Journal of Foreign Language Teaching and Translation Studies,
Vol. 3, No. 1, January 2014

A Comparative Study of Anticipatory *it* Lexical Bundles in Applied Linguistics and Analytical Chemistry Research Articles

Hassan Jalali
English Language Center, Isfahan University of Technology
h.jalali@cc.iut.ac.ir

Abstract

Lexical bundles, frequent word combinations that commonly occur in different registers, have attracted researchers in corpus linguistics in the last decade. While most previous studies of bundles have been mainly concerned with variations in the use of these word sequences across different registers, very few studies have focused on their use across disciplines. To address possible disciplinary variations in the use of these word combinations in academic register, this quantitative and qualitative study chose to investigate, compare, and contrast range, frequency, and function of anticipatory *it* bundles as a sub-set of these word clusters with important metadiscursive functions in published writing. For this purpose, the study zoomed in research articles of applied linguistics and analytical chemistry as a soft and hard science, respectively. The results indicated that generally anticipatory *it* bundles could be regarded as a distinctive characteristic of academic writing in both disciplines. At the same time, each discipline was found to draw almost on a particular set of *it* bundles in the development of its discourse. However, analytical chemistry seemed to rely more on these bundles in the development of its discourse. Functional analysis also showed that *it* bundles served a wide variety of functions in both disciplinary areas. Therefore, this study called for a more robust pedagogical focus on different multi-word sequences like anticipatory *it* lexical bundles. The findings also highlighted the importance of a more genre-focused EAP (English for academic purposes).

Keywords: corpus linguistics, analytical chemistry, applied linguistics, research articles, anticipatory *it* lexical bundles

Introduction

The study of different formulaic word combinations has a history of more than five decades (Cortes, 2002). Among diverse categories of formulaic sequences, lexical bundles, also known as clusters and chunks

Received: 02/06/2013 Accepted: 03/10/2013

(Hyland, 2008a, 2008b), were first introduced and defined by Biber, Johansson, Leech, Conrad, and Finegan (1999) in their well-known English grammar as a group of word combinations with important functions. 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 most salient and defining characteristic of bundles; 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 or repetitive uses. To go further, Lexical bundles are identified empirically just 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 (e.g., the meaning of a four-word bundle like *in the presence of* is almost easily retrievable form the meaning of its individual parts) are other features of bundles. Lexical bundles have been found to be an important part of academic discourse (Biber et al., 1999).

Such word sequences have been classified structurally (Biber et al., 1999; Biber, Conrad, & Cortes, 2004; Biber, 2006; Jalali, Eslami Rasekh, & Tavangar Rizi, 2008, 2009) as well as functionally (Cortes, 2002, 2006; Biber, Conrad, & Cortes, 2003; 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 corpus-based and mostly comparative 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 & Conrad, 1999; Biber et al., 2004, Biber & Barbieri, 2007), genres (Hyland, 2008b; Jalali, 2013), and different degrees of writing expertise (Cortes, 2002, 2004; Levy, 2003; Jalali, 2009; Jalali et al., 2008, 2009). Overall, these studies have indicated that lexical bundles are strong discipline, genre, and register discriminators (Biber et al., 1999). This means that apart from some overlaps, each discipline, genre, or register draws on its own specific 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 stressed that many lexical bundles favored by experts in any given disciplinary area may not be used by novices who could be students or developing writers with varying degrees of language proficiency and disciplinary expertise (see for example Cortes, 2004, 2006).

Interestingly, there is also usually a correlation between structural type of bundles and the function they serve in the discourse (Biber et al., 2004). Anticipatory *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; Jalali et al., 2009) or expressions of stance (Biber, 2006). The structure of these bundles is made of anticipatory *it*, *is* (as a present copula verb), a predicative adjective (*e.g*., *necessary*) and one of two complementizers, *to* (as in *it is necessary to*) or *that* (as in *it is clear that*). The clause-initial anticipatory *it* is often part of a multi-word fixed word combination or bundle which can act like a frame within which the following propositional meaning could be embedded (Biber & Barbieri, 2007). A straightforward and possibly accessible way for writers to express their stance toward the reader, and the content could be the use of those bundles beginning with anticipatory *it*, copula *is*, a predicative adjective (e.g., *interesting*, *necessary*, *clear*, etc.) followed by a subordinate clause which is usually introduced with complemetizers *to* or *that* (*e.g*., *it is interesting to*, *it is possible that*). Such bundles seem to convey a range of epistemic, evaluative, and attitudinal meanings (Jalali et al., 2009). Most extraposed complement clauses beginning with anticipatory *it* can also reflect the speaker or writer's assessment (Hewings & Hewings, 2002).

There are different reasons for the use of this type of *it* clausal bundles (Hewings & Hewings, 2002). Grammatically, there is a marked tendency in English to put the longer subjects at the end of the clause (Quirk, Greenbaum, Leech, & Svartvik, 1985). Thematically, elements with higher information usually occur at sentence final position (Hasselgard, Johansson, & Lysvag, 1998). From a metadiscursive point of view, the use of *it* bundles enables the writer to distance herself or himself away from the propositional content and thereby project an objective and impartial persona (Quirk et al., 1985; Hyland, 2004). One the other hand, by embarking on such a structure, the writer is able to evaluate the ensuing proposition (Hunston & Sinclair, 2000), and finally to depersonalize the opinions (Hewings & Hewings, 2002) as can be seen in the following examples taken form the corpus of applied linguistics research articles used in this study:

(1) First, *it is important to* help student-teachers to look at teaching through multiple but complementary lenses.
(2) *It is necessary to* bear in mind the mix of regional and national languages and accents that the participants represent as well as those that they are familiar with.

Biber et al. (1999) have shown 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). According to Hewings and Hewings (2002), clauses starting with an anticipatory *it* 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 credibility of the propositional meaning), and attribution (convincing the reader through a general or specific reference).

However, very few studies have focused on the use of anticipatory *it* bundles within high-stake genres of academy (see Hewings & Hewings, 2002; Hyland, 2008a; Jalali et al., 2009). Especially striking is the scarcity of studies that would address specific phraseological practices in different disciplinary areas with an aim to describe and explain possible differences and/or similarities between experts in their use of these word combinations. It must be noted that his structural group of lexical bundles was investigated in this study for two reasons. First, there is some evidence to suggest that for many non-native speakers of English, this structure can pose serious degrees of difficulty mostly because of the absence of anticipatory *it* structure in some languages (Jacobs, 1995; Hewings & Hewings, 2002). Second, recognizing the importance of this structure as a metadiscursive element or a stance expression, the study seeks to identify the range of interpersonal meanings conveyed by these word clusters as such *it* bundles are usually good means by which writers can express their opinions, evaluate the subject matter, and engage with readers (Hewings & Hewings, 2002).

The purpose of this study, therefore, was to compare the use of one structural class of bundles in research articles as a key written academic genre within two disciplinary areas of applied linguistics and analytical chemistry. This study, therefore, addressed the following questions:

1. What are the most frequent four-word anticipatory *it* lexical bundles in analytical chemistry research articles?

2. What are the most frequent four-word anticipatory *it* lexical bundles in applied linguistics research articles?
3. To what extent is there evidence to support similarity or contrast in the range, frequency, and function of anticipatory *it* lexical bundles across the two disciplines?

Method

Corpora

Two corpora were used in this study. The first corpus consisted of research articles in applied linguistics, representing published writing in this disciplinary area. The second corpus included published writing in the discipline of analytical chemistry. Each of these corpora will be described more below.

The first corpus had been originally prepared and used by Jalali (2009) and Jalali et al. (2008, 2009) for their study on variations in the use of all lexical bundles within applied linguistics. The basis for the selection of journal articles was mostly previous corpus-based studies done on the discipline of applied linguistics as well as the advice given by experts in the field and access to the electronic files of journals (Table 1). Applied linguistics was chosen in this study to stand as a kind of soft science to compare with a member of hard science, analytical chemistry.

Table 1. Corpora used in the study

Corpora	No. of journals	No. of texts	No. of words
Corpus of applied linguistic	7	201	1217963
Corpus of analytical chemistry	8	270	1003901

The second corpus**,** the corpus of analytical chemistry research articles, included analytical chemistry articles from 8 journals in this disciplinary area. This corpus was part of a larger corpus of 4 million words in the discipline of chemistry prepared by Valipoori Goodarzi (2010) that covered four main sub-disciplines of chemistry: analytical chemistry, inorganic chemistry, organic chemistry, and physical and theoretical chemistry. The reason for choosing analytical chemistry among other sub-disciplines was the large number of texts in this corpus and the wider variety of journals from which these texts had been taken.

Data Analysis Tools

Computer programs

Two computer programs were used in this study: Antconc3.2.1w (Anthony, 2007), and Wordsmith5 (Scott, 2008). The former was used for the identification of lexical bundles and concordancing while the

latter was only used to find the number of texts within which each bundle had been used. Through the use of Antconc3.2.1.w, all lexical bundles in the two corpora with their actual frequencies were found and displayed by inserting a set of commonly key words with which the bundles collocated such as prepositions (in this study, only anticipatory *it*), and deciding on the minimum optimal frequency (*e.g.*, twenty in a corpus of one million words) and specifying the required number of words in clusters (i.e., three, four, five, or six) (see Figure1). It must be mentioned here that in this study, like that of Biber et al. (1999), the frequency cut-off of ten was chosen. A four-word combination starting with anticipatory *it* had to occur at least ten times and in five different texts to be included in the study.

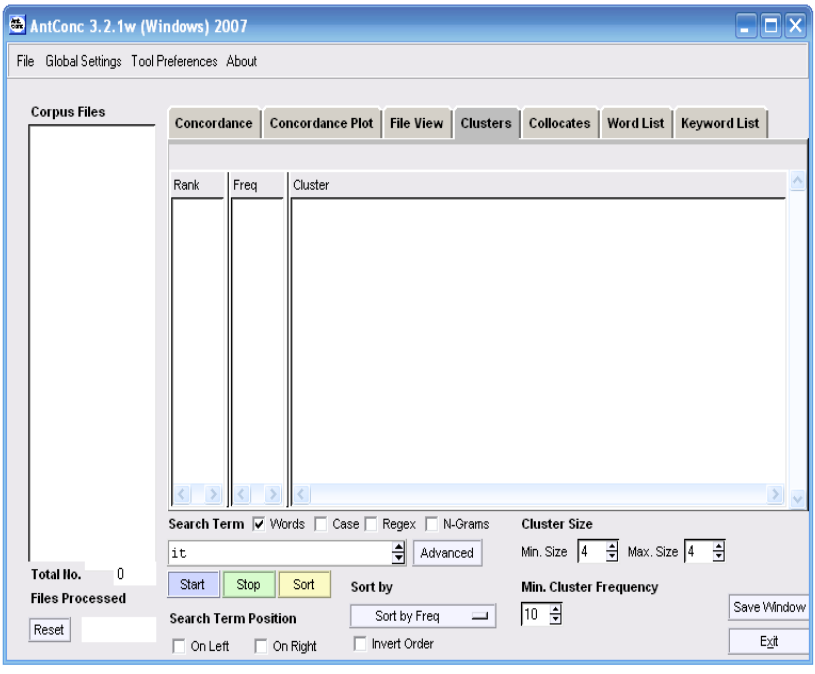

Figure 1. Tools of Antconc 3.2.1. software

When all candidate lexical bundles were identified by the first computer program, each of them was again searched on Wordsmith tools5 to find the number of texts in which they have 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). This was to guard against idiosyncratic and repetitive use of the same bundle in the same text by the same writer (see Figure 2).

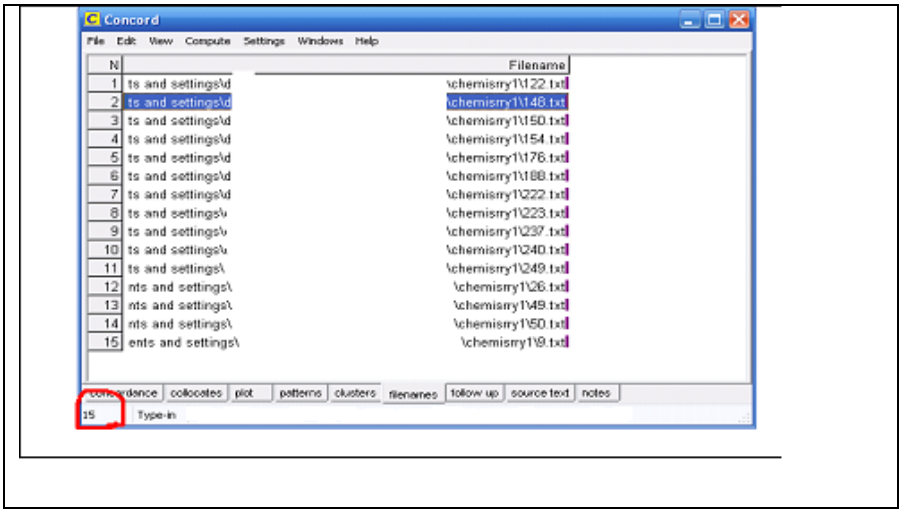

Figure 2. Exact number of texts

Analyses

The data were analyzed in three steps. First, through the two computer programs, all anticipatory *it* lexical bundles of interest were identified in the two corpora along with their actual frequencies and the number of texts in which they had been used. Second, using the functional typology of *it*-clauses developed by Hewings and Hewings (2002) (see Table 2) and examining the actual use of anticipatory *it* bundles through concordancers in the respective contexts, the researcher decided on the most predominant functions to which they had been put. In the third stage, the results were compared to determine the extent to which research articles in the two disciplines of applied linguistics and analytical chemistry might be different from each other in terms of range, frequency, and function of this group of bundles.

It must be noted that there are already some functional classifications of lexical bundles (e.g., Biber & Conrad, 1999; Cortes, 2002; Biber et al., 2004; Hyland, 2008a, 2008b). Hewings and Hewings's (2002) functional taxonomy of *it*-clauses was used in this study since it specifically dealt with interpersonal functions of clauses starting with anticipatory *it*. However, as the developers of this model confirmed

themselves, no functional classification of language could be totally objective and watertight; therefore, there was a good degree of subjectivity in functional grouping, and no clear-cut divisions between all categories, especially between the second , attitude markers, and the third one, emphatics could be firmly established.

Table 3. Interpersonal Functions of it clauses (Hewings & Hewings, 2002: 372)

Interpersonal functions *It*-clauses of	**Subcategories**	**Example 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)

Furthermore, as the corpora become more specific, there emerged a need for developing new categories to cater for new functions not

observed before. So, based on the analysis of *it* bundles, a new category (i.e., epistemic) was developed to describe a new function not present in the original taxonomy.

Results

Lexical Bundles in Published Writing in Analytical Chemistry

As can be seen in table 3, there were nineteen different anticipatory "it" lexical bundles in the corpus of analytical chemistry research articles. In terms of variety, this corpus had more bundles than the corpus of applied linguistics research articles.

Table 3. Anticipatory it lexical bundles in the corpus of analytical chemistry research articles

Lexical Bundles	Frequency	Number of texts
it can be seen	66	39
it is possible to	54	39
it was found that	52	37
it should be noted	42	23
it is important to	38	28
it is well known	36	29
it has been shown	28	24
it can be concluded	23	19
it is known that	21	19
it was observed that	20	13
it is expected that	19	17
it is clear that	17	15
it has been reported	16	14
it is necessary to	16	14
it is obvious that	15	10
it is difficult to	14	15
it is seen that	13	10
it is possible that	12	11
it is noteworthy that	10	8

It can be seen that, it is possible to, it was found that, and *it should be noted* were some of the most frequent bundles in this corpus. The overall frequency of all bundles was 512, accounting for 0/05% of the whole corpus. As for functions, it seemed that anticipatory "it" lexical bundles had a relatively considerable range of use among published writers in analytical chemistry (see Table 4). A large number of these bundles were those starting with anticipatory "it" followed by copula "is" and an

adjective and complementizer "that" or infinitival "to" (e.g., *it is possible to, it is important to, it is clear that ,etc.)*. From table 4, it can be seen that epistemic category with a portion of 33.39% was the most predominant function. This category, which was not used in the study conducted by Hewings and Hewings (2002), was borrowed from Cortes (2002, 2004) and described that function of bundles by which they reflected the factual status of the ensuing propositions (e.*g., it is known that*). This function was followed by attribution (24.02%), emphatics (17.57%), hedges (12.89%), and attitude markers (12.1%). The following examples can show the use of some of these bundles in their actual contexts of use:

(1) From the above figure, *it is possible to* observe that the proposed circuit fits better the experimental data in the whole range of frequencies studied (hedge).

(2) *It is important to note that*, in contrast to probe-based techniques, mass spectrometry-based determination of base composition does not require prior knowledge of the composition in order to make the measurement, only to interpret the results (attitude marker).

(3) *It is difficult to* find a balance between the health benefits and risks stemming from fish consumption or even, indeed, to draw any conclusion about this issue (attitude marker).

(4) Moreover, *it should be noted* that the frequencies of the calculated bands were obtained using the harmonic approximation, whereas the experimental frequencies include some anharmonicity by nature (emphatic).

(5) From Fig. 2, *it can be seen* absorption spectra of the complex in aqueous solution and the ionic liquid phase have a very similar shape, this show the extraction does not influence the complex's composition (attribution).

(6) *It is well known* that a wide variety of organic molecules can intercalate between the interlayer regions of expandable clays [5], and clays have recently been investigated as drug carriers [2, 6] (epistemic).

Table 4. Overall functional description of it-bundles in the corpus of analytical chemistry research articles

Subcategories	Number	Frequency	Percentage %
Hedges:			
1a likelihood/possibility/ certainty; importance/value/necessity etc.	2	66	12.89
1b what a writer thinks/assumes to be//will be/ was the case	0	0	0
Attitude markers:			
2a the writer feels that something is worthy of note	1	10	1.95
2b the writers evaluation	2	52	10.15
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	42	8.20
3c the writer expresses a strong conviction of what is possible/ important/necessary, etc.	3	48	9.37
Attribution:			
4a specific attribution (with a reference to the literature)	4	123	24.02
4b general attribution (no referencing)	0	0	0
Epistemic:			
5a Certain	2	57	11.13
5b uncertain	2	42	8.20
5c impersonal	2	72	14.06
total	19	512	100

Lexical Bundles in Published Writing in Applied Linguistics

Table 5 shows anticipatory "it" lexical bundles in the corpus of published writing in applied linguistics along with the frequency and the number of texts in which they have been used. A total of seventeen different *it*-bundles with different functions were found in this corpus. The overall actual use of these bundles was 449, accounting for 0/036% of the whole corpus. In terms of function, this corpus drew heavily on attitude markers (43.20%) and minimally on attribution markers (3.80%) (see Table 6). Some of the most frequent it-bundles were: *it is important to*, *it should be noted*, *it is possible that*, and *it is difficult to*. A large number of

anticipatory "it" lexical bundles in this corpus also had the pattern of *it +Vbe+ adjective + that/to*. It seemed that the use of such bundles by published writers in applied linguistics was a good means by which they encoded different interpersonal meanings (Jalali, 2009) as can be seen below:

Table 5. Anticipatory it lexical bundles in the corpus of applied linguistics research articles

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

(7) 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 (hedge).

(8) *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; others seemed rather more open to acknowledging more personal gains (hedge).

(9) *It seems that the* 2004 version of the CARS model successfully accounts for most of the limitations mentioned above (hedge).

Table 6. Overall functional description of it-bundles in the corpus of applied linguistics research articles

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

(10) *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 (emphatic).

(11) *It is important to notice* the scare quotes around 'unit'—this term is itself a metaphor, bringing with it a sense of concreteness, solidity,

and discreteness that may not be appropriate in a complex dynamic systems or emergentist framework (emphatic).

(12) 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 acquisitional processes involved in the development of implicit L2 knowledge from the general deductive learning strategies involved in the development of explicit knowledge (attitude marker).

(13) *It is clear that* the trend towards informality in academic writing has undermined old certainties and accepted conventions and this not only adds to the 'burden' of non-native speaker writers of English, but also to the difficulties of non-native speaker teachers too (attitude marker).

Comparisons

Comparisons in Terms of Variety and Frequency of Bundles

Probably, the most surprising finding of this study was the similarity between the two corpora under investigation in terms of the range of it-bundles employed. The results showed that in the corpus of analytical chemistry research articles, there were nineteen it-bundles while in the corpus of applied linguistics articles, they were seventeen.

Since the two corpora used in this study were not parallel in size, a normalization procedure had to be employed in order to make it possible to compare the corpora in terms of overall frequency of bundles used (Jalali et al., 2008, 2009). It should be, however, noted that the use of such a procedure could not always give a very true picture of the results and there are some reservations against it (Biber et al, 1999; Biber & Conrad, 1999; Cortes, 2002, 2004; Biber, 2006), but to make the comparisons possible, it was inevitable to employ it. So this normalization procedure was utilized to obtain the overall frequency of lexical bundles in each of the two corpora in a scale of one million words. It was quite easy to run this normalization procedure. First, the actual total frequency of all bundles in a corpus was multiplied by one million and then the result was divided by the actual size of each corpus. It must also be noted that this procedure had already been used in some previous corpus-based studies of lexical bundles such as Cortes (2002, 2004), Biber et al. (2004), Biber and Barbieri (2007), and Hyland (2008a, 2008b).

The results obtained in this way first showed that overall frequency of it-bundles in analytical chemistry research articles was considerably

more than that of applied linguistics published writing (see Table 7). Therefore, the overall frequency of it-bundles in analytical chemistry research articles was 38% more than that of applied linguistics published writing. It seems, therefore, that analytical chemistry published writing was ahead of applied linguistics not only in the variety of bundles used but also, in the overall use of these word combinations.

Table 7. Variety and overall use of it-bundles in the two corpora

Disciplines	Chemistry	Applied Linguistics
Number of bundles	19	17
Actual frequency	512	449
Normalized frequency	510	368

In terms of variety of it-bundles used, out of 19 different bundles employed in chemistry research articles, 47% were used in the other corpus. On the other hand, out of 17 bundles used in applied linguistics research articles, 53% were used in analytical chemistry. Table 8 shows shared it-bundles in these two corpora. As can be seen, half of the bundles were used in published writing of both disciplines.

Table 8. Shared it-bundles in the two corpora

Lexical bundles	**Frequencies in**	
	Chemistry	**Applied Linguistics**
it can be seen	66	17
it is possible to	54	25
it was found that	52	19
it should be noted	42	40
it is important to	38	88
it is clear that	17	33
it is necessary to	16	34
it is difficult to	14	38
it is possible that	12	38
total	311	330

Comparisons in Terms of Functions of Bundles

Table 9 shows the varieties, overall frequencies (normalized in one million), and percentages of anticipatory "it" lexical bundles in terms of the five functional categories used in this study based on Hewings and Hewings (2002) functional taxonomy of it-bundles and the category of epistemic which was added in this study (see Jalali, 2009). As for disciplinary differences in the variety of bundles used in each major functional category, analytical chemistry research articles outweighed

their applied linguistics counterparts in the range of all it-bundles employed. However, while the variety of bundles serving as attribution and epistemic in the former was more than latter, the variety of it-bundles serving as hedges and attitude markers in the latter was more than that of the former.

Table 9. Functional comparison of it-bundles in the two corpora (analytical chemistry/ applied linguistics)

Categories	Number of bundles	Frequency (normalized)	Percentage %
Hedges	2/5	65/80	12.74/21.73
Attitude markers	3/6	62/159	12.15/43.20
Emphatics	4/3	90/88	17.64/23.91
Attribution	4/1	123/14	24.11/3.80
Epistemic	6/2	170/27	33.33/7.33/
Total	19/17	510/368	100/100

This was in accordance with the findings of some previous research (e.g., Hyland, 1996, 1999, 2000; Thompson, 2001) that had attested the frequent use of hedging devices in humanities and social sciences. The use of attitude markers in applied linguistics published writing was also much heavier than the other corpus. As can be seen in Table 9, the overall use of attitude markers in applied linguistics research articles was three times more than that of chemistry. There were some 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,* which was used twice more than the other corpus. *It is difficult to* was another bundle used more heavily by applied linguistics published writers.

The absence of the subcategory (3a) of emphatics, where the writer indicated a conclusion or deduction should be reached and a proposition was true, in both corpora was also noteworthy. Surprisingly, the use of attributive markers in chemistry was much more that of applied linguistics in terms of range and frequency of use. While *it can be seen* was the only bundle used in the two corpora, there were other it-bundles (i.e., *it has been shown, it has been reported, it is seen that*) that were just used by published writers in analytical chemistry. The analysis of *it can be seen* also showed that in terms of the textual references to which this bundles was put, there was more diversity in analytical chemistry corpus. Bundles expressing epistemic meanings had a relatively wide

range in the two corpora. *It was found that* was the most frequent bundle of this type in the two corpora. The heavier use of this bundle in chemistry corpus was also noticeable. Generally, the use of epistemic bundles in chemistry was more than that of corpus of applied linguistics articles. At the same time, except for *it was found that*, each corpus had its own repertoire of bundles for epistemic functions.

Conclusions

In line with findings of some previous research (e.g., Biber et al, 1999; Cores, 2002, 2004; Biber et al., 2004; Hyland, 2008a, Jalali, 2009; Jalali et al., 2008; 2009), this study showed that anticipatory *it* lexical bundles are generally a distinctive feature of academic writing in both hard and soft disciplines. Also, similar to the findings of fewer studies on the use of bundles in disciplinary writing (Cortes, 2002, 2004), this study showed that the use of lexical bundles in hard sciences seems to be more than that of soft sciences. This could be due to the more formulaic nature of expressions in the former (Hyland, 2008a).

It seems, therefore, necessary for EAP practitioners in general and academic writing instructors in particular to leave a good space in their instruction for a more pedagogically focused treatment of anticipatory *it* bundles (Jalali, 2013), which are for the most part a characteristic of academic writing (Biber et al., 1999). The use of noticing (Cortes, 2004, 2006), conscious raising tasks (Lewis, 2000a, 2000b), clusters lists, and concordances (Hyland, 2008a) could be some of the means by which students could come to a possibly better understanding and more frequently appropriate use of these word combinations.

References

Anthony, L. (2007). Antconc 3.2.1: A free text analysis software. Available on line at http://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. (1999). Lexical bundles in conversation and academic prose. In H. Hasselgard & S. Oksefjell (Eds.), *Out of corpora: Studies in honor of Stig Johansson* (pp. 181–189). Amsterdam: Rodopi.

Biber, D., Conrad, S., & Cortes, V. (2003). Lexical bundles in speech and writing: An initial taxonomy. In A. Wilson., P. Rayson, & T.

McEnery (Eds.), *Corpus linguistics by the lune: A festschrift for Geoffrey Leech* (pp. 71-92). Frankfurt/Main: Peter Lang.

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* Harlow: Pearson.

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.

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

Hasselgard, H., Johansson, S., & Lysvag, P. (1998). *English grammar*. Oslo: Universitetsforlaget.

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*, *21*, 367-383.

Hunston, S., & Sinclair, J. (2000). A local grammar of evaluation. In S. Hunston, & G. Thompson (Eds.), *Evaluation in text* (pp. 74–101). Oxford: Oxford University Press.

Hyland, K. (1996). Writing without conviction? Hedging in scientific research articles. *Applied Linguistics*, *17*, 433-454.

Hyland, K. (1999). Talking to students: Metadiscourse in introductory coursebooks. *English for Specific Purposes*, *18*, 3-26.

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

Hyland, K. (2004). Disciplinary interactions: Metadiscousre in L2 postgraduate writing. *Journal of Second Language Writing*, *13*, 131-151.

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., Eslami Rasekh, A., & Tavangar Rizi, M. (2008). Lexical bundles and intradisciplinary variation: The case of applied linguistics. *Iranian Journal of Language Studies, 2* (4), 447-484.

Jalali, H., Eslami Rasekh, A., Tavangar Rizi, 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.

Levy, A. S. (2003). Lexical bundles in professional and student writing (Unpublished doctoral dissertation). University of the Pacific, San Francisco.

Lewis, M. (2000a). Learning in the lexical approach. In M. Lewis (Ed.), *Teaching collocation. Further developments in the lexical approach* (pp. 155–185). Hove: Language Teaching Publications.

Lewis, M. (2000b). Materials and resources for teaching collocation. In M. Lewis (Ed.), *Teaching collocation. Further developments in the lexical approach*. Hove: Language Teaching Publications.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language.* London: Longman.

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.

Valipoori Goodarzi, L. (2010). *A corpus-based study of words and bundles in chemistry research articles* (Unpublished master's thesis). University of Kashan, Iran.