Using *it* Bundles in Published and Unpublished Writings

Hassan Jalali

Assistant Professor, Isfahan University of Technology

h.jalali@cc.iut.ac.ir

Abstract

Lexical bundles are known as important elements of coherent discourse that have been the subject of much research. While the previous research has been mainly concerned with exploring variations in the use of these word sequences across different registers and disciplines, very few studies have addressed the use of some particular groups of lexical bundles within some types of academic writing. To probe in to the possible generic variations, this research focused on anticipatory *it* bundles as a particular structural group of these word sequences. More specifically, this study chose to investigate range, frequency, and function of these word clusters in applied linguistics published and unpublished writings. Through the use of two big corpora and text analysis programs, this study found that *it* bundles were used relatively frequently in both published and unpublished writings. Functional analysis also showed that *it* bundles served a wide variety of functions in both types of writings. Notwithstanding these findings, this study also revealed that some *it* bundles commonly used in unpublished writing did not count as bundles in published writing.

Keywords: Applied Linguistics, Published Writing, Unpublished Writing, *It* Bundles

1. Introduction

Lexical bundles, also known as clusters and chunks (Hyland, 2008a, 2008b), were first introduced and defined by Biber, Johansson, Leech, Conrad, and Finegan (1999). They defined lexical bundles as "recurrent expressions, regardless of their idiomaticity, and regardless of their structural status" (p. 990). More importantly, they referred to frequency as the defining characteristic of bundles; to put it more succinctly, in order for a word combination (e.g., *on the other hand, at the same time, it is necessary to*, etc.) to count as a bundle, it must occur at least twenty times in a corpus made of one million words with the additional requirement that this rate of occurrence be realized in at least five different texts to guard against idiosyncratic and repetitive uses. Lexical bundles are identified on the basis of frequency and breadth of use (Cortes, 2002, 2004). Fixedness in form (e.g., *on the basis of* not **on a basis of)* and non-idiomatic meaning are other properties of bundles that must be noted. Among other registers, lexical bundles have been found to be an important part of academic discourse too (Biber et al., 1999).

Such word sequences have been classified structurally (Biber et al., 1999; Biber, Conrad, & Cortes, 2004; Biber, 2006; Jalali, EslamiRasekh & TavangarRizi., 2008, 2009) as well as functionally (Cortes, 2002, 2006; Biber & Barbieri, 2007; Hyland, 2008a, 2008b; Jalali, 2009, 2013; Jalali & Ghayoumi., 2010).These word clusters can serve a wide range of discursive functions such as organization of discourse, expression of stance, and reference to textual or external entities (Biber & Barbieri, 2007; Jalali, 2013). Some studies conducted in this regard are briefly reviewed here.

Since 1999, a number of studies have been specifically launched to explore possible differences and\or similarities in the use of bundles between a few disciplinary fields (Cortes, 2002, 2004; Hyland, 2008a, 2008b), registers, such as

conversation, fiction, news, academic prose, classroom teaching and non-conversational speech (Biber et al., 1999; Biber et al., 2004, Biber & Barbieri, 2007), genres (Hyland, 2008b; Jalali, 2013), and different degrees of writing expertise (Cortes, 2002, 2004; Jalali, 2009; Jalali et al., 2008, 2009).

Overall, these studies have indicated that lexical bundles are strong discipline, genre, and register discriminators. This means that apart from some overlaps, each discipline, genre, or register draws on its own set of bundles to organize its discourse, express stance, and refer to different parts of the evolving text or elements outside the text. The findings have also highlighted that many lexical bundles favored by experts in given disciplinary areas may not be used by novices, who could be students or developing writers with varying degrees of language proficiency and disciplinary expertise.

Interestingly, there is also usually a correlation between the structural type of bundles and the function they serve in the discourse (Biber et al., 2004); for example, *it* bundles (e.g., *it should be noted, it can be seen*), the subject of the present study, are usually used to act as metadiscourse elements (Hyland, 2000, 2008a, 2008b) or expressions of stance (Biber, 2006). Biber et al (1999) show that *it* clauses followed by either *to* (as in *it is important to note that this relationship may always be true*) or *that* (as in *it is clear that this policy is unlikely to lead to fruitful results*) are common in academic writing and their relatively frequent presence has been substantiated in a range of academic genres (Hewings & Hewings, 2002).

The study of this structural group of lexical bundles can be important for two reasons. First of all, there is some evidence suggesting that for many non-natives, this structure can pose serious degrees of difficulty mostly because of the absence of an *it* structure in some languages (Jacobs, 1995, Hewings and Hewings, 2002). Second, if we could recognize the importance of this structure

as a metadiscursive element or a stance expression, it would be important to identify the range of interpersonal meanings conveyed by such word clusters. It is now axiomatic that such *it* clauses or bundles are usually good means by which writers can express their opinions, evaluate the subject matter, and engage with readers (Hewings & Hewings, 2002).

According to Hewings and Hewings (2002), as explained in more detail in the methodology section, clauses with *it bundles* can have four metadiscoursal or interpersonal roles: hedges (showing speaker or writer's tentativeness and uncertainty about the following proposition), attitude markers (expressing writer's attitude toward the content), emphatics (stressing writer's certainty about the force, and the credibility of the propositional meaning), and attribution (convincing the reader through a general or specific reference). The review of the literature showed that very few studies have focused on the use of *it* bundles within some key genres of academy (see Hewings & Hewings, 2002; Hyland, 2008a). Especially important is the scarcity of studies that would address specific phraseological practices in different disciplinary areas to describe and explain possible differences and/or similarities between experts and novices in their use of these word combinations in their respective high-stakes genres.

The purpose of this study was to compare the use of one structural class of bundles in published and unpublished writings of one disciplinary area of applied linguistics through the use of two corpora representing these two types of academic writing. The assumption was that exploring possible variations in the use of such word combinations across these two categories of writings could contribute to a better understanding of phraseological preferences and practices in different discourse communities. This is especially important since lexical bundles are building blocks of coherent discourse with important

functions and the previous research (e.g., Cortes.2006) shows that an important feature of proficient writing is the appropriate use of these sequences. So comparing published and unpublished wiring can shed some light on these practices and show whether there are differences between published and unpublished writers with regard to the use of these bundles.

More specifically, the study probed the use of *it* bundles in two groups of applied linguistics academic writing: research articles, representing published writing, and postgraduate theses, as written by students at master's and doctoral levels, reflecting unpublished writing. Accordingly, two corpora of published and unpublished writings in applied linguistics were used in this study to find whether these two types of writing in one single disciplinary area are similar or different from each other in terms of their use of *it* bundles. At the same time, comparing the two types of writing could show the extent to which postgraduate students' use of *it* bundles is similar to or different from that of published writers. The results are likely to shed some light on the possible patterns of bundles choices among publsihed and unpublsihed writers.

2. The Study

2.1. Research Questions

This study, therefore, addressed the following questions:

1. What are the most frequent four-word *it* lexical bundles in applied linguistics published articles and unpublished writings?

2. Is there any similarity or difference in the range, frequency, and function of *it* lexical bundles in published and unpublished writings?

2.2. Corpora

Two corpora were used in this study. The first corpus included published writing in the discipline of applied linguistics, and the second one represented students' unpublished writing at the postgraduate level. More specifically, the first corpus consisted of research articles published in some applied linguistic prestigious journals while the second one comprised master theses and doctoral dissertations written by some EFL students in the discipline of applied linguistics. Each of these corpora is further described in more details below.

The first corpus had been originally prepared by Jalali (2009) for his study on variations in the use of lexical bundles within a single discipline: applied linguistics. The basis for selection of journal articles was mostly previous corpus-based studies done on the scientific discipline of applied linguistics, the advice given by experts in the field, and the prior research done on applied linguistics published literature (e.g., Hyland, 2008a, 2008b). Table 1 describes the journals, the numbers of texts, and the number of words in this corpus. All these journals had some relevance to English language education. The authors of these articles were native or nonnative speakers of English who could be regarded as expert writers in the field given their publication in these high-rank prestigious journals. The second corpus, also collected by Jalali (2009), was made of master theses and doctoral dissertations written by some postgraduate EFL students (L1 Persian) in the field of English language teaching to represent postgraduate unpublished writing.

Table 1. *Published Writing Corpus Word Count*

Journal	No of texts	No of words
Applied Linguistics	29	240212
English Language Teaching	45	151506
English for Specific Purposes	37	250576
English for Academic Purposes	20	125236
Second Language Writing	14	108663
Linguistics and Education	11	94614
System Journal	45	247156
Total	**201**	**1217963**

Table 2. *Unpublished Writing Corpus Word Count*

Students' genres	Number of texts	Number of words
Master theses	22	441033
Doctoral dissertations	12	476922
Total	34	917955

2.3. Data Analysis Tools

2.3.1. Computer Programs

Two computer programs were used in this study: Antconc3.2.1w (Anthony, 2007), and Wordsmith (Scott, 2008). The former was used for the identification of lexical bundles and concordance while the latter was only used to find the number of texts within which each bundle had been used. These two are described more below.
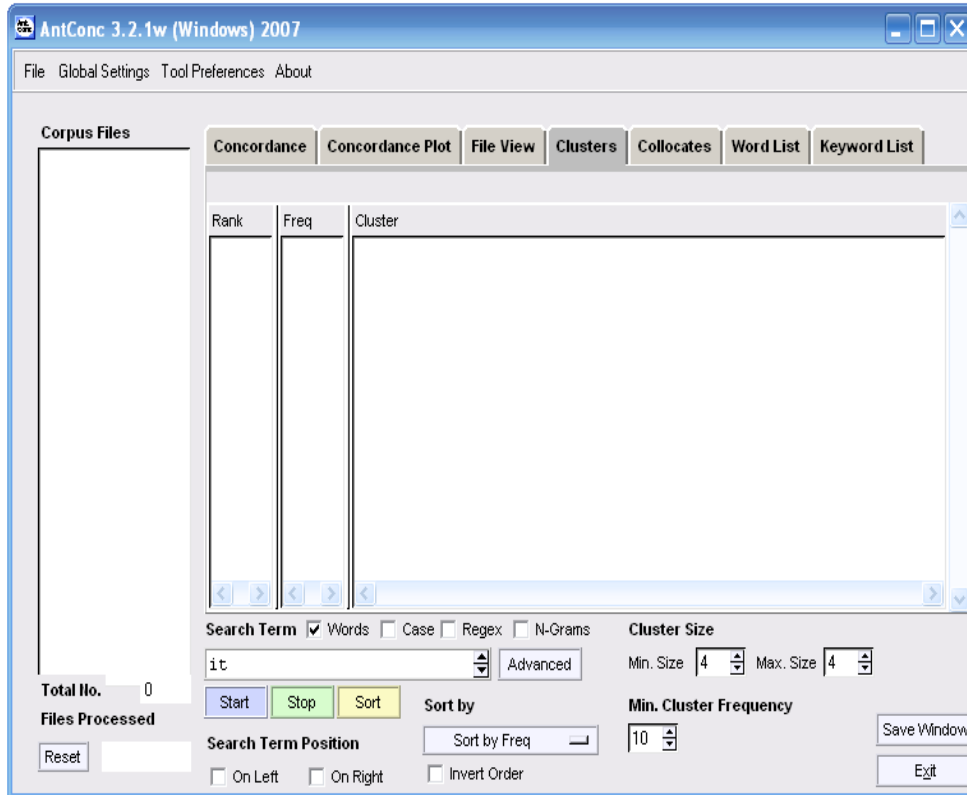
Antconc3.2.1.w is a free concordance program designed and developed by Anthony (2007)(seeFig.1). This study used it to identify "it" bundles and find their frequency. It has useful tools such as concordance, concordance plot, file view, N-grams, collocates, word list, and keyword list that are used to analyze

the text. The concordancer also makes it possible to see each of the clusters in actual textual context within which it has originally been used.

Among all these tools, there is one by which it is possible to identify word combinations, clusters, or lexical bundles of different lengths and frequencies in small or large corpora. All lexical bundles in corpora of different sizes with their actual frequencies are found and displayed by inserting a set of commonly key words with which the bundles collocate such as prepositions (e.g., at, of, on, etc), modals (e.g., can, should, could, may, etc), etc, and deciding on the minimum optimal frequency (e.g., twenty in a corpus of one million words); then the required number of words in clusters (i.e., three, four, five, or six) is specified. As Antconc3.2.1.w could not count and display the number of different texts, WordSmith tools5 (Scott, 2007) was applied for the identification of lexical bundles. This program is similar in many ways to Antconc3.2.1.w, but it can display the number of texts in which bundles have been used. This is very important because in order for a cluster to count as a bundle, it must be used by a number of writers in different texts to guard repetitive uses by the same writer in the same text. So when all candidate lexical bundles were identified by the first computer program, each of them was again searched on Wordsmith tools 5 to find the number of texts in which they had been used. Only those four- word combinations could count as lexical bundles that had been used ten times and in at least five different texts no matter how frequent they were (Biber et al., 1999). It must be noted that like some other studies (e.g., Hyland, 2008s; Cortes, 2004), four-word lexical bundles were of interest in this study as they are more frequent than five or six-word bundles and at the same time, serve more diverse functions than three-word bundles.

Figure1. Tools of Antconc3.2.1.

2.3.2. Functional Analysis of Bundles

The focus of this study was on four-word bundles because previous research has shown that they are far more common than five-word strings and offer a clearer range of structures and functions than three-word bundles. Bundles are essentially extended collocations defined by their frequency of occurrence and breadth of use, but the actual frequency cut offs are somewhat arbitrary. This study took a conservative approach by setting a minimum frequency of 10 times per million words and an occurrence in at least 10% of texts, i.e., word

combinations had to appear in five or more texts to be regarded as bundles. These specifications were according to the extensive research carried out by Biber et al. (1999).

The data was analyzed in three steps. First, all *it* bundles were identified in the two corpora along with their actual frequencies and the number of texts in which they had been used. Second, by using a functional typology of *it*-clauses developed by Hewings and Hewings (2002) (see table 3) and the AntConc 3.2.1 concordance (Anthony, 2007) and Wordsmith tool5 (Scott, 2008) for doing the quantitative analysis of bundles, an attempt was made to probe the context in which bundles had been used to decide on the most predominant functions. This was done by two raters to ensure the accuracy of classification. In the third stage, the results were compared to find whether applied linguistics published and unpublished writings were different from/or similar to each other in terms of range, frequency, and function of anticipatory *it* bundles. While there are already some functional classifications of lexical bundles (e.g., Cortes, 2002; Biber et al., 2004; Hyland, 2008a, 2008b), Hewings and Hewings' functional taxonomy of *it*-clauses (2002) was used in this study since it specifically dealt with the interpersonal functions of this structural group. These categories were described in the introduction section.

Table 3. *Interpersonal Functions of It Clauses (Hewings & Hewings, 2002, P.372)*

Interpersonal functions *It*-clauses of	subcategories	Realization
1 hedges	1a likelihood/possibility/certainty; importance/value/necessity etc. 1b what a writer thinks/assumes to be//will be/ was the case	It is likely, it seems improbable, it would certainly appear, it could be argued, it was felt
2 attitude markers	2a the writer feels that something is worthy of note 2b the writers evaluation	It is of interest to note; it is worth pointing out; it is noteworthy; it is important
3 emphatics	3a the writer indicates that a conclusion/deduction should be reached; that a proposition is true 3b the writer strongly draws the reader's attention to a point 3c the writer expresses a strong conviction of what is possible/ important/necessary, etc.	It follows; it is evident; it is apparent It is important to stress; it should be noted; it must be recognized It is clear; it is impossible; it is safe to assume
4 attribution	4a specific attribution (with a reference to the literature) 4b general attribution (no referencing)	It has been proposed (+ reference) It is estimated (+ no reference)

3. Results

3.1. Lexical Bundles in Applied Linguistics Published Writing

Table 4 shows *it* bundles in the corpus of published writing along with the frequency and the number of texts in which they have been used. A total of seventeen different *it*-bundles were drawn from this corpus. The overall use of

these bundles was 449, tantamount to 0/036% of the whole corpus. In terms of function, this corpus capitalized maximally on attitude markers (43.20%) and minimally on the attribution markers (3.80%) (See table 5). Some of the most frequent it-bundles were: *it is important to* (88 times), *it should be noted* (40 times), *it is possible that* (38 times), and *it is difficult to* (36 times). A large number of it bundles in this corpus had also the pattern of *it + V be + adjective + that/to*. It also seemed that the use of such bundles by published writers in applied linguistics helped writers to encode different interpersonal meanings.

Table 4. *Anticipatory It Lexical Bundles Applied Linguistics Published Writing*

Lexical bundles	Frequency	Number of texts
it is important to	88	58
it should be noted	40	32
it is possible that	38	23
it is difficult to	36	31
it is necessary to	34	29
it is clear that	33	26
it is possible to	25	22
it is interesting to	25	18
it was found that	19	15
it is important that	17	15
it can be seen	17	11
it is hoped that	14	12
it is not clear	14	11
it is suggested that	14	10
it could be argued	12	12
it may be that	12	8
it seems that the	11	10

The following examples from this corpus can show the use of some of such bundles by published writers:

As a result of these experiences, *it is possible that* these students retrospectively constructed the mainstream basic writing section as being "for American students" and assumed that such an environment would have been more stressful for them than the multilingual one.

It may be that students in the sciences, all PhD students in our case, focused more on the explicit goals of the courses, which answer an urgent need to publish.

It is important to emphasize in this section that although the majority of the words that remind us of a non-Spanish spelling are grouped among those which form their plural by adding the suffix -s, we have found two examples of zero plural morpheme: Bluetooth and reflex.

By way of final comment, *it is interesting to* note that the results of the study are compatible with a view of language learning that distinguishes the acquisition processes involved in the development of implicit L2 knowledge ….

Table 5. Overall Functional Description of it-Bundles in Applied Linguistics Published Writing

Subcategories	Number	Frequency	Percentage %
Hedges:			
1a likelihood/possibility/certainty; Importance/value/necessity etc.	2	63	14.03
1b what a writer thinks/assumes to be//will be/ was the case	3	35	7.79
Attitude markers:			
2a the writer feels that something is worthy of note	0	0	0
2b the writers evaluation	6	194	43.20
Emphatics:			
3a the writer indicates that a conclusion/deduction should be reached; that a proposition is true	0	0	0
3b the writer strongly draws the reader's attention to a point	1	40	8.90
3c the writer expresses a strong conviction of what is possible/ important/necessary, etc.	2	67	14.92
Attribution:			
4a specific attribution (with a reference to the literature)	1	17	3.78
4b general attribution (no referencing)	0	0	0
Epistemic:			
5a Certain	0	0	0
5b uncertain	1	14	3.11
5c impersonal	1	19	4.23
Total	17	449	100

3.2. Lexical Bundles in Applied Linguistics Unpublished Writing

As shown in table 6, there were again seventeen different anticipatory *it* bundles in the corpus of postgraduate writing: *it was found that, it is important to,* and *it should be noted* were some of the more frequent lexical bundles used by postgraduate students. The overall frequency of all *it*-bundles in this corpus was 354, covering 0.038% of the whole corpus. However, the frequency in this corpus was lower than that of applied linguistics published writing (0/036%). Interestingly, however, there were some bundles in this corpus (i.e., *it should be mentioned, it was revealed that,* and *it is assumed that*) that were just used in postgraduate writing, not published writing.

Table 6. *it Lexical Bundles in Applied Linguistics Unpublished Writing*

Lexical bundles	Frequency	Number of texts
it was found that	42	13
it is important to	38	17
it should be noted	28	16
it can be concluded	26	13
it seems that the	24	13
it should be mentioned	22	11
it is possible to	20	11
it is believed that	19	11
it is difficult to	18	13
it is necessary to	18	13
it was revealed that	18	6
it is assumed that	17	11
it is not clear	14	12
it is possible that	14	10
it is clear that	12	10
it can be seen	12	8
it is obvious that	12	8

Functional analysis also showed that unpublished writers, like published counterparts, were able to employ lexical bundles in the discourse to serve a wide variety of different functions (see table 7). As can be seen, among the five categories, 34.45% of all it-bundles were devoted to those bundles expressing epistemic meanings. Emphatics were the second group of bundles in terms of the occurrence covering around 26% of all it-bundles, with attitude markers (19.77%) and hedges (16.37) being the next. And finally, the category of attribution was found to be in the last rank with a portion of 3.38%. The following examples can show some of these different uses:

In general, *it seems that the* newspapers encode and reinforce asymmetries between EU and Iran in their representation, in the context of west-dominated international politics.

it is possible to speculate the meaning of unknown words when 95 percent of the words in the text are familiar to the reader.

Thus, when authors use expressions such as my purpose for you in this chapter is to, *it is important to* note that, perhaps, and surprisingly, they are using meta-discourse.

it is not clear that this is due to their sharing a memory representation, as there is a great deal of debate over how bilingual memory is organized.

Table 7. *Overall Functional Description of it-Bundles in Applied Linguistics Unpublished Writing*

Subcategories	number	frequency	Percentage %
Hedges:			
1a likelihood/possibility/certainty; importance/value/necessity etc.	2	34	9.60
1b what a writer thinks/assumes to be//will be/ was the case	1	24	6.77
Attitude markers:			
2a the writer feels that something is worthy of note	0	0	0
2b the writers evaluation	3	70	19.77
Emphatics:			
3a the writer indicates that a conclusion/deduction should be reached; that a proposition is true	0	0	0
3b the writer strongly draws the reader's attention to a point	2	50	14.12
3c the writer expresses a strong conviction of what is possible/ important/necessary, etc.	3	42	11.86
Attribution:			
4a specific attribution (with a reference to the literature)	1	12	3.38
4b general attribution (no referencing)	0	0	0
Epistemic:			
5a Certain	2	37	10.45
5b uncertain	2	43	12.14
5c impersonal	1	42	11.86
Total	17	354	100

3.3. Comparisons

3.3.1. Comparisons in Terms of Variety and Frequency of Bundles

Probably, the most important finding of this study was related to the similarity between the two corpora under investigation in terms of the variety of it-bundles employed. Although the number of texts used in the corpus of applied linguistics articles was six times more than that of postgraduate writing, these two corpora were very similar in terms of the number of different bundles used. In terms of the variety of shared it-bundles, out of seventeen (17) bundles used in applied linguistics research articles, fifty-three percent were used in the other corpus. Table 8 shows shared it-bundles in the two corpora. The results obtained also showed that it-bundles were used quite frequently both in applied linguistics published and unpublished writing (368, and 386, respectively) as can be seen in table 9.

Table 8. *Shared it-Bundles in the Two Corpora*

Lexical bundles	Frequencies in ALRAs/ ALPGs
it can be seen	17/12
it is possible to	25/20
it was found that	19/42
it should be noted	40/28
it is important to	88/38
it is clear that	33/12
it is necessary to	34/18
it is difficult to	38/18
it is possible that	38/14
Total	330/202

Table 9. *Variety and Overall Use of it-Bundles in the Two Corpora*

Genres	ALRAs	ALPGs
Number of bundles	17	17
Actual frequency	449	354
Per million	368	386

3.4.2. Comparisons in Terms of Functions of Bundles

In terms of generic differences in the variety of bundles used in each major functional category, it was found that the variety of it-bundles serving as hedges and attitude markers in applied linguistics published writing was more than that of unpublished writing. While in the case of emphatics and attributions, there was no significant difference, for epistemic meanings, it was the unpublished writing that made a considerably heavier use.

Table10. *Functional Comparison of it-Bundles in the Two Corpora (Applied Linguistics Published Writing, Unpublished Writing)*

Categories	Number of bundles	Frequency (normalized)	Percentage %
Hedges	5/3	80/64	21.73/16.58
Attitude markers	6/3	159/76	43.20/19.68
Emphatics	3/5	88/101	23.91/26.16
Attribution	1/1	14/13	3.80/3.36
Epistemic	2/5	27/132	7.33/34.19
Total	17/17	368/386	100/100

There were attitude markers (i.e., *it is interesting to, it is important that,* and *it is hoped that*) that were only used by published writers in applied linguistics. Especially important was the higher frequency of *it is important to* in the corpus of published writing. *It is difficult to* be another bundle which was also used more heavily by applied linguistics published writers. Interestingly

19

and in contrast to findings of some previous studies (e.g., Hyland, 2008a, 2008b, Cortes, 2004), postgraduate students, as unpublished writers, who might not have established themselves as members of their disciplinary communities, were found to be confident in using those stretches that involved making emphasis. This, unlike the findings of the previous research (e.g., Hyland, 2008a), showed that postgraduate students could express their attitudinal meaning in a relatively straightforward manner.

4. Discussion

Relatively frequent use of *it* bundles in unpublished writing could be taken as an important finding in this study in spite of the previous research (e.g., Cortes, 2004) showing that students tended to rely less on bundles in the development of their discourses. The analysis of the corpus of unpublished writing showed that the number of different lexical bundles used by students in their writing was almost as many as those used by published writers. It seemed that students, both at the master's and doctoral levels, tended to use it lexical bundles for a wide variety of discursive functions.

Students' relatively frequent use of *it* bundles could be due to the fact that they had already been exposed to such word-sequences several times in their prior readings of applied linguistics published literature. There is almost no doubt that postgraduate students have repeatedly observed different lexical bundles in different research articles they may have studied for doing and writing their own research. Furthermore, given that *it* lexical bundles are very pervasive in university written language (Biber at al., 1999; Biber & Barbieri, 2007) and they may have a formulaic status (Wray, 2000), the acquisition of such word combinations may not confront students with a very difficult task,

especially for students at master's and doctoral levels who have a high language proficiency.

Probably also, lexical bundles are retrieved and stored whole from memory through holistic rather than analytical processes (Conklin & Schmitt, 2008) and therefore, postgraduate students may have little if any difficulty not only in understanding but also in producing *it* bundles. There may be a processing advantage in the use of bundles as some formulaic sequences have been shown to be easier to use (Conklin & Schmitt, 2008). It can also be postulated that lexical bundles can act as handy short-cuts or frames (Biber & Barbieri, 2007) through which writers can scaffold their propositional meanings with a relative ease. However, automatic acquisition of all lexical bundles may not always occur as this study also showed that there were it bundles in applied linguistics published writing on which students did not draw quite frequently. These word sequences are not idiomatic in meaning and hence they may be easy to understand, but they do not seem to be marked and perceptually salient. Consequently, there may still be a need to incorporate them in L2 syllabus or EAP (English for academic purposes) courses for an increased pedagogical focus. This can be especially helpful for those who need to understand and use bundles in their future target genres (Hyland, 2008b).

One other explanation that may provide a partial account for the frequent use of *it* bundles in unpublished writing could be provided by referring to a notion of limited vs. extended discourses. Unlike published writing, which is an academic genre characterized by compactness of information, less repetition and redundancy, and limitations in space, theses and dissertations are extended discourses which allow student writers to express themselves, evaluate the subject matter and engage with readers (Hyland, 2008b). As Hyland (2008b) holds, theses carry a strong burden of assessment and postgraduate students go

through lots of pains to show their knowledge, familiarity with research procedures, and adherence to disciplinary accepted practices with as much recycling, explanation, persuasion, and expansion of their ideas as possible. While too much use of formulaic expressions in general and lexical bundles in particular by students could be partly akin to repetitiveness and at times verboseness, especially by L2 users, in most cases, these multi-word sequences served important functions in high-stake student genres like master theses and doctoral dissertations.

5. Conclusion

In spite of one decade research on lexical bundles, much still remains to be explored about this group of word combinations which can contribute to an almost overlooked dimension of genre analysis (Hyland, 2008b). Identifying lexical bundles in other disciplines, registers, and genres, examining the formulaic status of these multi-word sequences (Biber & Barbieri, 2007) and better probing the effect of a pedagogical treatment on their acquisition could be areas worth exploring in future research.

References

Anthony, L. (2007). Antconc3. 2. 1: A free text analysis software. Available on line athttp://www.antlab.sci.waseda.ac.jp/.

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

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

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

Biber, D, Johansson, S, Leech, G, Conrad S, & Finegan, E. (1999). *Longman grammar of spoken and written English* (p. 990). Harlow: Pearson.

Conklin, K, & Schmitt, N. (2008). Formulaic Sequences: Are They Processed More Quickly than Nonformulaic Language by Native and Nonnative Speakers? *Applied linguistics, 29*(1), 72-89.

Cortes, V. (2002). *Lexical bundles in academic writing in history and biology.* Unpublished Doctoral dissertation: Northern Arizona University, Arizona.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes, 23,* 397–423.

Gibbs R, Bogadanovich J, Sykes J, Barr D. (1997). Metaphor in idiom comprehension. *Journal of Memory and Language,* 37, 141-154.

Hewings, M. & Hewings, A. (2002). "it is interesting to note that….." : A comparative study of anticipatory 'it' in student and published writing. *English for Specific Purposes, 2*1, 367-383.

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

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

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

Jacobs, R.A. (1995). *English syntax: A grammar for English language professionals.* New York: Oxford University Press.

Jalali, H. (2009). *Lexical bundles in applied linguistics: Variations within a single discipline.* Unpublished doctoral dissertation: Isfahan University, Isfahan.

Jalali, H. (2013). Lexical bundles in applied linguistics: Variations across postgraduate genres. *Journal of Foreign Language Teaching and Translation Studies, 2*(2), 1-29.

Jalali, H., EslamiRasekh, 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., EslamiRasekh, A., & TavangarRizi, M. (2009). Anticipatory 'it' lexical bundles: A comparative study of student and published writing in applied linguistics. *Iranian Journal of Language Studies, 3* (2), 177-194.

Jalali, H., & Ghayoomi, S. (2010). A comparative qualitative study of lexical bundles in three academic genres of applied linguistics. *Modern Journal of Applied Linguistics, 2* (4), 323-333.

Scott, M. (2008). *Wordsmith Tools 5.* Oxford: Oxford University Press.

Thompson, G. (2001). Interaction in academic writing: Learning to argue with the reader. *Applied linguistics, 22*(1), 58-78.

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