepositional phrases such as *"a little bit about"*.

Chi-square was run to find whether Persian speaking writers use lexical bundles concerning parts of speech (structures) differently and more than native writers do in DPS articles and to what extent this difference statistically significant. Table 12 shows the frequencies, percentages and standardized residuals of the lexical bundles concerning their structure by native and Persian speakers. The native speakers employed the first (Std. Residual=1.9) and second (Std. Residual=1.6) categories beyond what was expected. The positive values of Std. Residuals paved the way to arrive at that conclusion. On the other hand, the Persian speaking writers more used the third category (Std. Residual=1.5). Contrary arrangements can be detected considering the negative Std. Residuals. The native speakers used the third category below expectation and Persian speaking writers used the first and second categories below expectation.

Table 10. Frequencies, Percentages and Std. Residuals Lexical Bundles (Structure)

		Structure			
		First Category	Second Category	Third Category	Total
NATIVE	Count	39	17	122	178
	% within NATIVE NON-NATIVE	21.9%	9.5%	68.5%	100%
	Std. Residual	1.9	1.6	-1.3	
PERSIAN	Count	12	4	116	132
	% within NATIVE NON-NATIVE	9.0%	3.0%	96.6%	100%
	Std. Residual	-2.1	-1.8	1.5	
Total	Count	51	21	238	310
	% within NATIVE NONNATIVE	16.4%	6.7%	76.7%	100%

The Chi-square value of 17.46 (P=.000<.05) shows that there are significant differences concerning the native and Persian speaking writers in the use of lexical bundles in terms of their structure. Namely the Persian speaking

writers' use of the first category is significantly lower than what was expected. Considering these results it can be said that Persian speaking writers do not use lexical bundles in various parts of speech (structure) the same as NSs do in DPS articles. Table 13 illustrates the case.

Table 11. Chi-square Results for the Identification of Lexical Bundles Concerning Structure

	Value	Df	Asymp. Sig. (2-sided)
Pearson Chi-Square	17.468a	2	.000

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 10.87.

5.4. The Use of Lexical Bundles: Function

The final question to be answered in the present study was concerned with the functions of lexical bundles. To this end, another analysis of Chi-square was run to see whether Persian speaking writers use lexical bundles in various functions in DPS articles more than NS writers do. Native writers made more use of the first category (Std. Residual=2.2) while Persian speaking writers made more use of the second category (Std. Residual=0.6) and the third category (Std. Residual=1.4). Opposite arrangements can be detected considering the negative Residuals. The Native writers used more the second and the third categories. Table 14 illustrations the frequencies, percentages and standardized residuals of the lexical bundle units concerning their function between NS writers and Persian speaking writers.

Table 12. Frequencies, Percentages and Std. Residuals Lexical Bundles (Functions)

		Function			
		First Category	Second Category	Third Category	Total
NATIVE	Count	67	13	98	178
	% within NATIVENONNATIVE	37.6%	7.3%	55.0%	100%
	Std. Residual	2.2	-.6	-1.3	
PERSIAN	Count	24	13	95	132
	% within NATIVENONNATIVE	18.1%	9.84%	71.9%	100%
	Std. Residual	-2.4	.6	1.4	
Total	Count	91	26	193	310
	% within NATIVENONNATIVE	29.3%	8.3%	62.2%	100%

The Chi-square value of 15.27 (P=.000<.05) shows that there are significant differences concerning the native and Persian speakers' use of lexical bundles concerning their functions. It can be realized that the significant Residuals are related to the first category where the two values are more than the ranges of +/- 1.96. Namely the NS writers' use of the first category is significantly beyond the expected value whereas Persian speaking writers' use of the first category is significantly lower than what was expected. Considering these values we can conclude that Persian speaking writers did not use lexical bundles for the same functions as Native speaking did in DPS articles. Table 15 illustrate show the functions were used.

Table 13. Chi-square Cells for the Identification of Lexical Bundles (Function)

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	15.279a	2	.000

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.64.

6. Discussion

Hyland (2008) emphasized the nature of science and its influence on the manners of the lexical units. That is to say, those sciences purely prefer empirical approach, for instance engineering, which are called hard sciences; however, those which make use of interpretive approach in studies are called soft sciences. If we identify political sciences as belonging to the second category, and if we are inspired by the findings of this study, it would be possible to argue that such strong dichotomous categorization on the nature of the science appears to be artificial since within political context we can come up with another dichotomy. One category includes sciences which make use of experimental framework, somehow similar to the first category in soft/hard dichotomous categorization, and the other one comprises sciences which follow hermeneutic approach, somehow similar to the second category in the former dichotomy. Thus, there seems to be a kind of relative inclination in each type of science, the hard and the soft one, to fluctuate against their expected behaviors by showing characteristics of bundles of the opposite side of the dichotomy.

The categorization of the functional subcategories cannot be soundly detected all the time. Biber et al. (2004) and Cortes (2004) emphasized the relative intuitive approach of assigning one specific lexical unit under one functional subcategory. Considering sensitive nature of political sciences, especially in the contexts with some degree of political constraints, writers have the tendency to use words with unbiased characteristics to hold onto the

conservative mode of rhetoric, to avoid uncontrolled dramatizations, and to tone down the weight of authorship. Moreover, it is probable for the writers writing within this field of science to employ particular lexical units with a certain function in mind to communicate meaning ironically.

Though, when analyzing the text, the lexical bundle comes to belong to the irrelevant functional category concerning what the writer primarily directed to use so as to communicate meaning with the intended audience. This categorization can even be more artificial once the discourse freely yields itself to open interpretations. Namely, in contrast to experimental approach where some lexical bundles are specially used to carry out particular discourse functions for (e.g., *"as can be seen"*), interpretive approach use potential coincidences. The nature of hermeneutic reasoning is to explain particular unexplainable events rather than to provide statistical tables, to shed light on issue and pave the way for further study. While papers from Native speaking writers indicated that these writers communicate the political concepts as other writers do in other fields of science and follow channelized method of interpretation. In contrast Persian speaking writers would possibly, due to the lack of empirical work in the field, and prefer more open interpretations.

7. Conclusions

Analyzing the use of lexical bundles across the corpora of the present study, the following findings came up:

1. Native speaking writers used a systematic pattern of use since they purposefully made use of lexical bundles to communicate meaning, while Persian speaking writers had a typical style of use since they used bundles irregularly and in a predictable mode, not adjusted to the functions they planned to develop the discourse upon.

2. Four-word lexical clusters were still the most dependable form of bundles, and the genre of DPS articles which was the focus of the present study did not offer any new form of lexical bundles.
3. Nominal phrases and prepositional phrases were the most appropriate categories to deal with the abstract nature of the concept incorporated in academic writing, and the same was true for writing in the discourse of political sciences.
4. As was mentioned, methodological approaches (empirical or interpretive) influence the use of particular categories over the others concerning form and above all function; considering the findings of this study, Persian speaking writers were more likely to make statements with greater degrees of confidence, whereas NS writers were expected to hedge when making statements even when there was statistical logic to minimize authorship.

References

Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 101–122). Oxford: Oxford University Press.

Anthony, L. (2004). *AntConc: A Learner and Classroom Friendly, Multi-Platform Corpus Analysis Toolkit.* Center for English Language Education in Science and Engineering School of Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo 169-8555, Japan.

Ari, O. (2006). Review of three software programs designed to identify lexical bundles. *Language Learning and Technology*, 30-37, Georgia State University, USA.

Baradaran, A., & Sarfarazi, B. (2011). The impact of scaffolding on the Iranian EFL learners' English academic writing. *Australian Journal of Basic and Applied Sciences*, *5*(12), 2265-2273.

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

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

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

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

Brown, C. L. (2007). Strategies for making social studies texts more comprehensible for English-language learners. *The Social Studies*, (September-October), 185-188.

Brown, H. D. (1994). *Teaching by principles: An interactive approach to language pedagogy*. New Jersey: Prentice Hall Regents.

Byrd, P., & Coxhead, A. (2010). On the other hand: lexical bundles in academic writing and in the teaching of EAP. *Papers in TESOL, University of Sydney,* Australia.

Chaudron, C., & Richards, J. (1986). The effect of discourse markers on the comprehension of lectures. *Applied Linguistics*, *7*(2), 113-127.

Chen, Y., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Special Issue on Technology and Learning Vocabulary*, University of Lancaster, England.

Cheng, W. (2007). Concgramming: A Corpus-Driven Approach to Learning the Phraseology of Discipline-Specific Texts. *CORELL: Computer Resources for Language Learning*, *1*, 22-35.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, Department of English, Iowa State University, 203 Ross Hall, Ames, IA 50014, USA.

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

Erman, B. (2007). Cognitive processes as evidence of the idiom principle. *International Journal of Corpus Linguistics, 12*(1), 25-53.

Firth, J. R. (1951). Modes of meaning. *Essays and Studies* (The English Association), 118-149.

Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: Countering criticisms against corpus-based methodologies. *English for Specific Purposes, 24*, 321-332.

Gibbons, M. T. (2006). Hermeneutics, political inquiry, and practical reason: an evolving challenge to political science. *American Political Science Review, 100(4),* 563-571.

Giles, M. W., & Garand, J. C. (2007). Ranking political science journals: Reputational and citational approaches. *Political Science & Politics*, *40*(4), 741-751.

Grabe, W., & Kaplan, R. (1996). Theory and practice of writing: *An applied linguistic perspective*. New York: Longman.

Haswell, R. (1991). *Gaining ground in college writing: Tales of development and interpretation*. Dallas: Southern, Methodist University Press.

Hix, S. (2004). A global ranking of political science departments. *Political Studies Review*, *2*(3), 293–313.

Hoey, M. (1991). *Patterns of lexis in text*. Oxford: Oxford University Press. England.

Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *Center of Academic and Professional Literacies, Institute of Education*, University of London, 20 Bedford Way, London WC1H 0AL, United Kingdom.

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

Jalali, H., EslamiRaskh, A., & TavangarRizi, M. (2008). Lexical bundles and intradisciplinary variation: The case of applied linguistics. *Iranian Journal of Language Studies, 2*(4), 447- 484.

Jalali, H., EslamiRaskh, A., &TavangarRizi, M. (2009). Anticipatory 'it' lexical bundles: A comparative study of student and Published writing in applied linguistics. *Iranian Journal of Language Studies*, 177- 194.

Kubota, R., & Lehner, A. (2004). Toward critical contrastive rhetoric. *Journal of Second Language Writing, 13*, 7-27.

Kumaravadivelu, B. (2001). Toward a postmethod pedagogy. *TESOL Quarterly, 35*(4), 537-560.

Jespersen, O. (1924). *The philosophy of grammar*. Allen & Unwin, London.

Juknevičien, R. (2009). Lexical Bundles in learner language: Lithuanian learners vs. native speakers. *KALBOTYRA*, *61*(3), 1-12.

Laane, M. (2011). *Lexical bundles in engineering research articles*. 10th International Symposium„ Topical Problems in the Field of Electrical and Power Engineering "Pärnu, Estonia, January 10-15, Tallinn University of Technology.

Moon, R. (1998). *Fixed expressions and idioms in English: A corpus-based approach*. Oxford: Clarendon Press. London.

Nattinger, J., & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: OUP.

Neely, E., & Cortes, V. (2009). *A little bit about: Analyzing and teaching lexical bundles in academic lectures*. Language Value, Georgia State University, USA

Nekrasova, T., M. (2009). English L1 and L2 Speakers' Knowledge of Lexical Bundles. *Language Learning Journal*, *59*(3), 647-686.

Pang, W. (2010). Lexical bundles and the construction of an academic voice: A pedagogical perspective. *British Columbia Institute of Technology, Asian EFL Journal Professional Teaching Articles, 47.*

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Native like selection and native like fluency. In J. C. Richards & R.W. Schmidt (Eds.), *Language and Communication* (pp. 191-230). London: Longman.

Ping, P. (2009). A study on the use of four-word lexical bundles in argumentative essays by Chinese English majors: A comparative study based on WECCL and LOCNESS. *CELEA Journal (Bimonthly), 32*(3).

Rafiee, M., & Tavakoli, M. (2011). Structural Analysis of Lexical Bundles Across Two Types of English News Papers Edited by Native and Non-native speakers. *MJAL, 3*(2), 218-236.

Sinclair, J. (1991). *Corpus, concordance, collocation: describing English language*. Oxford: Oxford University Press.

Sinclair, J. M. (1999). The lexical item, in E. Weigand (Ed.) *Contrastive Lexical Semantics*. Amsterdam: John Benjamins.

VahidDastjerdi, H. & Rafiee, M. (2010). Corpus study of lexical bundles: Journalistic discourse in focus. *The Asian ESP Journal, 7*(3), 59-78.

Van Manen, M. (1997). *Researching lived experience: Human science for an action sensitive pedagogy*. London, ON: The Althouse Press.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Appendix A

Distribution of Lexical Bundles Native Speaking Writers' Corpus

N	Freq	Lexical Bundles	N	Freq	Lexical Bundles	N	Freq	Lexical Bundles
1	13	the nature of the	61	4	annual meeting of the	120	3	extent to which a
2	12	at the same time	62	4	are likely to be	121	3	for a number of
3	12	in the case of	63	4	as in the case	122	3	for a variety of
4	11	in the united states	64	4	at one point in	123	3	for many of the
5	11	on the basis of	65	4	be thought of as	124	3	for their helpful comments
6	10	as well as the	66	4	but it is not	125	3	has no effect on
7	10	in terms of the	67	4	by far the most	126	3	I am grateful to
8	10	in the context of	68	4	can be used to	127	3	in any of the
9	10	the extent to which	69	4	each of the three	128	3	in each of the
10	9	a large number of	70	4	for the most part	129	3	in order to win
11	9	on the other hand	71	4	important to note that	130	3	in the area of
12	8	are more likely to	72	4	in a series of	131	3	in the house of
13	8	as a function of	73	4	in addition to the	132	3	in the last column
14	8	fit of the model	74	4	in the face of	133	3	in the next section
15	8	it is important to	75	4	in the first row	134	3	in the previous section
16	8	the size of the	76	4	in the number of	135	3	in the second column
17	7	a special case of	77	4	in the set of	136	3	in the u s
18	7	department of political science	78	4	in this case the	137	3	in this paper we
19	7	for a discussion of	79	4	is assumed to be	138	3	is equal to the
20	7	it is useful to	80	4	it is difficult to	139	3	is important to note
21	7	one of the most	81	4	levels of political information	140	3	is not the case
22	7	the degree to which	82	4	may be interpreted as	141	3	is proportional to the
23	7	the magnitude of the	83	4	more likely to be	142	3	it is also possible
24	7	to the extent that	84	4	of political science vol	143	3	it is possible that
25	6	in the presence of	85	4	of the impact of	144	3	last column of table
26	6	of the dependent variable	86	4	of the relationship between	145	3	of a set of
27	6	the assumption that the	87	4	one point in time	146	3	of the house of
28	6	the effect of a	88	4	our discussion of the	147	3	of the independent variables
29	6	the end of the	89	4	political science vol no	148	3	of the most important
30	6	the impact of the	90	4	that the effects of	149	3	of the national election
31	6	the study of political	91	4	the core of the	150	3	of the paper is
32	6	to the study of	92	4	the effect of the	151	3	of the public s
33	6	we would like to	93	4	the importance of the	152	3	of the university of
34	5	a great deal of	94	4	the increase in the	153	3	of this paper is
35	5	American journal of political	95	4	the results for the	154	3	on the one hand
36	5	are presented in table	96	4	the ways in which	155	3	one of the best
37	5	can be thought of	97	4	to the problem of	156	3	or to put it
38	5	for each of the	98	4	to the use of	157	3	ordinary least squares of
39	5	in the form of	99	4	we do not have	158	3	political science university of
40	5	in the sense that	100	4	with a discussion of	159	3	presented at the annual

41	5	in this section we	101	3	a large proportion of	160	3	should come as no
42	5	is one of the	102	3	a number of reasons	161	3	since there is no
43	5	journal of political science	103	3	a time varying covariate	162	3	state university of new
44	5	of California Los Angeles	104	3	also more likely to	163	3	statistically significant at the
45	5	of the number of	105	3	and there is no	164	3	take into account the
46	5	professor of political science	106	3	are based on the	165	3	that it can be
47	5	the baseline hazard function	107	3	as a consequence of	166	3	that the number of
48	5	the case of the	108	3	as a result of	167	3	the annual meeting of
49	5	the discussion of the	109	3	as a set of	168	3	the basis of this
50	5	the effects of the	110	3	as can be seen	169	3	the coefficient for the
51	5	the fit of the	111	3	as the number of	170	3	the context of a
52	5	the house of representatives	112	3	as well as to	171	3	the distribution of the
53	5	the strength of the	113	3	at the annual meeting	172	3	the fact that the
54	5	through the use of	114	3	at the end of	173	3	the first is that
55	5	university of California los	115	3	be interpreted as the	174	3	the goal is to
56	5	we find that the	116	3	but there is no	175	3	the last column of
57	4	a change in the	117	3	by the university of	176	3	the limiting case of
58	4	a discussion of the	118	3	can be expressed as	177	3	the presence of a
59	4	a function of the	119	3	estimates are presented in	178	3	the problem is that

Appendix B

Distribution of Lexical Bundles in Persian Speaking Writers' Corpus

N	Freq	Lexical Bundles	N	Freq	Lexical Bundles	N	Freq	Lexical Bundles
1	24	in the middle east	45	5	as one of the	90	4	one of the most
2	16	on the other hand	46	5	as well as a	91	4	parts of the world
3	15	the council of guardians	47	5	at the end of	92	4	paved the way for
4	15	the end of the	48	5	in post-revolutionary Iran	93	4	the collapse of the
5	14	as well as the	49	5	is no doubt that	94	4	the idea of a
6	14	at the same time	50	5	of the armed forces	95	4	the one hand it
7	10	as a result of	52	5	of the armed forces	96	4	the power of the
8	10	of the constitutional revolution	53	5	of the Islamic republic	97	4	the private sector in
9	10	the constitutional revolution of	54	5	of the middle east	98	4	the relationship between the
10	10	the fact that the	55	5	of the post revolutionary	99	4	the rest of the
11	9	in the price of	56	5	on the part of	100	4	the same way as
12	9	Islamic republic of Iran	57	5	the aim of this	101	4	the time of the
13	8	at the time of	58	5	the destruction of the	102	4	to percent of the
14	8	of the private sector	59	5	the outcome of the	103	4	united states and the
15	8	of the regime s	60	5	the political economy of	104	4	vis a vis the
16	8	on the one hand	61	5	the shah and the	105	4	year war with Iraq
17	8	the role of the	62	5	the study of the	105	4	year war with Iraq
18	7	in the third world	63	5	the united states has	106	3	about the nature of
19	7	of the country side	64	5	was one of the	107	3	an increasing number of
20	7	per cent of the	65	5	would have to be	108	3	as far as the
21	7	the impact of the	66	5	a great deal of	109	3	at the expense of
22	7	the Iran Iraq war	67	5	a result of the	110	3	be described as the
23	7	the Islamic republic of	68	5	and on the other	111	3	been one of the
24	7	the middle east and	69	4	and the private sector	112	3	between Iran and the
25	7	the united states and	70	4	and the united states	113	3	between the united states
26	6	in the case of	71	4	as a result the	114	3	for the first time
27	6	no more than a	72	4	as well as in	115	3	from the perspective of
28	6	percent of the total	73	4	be attributed to the	116	3	half of the twentieth
29	6	the price of oil	74	4	by a group of	117	3	in favor of the
30	6	to the extent that	75	4	by the council of	118	3	in opposition to the
31	5	bear in mind that	76	4	be described as	119	3	in terms of the
32	5	in a number of	77	4	eight year war with	120	3	in the aftermath of
33	5	in the absence of	78	4	first half of the	121	3	in the annals of
34	5	in the course of	79	4	in addition to the	122	3	in the face of
35	5	in the midst of	80	4	the bazaar and	123	3	in the name of
36	5	of the state in	81	4	in the country s	124	3	in the nature of
37	5	one of the main	82	4	in the form of	125	3	in the pages of
38	5	the case of Iran	83	4	in the Islamic republic	126	3	in the works of
39	5	the course of the	84	4	Iran Iraq war and	127	3	Iran and the united
40	5	the emergence of a	85	4	is based on the	128	3	it was in the
41	5	the fall of the	86	4	it is clear that	129	3	more than a few
42	5	the ranks of the	87	4	of the nineteenth century	130	3	of religion and state
43	5	the state and the	88	4	of the twentieth century	131	3	of the Iranian revolution
44	5	there is no doubt	89	4	on the basis of	132	3	of the revolution the