Teaching English as a Second Language Quarterly (TESLQ)
(Formerly Journal of Teaching Language Skills)
44(2) 2025, pp. 1-23
https://doi.org/10.22099/tesl.2025.50599.3306
Online ISSN: 2981-1546
Research Paper

It-Lexical Bundles Revisited: The Role of Disciplinary Variations and First Language

Nahid Shahmoradi [1] **Hassan Jalali** [2 *]

Abstract

Lexical bundles, starting with the anticipatory *it*, mostly serve as stance expressions (e.g., *it is necessary to*). While most research centers on anticipatory, *it-bundles* in one single discipline, not much has been done across different disciplinary areas. Therefore, based on an adaptation of the functional taxonomy proposed by Hewings and Hewings (2002), this research attempted to detect *it-bundles* using a corpus of 400 research papers in L1-English and L1-Persian in applied linguistics (AL) and information technology (IT) to probe the possible significant resemblances and disparities. According to the results, IT writers employed fewer bundles than their AL counterparts, and their overuse was more impressive than that of their L1-English peers in AL. However, AL and IT writers showed similarity in their use of functional categories; AL writers also made heavier use of two functional categories: emphatic and epistemic. As for practice, writing instructors can exploit the findings of this study to facilitate academic writing instruction. They can also help students to achieve a better comprehension of anticipatory *it-bundles*.

Keywords: Applied Linguistics; Information Technology; *It-bundles*; Corpus; Function

Biber et al. (1999), for the very first time, introduced lexical bundles as complex phrases serving diverse functions in different registers as word combinations. These scholars interpreted these phrases as "recurrent word expressions regardless of their idiomaticity and regardless of their structural status" (p. 990). They stated that these complex phrases are utilized in different registers as word combinations. Previously, Jespersen (1927) and Firth (1951,1957) had specified these phrases as special collocations. These sequences of words are fixed expressions that have a notable contribution to communication (Cortes, 2002, 2004). Biber (2007) also defined these sequences as the most frequently continuous arrangements of terms. Readers can acknowledge the denotations of these complex phrases by realizing the meaning of each component. They can also help to acquire communicative competence and experience fluent expression. In addition, these expressions promote the learning of pragmatic competence (Jones & Heywood, 2004).

* Review History:
Received: 04/07/2024 Revised: 13/01/2025 Accepted: 21/01/2025

[1] Instructor, Center of English Language, Isfahan University of Technology, Isfahan, Iran; nahid.shah1995@gmail.com
[2] Assistant Professor, Center of English Language, Isfahan University of Technology, Isfahan, Iran (Corresponding Author); h.jalali@iut.ac.ir

How to cite this article:
Shahmoradi, N. and Jalali, H. (2025). It-Lexical Bundles Revisited: The Role of Disciplinary Variations and First Language. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills)*, *44*(2), 1-23. doi: https://doi.org/10.22099/tesl.2025.50599.3306

Several studies have previously evaluated the usage of these phrases, their structures, and functions in the research papers of diverse academic fields across different levels of writing (Ädel & Erman, 2012; Chen & Baker, 2010; Conklin & Schmitt, 2008; Durrant, 2017; Fan et al., 2016; Herbel & Wagner, 2010;). In this regard, some research works have addressed these phrases in different disciplines; for example, we can note the studies done by Kashiha and Chan (2014) and Farvardin, Afghari, and Koosha (2012). Meanwhile, some other investigations have explored these complex phrases in different academic pieces of writing by particularly approaching the L1/ L2 distinction (Cortes, 2004; Jhang, Kim & Qi, 2018; Wei & Lei, 2011).

In academic fields, learners who have mastered the use of these structures gain fluency in writing. The use of frequent words and expressions could be regarded as proof of the disciplinary expertise of the members of the respective communities. Lexical bundles also make a substantial contribution to English for specific purposes (ESP), as they highlight the differences between research written by professional authors and that written by novices using discipline-specific frequent word expressions (Hyland, 2008). Thus, the lack of lexical bundles in writing could reflect that the writer(s) may not be able to foster involvement in academics.

Wei and Lei (2011) have also highlighted how complex phrases are employed by advanced Chinese EFL learners in academic writing. They shed light on the doctoral essays by the learners and research articles published by experienced authors. Novice authors tended to use bundles more frequently than their experienced counterparts. Moreover, Hyland (2008), based on an examination of expert and novice writing, pinpointed the differences in terms of discipline-specific word expressions. The findings have also revealed that some bundles are specific to expert writers, while students with different degrees of proficiency use other bundles (Cortes, 2006).

Biber and Barberi (2007) have also mentioned that clauses starting with an anticipatory *it* are part of the phrases regarded as a structure in which the meaning is inserted. Analyzing these bundles is a way by which writers express their stance toward the readers. In this regard, Hewing and Hewing (2002) functionally classified *it-bundles* into different categories, including hedges, which refer to the authors' speculative attitude against the prepositions; attitude markers, which express writers' attitude toward the content; emphatics, which define the extent to which the writers are certain about the prepositional meaning and finally, attribution, which specifies something already presented, such as tables or figures and something which is inferred and interpreted from the tables or conclusions already drawn.

Overall, while many researchers have concentrated on the use of it-bundles, few studies have examined the similarities and differences among different disciplines. It is important to note that not much in-depth research has been carried out to evaluate it-bundles as sequences that can contribute to specific phraseological practices in academic writing, such as research articles. Therefore, the current study attempted to address this disparity in articles between two different disciplines of Applied Linguistics (AL) and Information Technology (IT) among L1-English and L1-Persian writers, both structurally and functionally. Before we proceed to

describe our study, a review of the relevant literature is presented to show the state-of-the-art in the line of research.

Review of the Literature

Lexical bundles are a sequence of two or more words that co-occur naturally with high frequency, but their most important feature is that they are not idiomatic in meaning; this means that their meaning is relatively transparent. They are also fixed phrases in structure. It-lexical bundles, on the other hand, are phrases that start with an anticipatory *it* followed by a clausal subject. This structural class of bundles is typically used to make neutral statements and express writers' attitudes (Tove Larsson, 2017). Lexical bundles also play a very important role in academic prose, building the meanings in texts and transferring the communicative purpose of the writers to represent a more meaningful and coherent text.

Functionally, according to Hyland (2008), there are four types of lexical bundles: referential bundles, text organizers, stance bundles, and interactional bundles. Referential bundles are used as references to textual or external entities. Text–oriented bundles, also known as text organizers, include the phrases used for transition, resultative, structuring, and framing signals. Stance bundles indicate feelings, attitudes, and judgments. The last type refers to the combination of words that express politeness.

The present study aimed to develop a more functional classification for it-bundles in native and non-native studies to highlight the importance of these patterns in written academic discourse. These patterns are regarded as problematic for non-native writers, as some studies reported the underuse or overuse of bundles (Thompson, 2009). One of the main reasons why these patterns can be problematic for non-natives is that they do not have any equivalent in many languages like Persian (Jalali, 2017). Persian writers may, therefore, find it difficult to use these types of clusters. Thus, analyzing these clusters can address the differences between writers cross-linguistically.

Meanwhile, the survey of stance expressions known as explicit lexico-grammatical features has been of interest to many linguists as they have a good potential to show the writer and reader's relationships and express their attitudes and feelings. In the last decades, linguists have done research using such terms as 'hedging' (Brown & Levinson,1987; Hyland,1996a, 1996b), 'evidentiality' (Chafe & Nichols,1986),'evaluation' (Hunston & Thompson, 2000), 'intensity' (Labov,1984), and 'stance' (Biber et al.,1999). Accordingly, stance features have gained a lot of attention. Biber (2006) and Biber and Finegan (1988), along with Hyland and Guinda (2012) and Gray and Biber (2012), have made a clear and useful distinction between meanings that indicate writers' attitudes and the epistemic features of an entity. The former refers to affective meanings, including positive and negative ones (e.g., i*nterestingly*, *unfortunately*, and *hopeful*). Meanwhile, the latter refers to evidential meanings that represent certitude (e.g., *impossible*, *could*) and suspicion (e.g., *should, may*).

Analyzing it-bundles not only makes a distinction between meanings that indicate writers'

attitudes but also shows the epistemic features of entities. The former refers to affective meanings, including positive and negative ones such as *interestingly*, *unfortunately*, and *hopefully*. Meanwhile, the latter refers to evidential meanings that represent certitude (e.g., *impossible*, *could*) and suspicion (e.g., *should*, *may*).

Using a lexico-grammatical approach employing special tools and analyzing the lexical bundles, specially it-bundles, is regarded as a useful method to explore particular attitudinal, epistemic, and meanings that encode stance, with each type of text presenting its own specific bundles to express stance and referring to different parts of the texts to build a cohesive discourse. These clusters of coherent discourse organize the blocks of stance expressions and references to textual or external entities. Each structure of bundles, such as it-bundles addressed in this study, serves different discourse functions, usually acting as an expression of stance.

Biber (2006) also broadened the expressions of stance in both spoken and written registers, including classroom instructions, class speech, course books, and written discourse. Interestingly, it has been specified that all stance expressions are rarely used in textbooks, as compared to those employed in spoken university registers. Despite the expectations, all stance expressions are employed regularly in speaking rather than writing. This has been somewhat surprising as textbook authors are expected to be more precise, and writing is supposed to be more accountable (Gray & Biber, 2012). However, all stance devices have been less frequent in textbooks.

Additionally, many authors have been interested in analyzing hedges and boosters, attitude markers, and self-mention (Hyland, 2005). Hedges include words like *might* and *possible*. They indicate a commitment to a preposition that writers have when using these devices. Further, all the information is considered as an idea rather than a fact. Boosters include expressions like *clearly* and *obviously*, by which writers demonstrate their assurance toward what they say. Also, they represent how they are involved with the topic and the extent to which they agree with the audience. To further elaborate, attitude markers show the affective attitude of writers towards prepositions, unity, value, and so forth (Hyland, 2005). Self-mention, alternatively, brings up first-person pronouns or possessive adjectives that writers use to convey prepositions, emotions, and social knowledge (Hyland, 2001). Most of the previous studies have, therefore, evaluated the L2 English academic writings, and there has been almost no in-depth analysis of L1 and L2 learners in terms of *it*-bundles.

Some researchers have also found the distinctions between writers' L1 a useful analysis to reveal how they use lexical bundles (Shin, 2019). Fen Jiang (2015), for instance, evaluated these combinations of words in Chinese and American students' theses with the same age and level of education. Accordingly, their results indicated that L2 students employed them less than their L1 counterparts. However, their heavy use of evaluative stance in the attribute category was marked. This study also discussed that L2 students faced difficulties when using noun-complement, which could be due to their limited writing lexicon. Another corpus-driven study zoomed on the use of it-bundles by examining applied linguistics master theses and doctoral

dissertations. The results indicated that doctoral students applied more bundles to develop meaning in their texts. However, bundles were used infrequently in both genres (Jalali, 2015).

Further, Larsson (2017) used three corpora of student writing, including the advanced learner English corpus, the Michigan Corpus of Upper-Level Students' Papers, and the British Academic Written English Corpus. His study mapped out the functional distribution of the patterns, such as attitude markers (e.g., *it is surprising*) and hedges (e.g., *it seems like*). The results revealed that non-native speakers did not equally use the patterns; this finding was found to be due to their problems with certain functional categories. Concerning the academic differences, remarkable points were found. For example, linguistic papers outweighed the introductory patterns, showing difficulty, expectations, and importance compared to literature papers. Interestingly, non-native speakers, even with a lower level of language proficiency, used more attitude markers than non-native speakers at a higher level. However, native speakers and non-native speakers at higher levels used more frequent bundles in this study. Therefore, students who use bundles frequently do not necessarily display proficiency in using them.

Comparing the general novice academic stance, Lancaster (2016) found contextual specificity in two high and low-graded papers written in two undergraduate courses in the United States. His study revealed that high-graded papers represented a greater frequency of stance than low-graded papers. A closer look at the study marked the higher use of hedges, boosters, and attitude markers by the high-performers in comparison with their corresponding low-performers. Importantly, the high-performers used these devices that reflected "critical thinking", which expressed disagreement between writers, theories, or approaches. Lancaster also focused specifically on critical thinking because instructors could spell out the stance qualities in the process of learning to enhance the student's knowledge about the value of the stance orientations. However, the results of another study (Staples et al., 2013) indicated that the lower level learners used more bundles.

Jalali (2017) also argued the importance of stance expressions as linguistic devices, which indicated the relationship between the reader, the writer, and the prepositional meaning in three corpora of research articles, doctoral dissertations, and master theses in applied linguistics. In contrast to many other studies that had focused on the use of lexical bundles among native and non-native or novices and expert writers, this study outlined the stance expressions in different types of writing. Therefore, his results highlighted the greater use of bundles in research articles. The study also shed light on the urgent need to increase students' awareness of these bundles due to their relatively scarce use, as compared to doctoral dissertations and published papers. However, the study suggested that both groups of students, except doctoral ones, needed instruction on the use of *it-bundles*.

A more recent study by Akeel (2024) concentrated on the role of stance expressions. In this study, 237 writing tasks written in Saudi English as a foreign language by students were explored. These writers used the lexical bundles to serve specific functions; these included supporting a point, representing an item and making suggestions. In addition, Azadnia (2023)

addressed the structural and functional gaps in the use of lexical bundles in two non-native learning context modes of research papers to investigate the possible micro-functions variations regarding text-oriented and research-oriented functions. To add more, Yang (2024) expanded the evaluation of the functions of lexical bundles to more comprehensive research in which shared lexical bundles were linked with rhetorical moves to disclose similar discourse functions. In this regard, a similar exploration of move analysis was conducted by researchers who concluded that all the bundles could not express specific moves. However, previous studies have claimed that lexical bundles could occur in more than one move. The results have specifically shown that *stance* and *referential* bundles are used more frequently than discourse organizing bundles.

Although there have been some studies comparing the competency of L1 and L2 learners considering the employment of these complex phrases, especially regarding proficiency levels and the usage of these phrases, there is still a large gap calling for further studies to focus on the different disciplines in which the bundles are used. Analyzing *it*-clauses will allow us to gain significant insights into high-stakes genres since they are frequent in academic writing. Therefore, in this study, we attempted to have a closer look at those clauses known as *it-bundles* since they encode different stance expressions. The following questions are, therefore, posited for this investigation:

- What it-bundles are mostly utilized in AL and IT research papers?
- How can it-bundles in AL and IT research papers be described functionally?
- Are L1-English writers different from their L1-Persian peers concerning the frequency and function of *it-bundles*?

Method

This research encompassed four corpora—the first one comprised research papers published in the applied linguistics area. The second one represented published work in the field of information technology. Two other corpora were compiled from research papers by L1-English writers in both AL and IT disciplines. The choice of the journals included in each of AL and IT corpora was based on three criteria: previous corpus-based research done on these disciplinary areas, experts' views on the journals best representing each of these fields, and accessibility to the electronic files of the research papers. Therefore, 400 research articles were collected and applied in this study to fulfill its purposes. Academic works accomplished by both L1 English and L1 Persian writers were chosen according to their names to ascertain whether they were native or not. Co-authored articles were also selected based on the first writer's name, as this study assumed that the first author of research papers was more responsible for writing and preparing the final draft. The details of the four corpora have been presented in Tables 1 and 2.

Table 1

Corpora Word Count

Corpora	Number of texts	Number of words
Native corpus of AL	100	898,221
Native corpus of IT	100	1,281,003
Non-native corpus of AL	100	810,416
Non-native corpus of IT	100	613,710
Total	400	3,603,350

Table 2

Information Technology Corpus

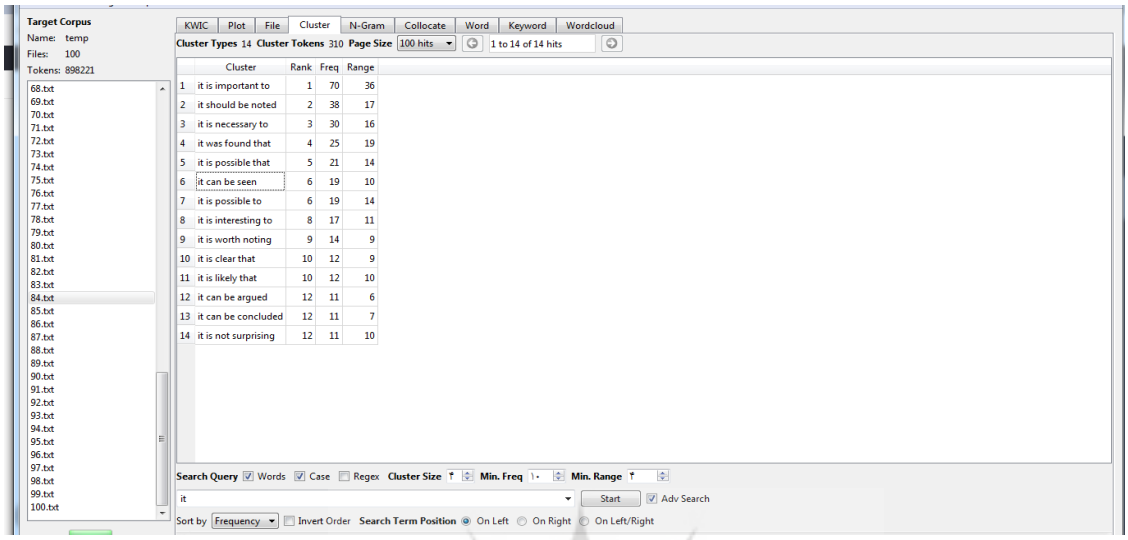
Native corpus of AL/journals	Number of texts	Native corpus of IT/ Journals	Number of texts
English for academic purposes	25	Data and knowledge engineering	25
English for specific purposes	25	Future generation computer system	25
English language Teaching	25	Information science	25
Linguistics and education	25	Information systems	25
Non-native corpus of AL		**Non-native corpus of IT**	
Applied language studies		Artificial intelligence and	
Applied research on	25	data mining	25
English learning	25	Information and	
English language research		communication technology	25
English for academic	25	research	
purposes	25	Information science and	25
		management	25
		Operation Research	

Instrument

To identify it-bundles and check their dispersion and frequency, we employed Antconc software. Antcoc 4.1.0.1(Anthony, 2007). While there are other corpus analysis programs to extract lexical bundles, Antconc has many useful features (e.g., concordances, concordance plots, keyword lists, and collocates). This program was used to identify and concord lexical bundles. After all of the bundles were found with their exact frequencies, they were displayed by the classification of keywords by which bundles were collocated, such as prepositions (only anticipatory *it* in this study), and the minimum optimal frequency was chosen (e.g., thirty in the corpus of one million). Then, the required number of words in clusters was specified (i.e., two, three, four) (see Figure 1). It must be mentioned that in this study, the frequency cut-off of 10 was chosen. A four-word bundle starting with the anticipatory *it* must occur at least ten times in five different texts to be analyzed in this study. Antconc can help researchers find multi-word expressions that contain various lengths and frequencies in every corpus. This convenient program will hand over all word combinations when given any set of words (figure 1).

Figure 1

Antconc.4.1.0.1



Procedure

Three criteria were considered for the identification of *it-bundles*: frequency, dispersion, and length. The frequency was settled on 10 to derive the bundles in the research papers. Dispersion is another criterion showing how many texts they are utilized in. As there are different dispersion rates for different corpora (Hyland, 2008 a, b), in this study, we chose five texts as the criterion for dispersion. So, a word sequence had to be used ten times in five different research papers to be recognized as a lexical bundle. The length of the bundles is the last standard considered here. Only four-word anticipatory *it-bundles* were detected in this research since they are the most frequent ones and also more varied when compared to the other lexical bundles.

Data Analysis

The data analysis for this study was done in three major phases: First, data coding was carried out. All of the research papers were prepared and numbered. Then, the researchers agreed on the identified bundles and their functions after they had negotiated all of the details. Strictly speaking, the researchers reached the full agreement. So, this approach helped them to ensure the flexibility of the coding. Second, all *it*-bundles were elicited from the research articles by adopting Antconc (Anthony,4.1.0). Then, the frequency and number of *it*-bundles were determined thoroughly. Third, the functions of the extracted bundles were explored by adapting the framework of *it-bundles* previously proposed by Hewings and Hewings (2002). So, the phrases were distinguished based on their frequency and function. Eventually, the results were compared to analyze the differences and similarities in research articles among L1-English and L1-Persian writers in terms of frequency and function. However, there are some other functional classifications of

bundles; Hewings and Hewings' taxonomy (2002) was chosen for this study since it contained all the interpersonal functions and subcategories needed for this study. To go for the details, we describe this functional typology in Table 3.

Table 3

Interpersonal function of it-clauses (Hewings and Hewings, 2002, p. 372)

Interpersonal functions	Subcategories	Examples
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,
Attitude markers	2a the writer feels that something is worthy of note 2b the writer's evaluation	It is of interest to note, it is worth pointing out, it is noteworthy, it is important
Emphatics	3a the writer indicates that a conclusion/deduction should be reached; that a preposition is true 3b the writer strongly draws the reader's attention to a point 3c the writer expresses the 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.
Attribution	4a specific attribution (with a reference to a literature) 4bgeneral attribution(no referencing)	It has been proposed (+ reference) It is estimated(no reference)

Results

Considering the results, a frequency of ten was chosen for the study because the corpora contained over one million words. Therefore, 14 *it-bundles* were retrieved in AL and 16 bundles in IT, both in L1-English writing. These bundles accounted for 22% and 26% of the whole corpus. As shown in Tables 4 and 5, regarding the recurrence and dispersion of *it*-bundles, '*it is important to'* occurred more often than the other phrases with a frequency of 70 for AL and 58 for IT; the number of texts (dispersion) was 36 for AL and 31 for IT. In addition, some bundles such as *it should be noted, it is necessary* to, and *it was found that* were employed very often, while some other bundles like *it is likely that, it can be argued* and *it can be concluded* could not be found in the research papers with considerable frequency. Meanwhile, the least frequent bundle was '*it is not surprising,'* with a frequency of 11 and a range of 10 in AL. However, *it is not yet* at the bottom of the table that had the minimum occurrence, with a frequency of 10 times.

Further, it should be mentioned that *it is important to* use the pattern of *it + is + adjective + to* was represented at the top of the table, showing that both writers intended to draw readers' attention by using it similarly. Most frequent bundles contain this structure as the writers were eager to express something clearly. Another important point is that both writers made almost similar use of *it-bundles* with the slightly heavier use of IT writers, while AL writers had more exposure to the English language. Therefore, language proficiency could not lead to the

frequent use of *it-bundles*.

Table 4
It-lexical bundles in L1-English AL

Lexical bundles	Frequency	Range
It is important to	70	36
It should be noted	38	17
It is necessary to	30	16
It was found that	25	19
It is possible that	21	14
It can be seen	19	10
It is possible to	19	14
It is interesting to	17	11
It is worth noting	14	9
It is clear that	12	9
It is likely that	12	10
It can be argued	12	6
It can be concluded	11	7
It is not surprising	11	10

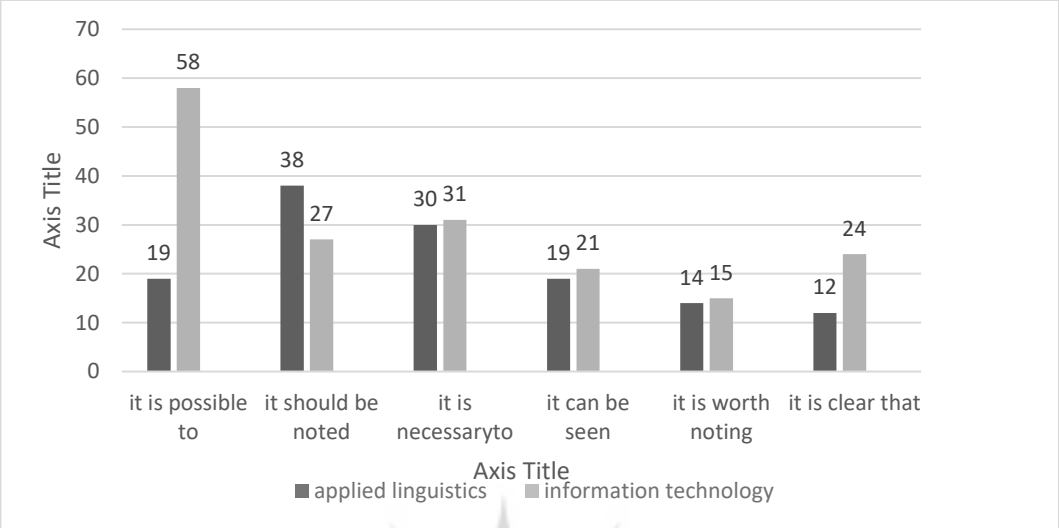
Table 5
It- lexical bundles in L1-English IT

Lexical bundles	Frequency	Range
It is important to	58	31
It is possible to	58	28
It is necessary to	31	21
It should be noted	27	13
It is clear that	24	13
It is easy to	24	14
It is published in	24	24
It can be observed	23	8
It can be seen	21	11
It is difficult to	20	16
It is worth noting	15	8
It is assumed that	12	12
It is expected that	11	8
It comes to the	10	7
It is not necessary	10	5
It is not yet	10	10

The overall occurrence of bundles in both AL and IT was 310 and 378, respectively. Therefore, IT writers not only used two more *it-bundles* but also drew on bundles more frequently than AL writers did. The three top frequent bundles in the former included '*it is important to'* (70), *it should be noted* (38), and '*it is necessary to'* (30). Meanwhile, in the latter, '*it is important to'* (58), *'it is possible to'* (58), '*it is necessary to'* (31) could be identified. Although '*it is important to'* was used by applied linguistics writers with the highest frequency, IT writers drew on the top three bundles more frequently, relative to their counterparts. Therefore, the two corpora only had six bundles in common: '*it is possible to,' 'it should be noted,' 'it is necessary to,' 'it can be seen,' 'it is worth noting',* and *'it is clear that.'* Figure 2 demonstrates these bundles.

Figure 2

Shared it-bundles in L1-English AL and IT corpora



Interestingly, the retrieved bundles in the AL corpus belonging to L1-Persian writers included twenty-two *it-bundles* (see Table 5). These bundles accounted for 36% of the whole corpus, which was more than that in the L1-English corpus of AL and all other corpora. However, only nine bundles were found in the IT corpus, with an overall percentage of 14%, which was the least one. *'It was found that'* was used 72 times as the most frequent one. It was quite similar to the occurrence of '*it is important to'* in the AL L1-English corpus. These frequencies were almost 7 times more than the minimum threshold of 10. Similarly, the range of this bundle was 32, which indicated that this bundle was used in 32 texts by the writers. Unlike the most frequent bundles, *'it is shown in'* was found at the bottom of the table as long as it occurred 10 times in 6 texts.*' It was found that', 'it can be concluded that'* and *'it should be noted'* were at the top of the list.

Not surprisingly, *'it is important* to,' similar to the L1-English corpora, was at the top; meanwhile, '*it was found that'* with a different structure was found to have the highest rank among L1-Persian writers. A reason for these differences might be the fact that L1-Persian writers intended to determine the degree of confidence and provide evidence for their statements throughout the study, whereas IT writers were not interested in preparing any verification, as mentioned earlier. Another point is the absence of one bundle with different structures: '*it is worth noting'* or '*it is worth mentioning'* in the L1-Persian IT corpus, though all other writers used this word combination to evaluate and express feelings.

The results, thus, indicated that five bundles, *it should be noted, it can be seen, it is necessary to, it can be concluded, and it can be said,* were the shared ones in L1-Persian AL and IT corpora. Among these bundles, attitude markers were common in both L1-English and L1-Persian corpora. These bundles included '*it should be noted,' 'it can be seen,'* and *'it is necessary to.'* To conclude, both writers with different L1s showed striking similarities,

regardless of the differences mentioned earlier. Figure 3 demonstrates shared bundles in the L1-Persian corpora. In the next section, we zoom on the similarities and disparities regarding the use of *it-bundles* in the investigated corpora.

Table 6

It lexical bundles in L1-Persian AL corpus

Lexical bundles	Frequency	Range
It was found that	72	32
It can be concluded	40	23
It should be noted	38	24
It is important to	29	11
It is worth mentioning	25	20
It can be seen	21	14
It can be inferred	20	13
It was revealed that	20	15
It can be argued	18	12
It was concluded that	17	13
It is believed that	16	10
It is better to	16	6
It is necessary to	12	6
It could be concluded	11	6
It should be	11	9
mentioned	11	10
It was also found	10	8
It can also be		
It can be said	10	9
It depends on the	10	4
It is noteworthy that	10	7
It is possible to	10	8
It is shown in	10	6

Table 7

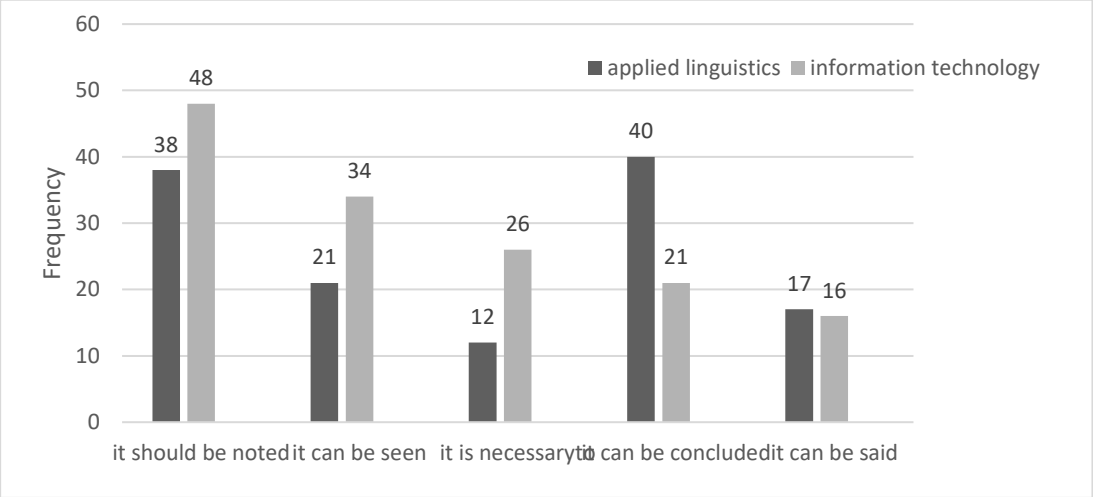
It lexical bundles in L1-Persian IT corpus

Lexical bundles	Frequency	Range
It should be noted	48	28
It can be seen	34	19
It is necessary	26	21
It can be concluded	21	14
It is assumed that	21	15
It is possible to	20	15
It can be said	16	10
It is clear that	16	15
It is obvious that	15	10

Only nine bundles were derived from the L1-Persian corpus of IT (see Table 7). In comparison with L1-English, this corpus contained fewer bundles. '*It should be noted'* was found to occur 48 times in 28 texts as the most frequent bundle. Meanwhile, '*It is obvious that'* was the least frequent bundle. '*It should be noted,' 'it can be seen,'* and '*it is necessary' to* see were at the top of three frequent bundles in this corpus. Among these bundles, *'it is necessary to'* was also found in both L1-English corpora.

Figure 3

Shared it-bundles in L1-Persian corpora of AL and IT



Hedges

As can be seen in Tables 4, 5, 6, and 7, in all four corpora, there were eight bundles of this kind: '*It is possible to,' 'it is possible that,' it is likely that, it can be argued, and it is assumed that.* These bundles show the authors' speculative attitude against the prepositions. AL writers utilized them to express hypothetical statements, albeit IT writers employed them to complete the arguments and made inferences about their findings. Moreover, L1-Persian writers had the tendency to employ this kind more than L1-English writers. So, they were more cautious in diminishing the risk of criticism or opposition: it can be argued, it could be concluded, it should be mentioned, *and it is possible to.* The following four examples show the interpretative and inferential use of these bundles from four different corpora:

(1) In-vivo coding uses the terms and concepts emerging from the participants' words; thus, ***it is possible to*** capture their experiences and perceptions while preserving the meaning of their views and actions in the coding itself.

(2) In fact, more improvement in lexical complexity due to task repetition resulted in less improvement in both measures of accuracy. ***It is possible to*** say that students used more complex language at the expense of accuracy. In other words, they utilized cutting-edge language.

(3) Second, extending our approach to mining rules from both numerical and categorical attributes can be considered. While ***it is possible to*** simply use additional categorical attributes to filter out some ranges and rules discovered by our method.

(4) By extracting meaningful features from the text using Natural Language Processing (NLP), ***it is possible to*** conduct spam detection using various machine learning techniques.

In respect of '*it is possible to,'* it may seem that there was no difference among the writers other than the frequency which was higher in IT, as compared to AL, in both L1-English and L1-Persian. They used this kind of bundle to make their statements less direct and express doubt.

Attitude markers

There were eight anticipatory *it*-bundles serving attitude markers in all corpora: '*it is important to,' 'it is interesting to,' it is not surprising,' 'it depends on the,' 'it is difficult to,' 'it is easy to,' 'it is better to,' 'it comes to the,' 'it is not yet,' 'it is worth noting,' 'it is worth mentioning'* and *'it is noteworthy that.'* Some of these bundles could be encoded as emphatic meanings too (e.g., '*it is important to'*) based on Hewings and Hewing's (2002) typology of *it*-bundles. Even though *'it is important to'* was at the top of the three tables, L1-Persian writers of IT had not used it. In the case of articles written by L1-English authors, it was employed more frequently, which was almost more than twice that of L1-Persian research articles. The following sentences exemplify this bundle:

(5) For practical reasons, however, this was not possible in this study. Nevertheless, in light of the findings of our study, ***it is important to*** note that even a study with as short a duration as 11 weeks, like the present one, surprisingly yielded results that reveal noticeable changes in the use of linguistic resources over time.

(6) The underlying impetus behind this study is that ***it is important to*** examine how these four components of individual differences are linked to learning and achievement and how they are related to each other.

(7) Finally, ***it is important to*** recognize that a process model and textual description may not describe exactly the same steps that comprise a process, whether intentional or not.

As can be seen in the above examples, there are some verbs with which '*it is important to'* is associated. These were like *notes, examinations, and recognitions*. In the case of *note,* it seems that writers tended to bring the reader's attention to an important point, while other verbs are mostly used to express attitudinal meanings. Also, other verbs have been found to collocate with this bundle: *understand, highlight, choose, address, realize, specify, and consider.*

Emphatics

Ten bundles were found with the emphatic pose: '*it is necessary to,' 'it should be noted,' 'it is clear that,' 'it is not necessary,' 'it is believed that,' 'it can be concluded,' 'it could be concluded,' 'it is expected that,' 'it is clear to,' 'it was concluded that'* and *'it is obvious that.'* In all four corpora, *it is necessary* was the most frequent one, while *it is important to* was the second frequent bundle since L1-Persian writers of IT had not used it, thereby showing some consequential variation between L1-English and L1-Persian writers who employ such an attitude marker. It is, therefore, obvious that L1-English writers employed many more *it*-bundles, which served as attitude markers. The infrequent use of *it*-bundles by L1 Persian can be an account of their less confidence in shaping their text. The following examples were extracted from four different corpora of this research:

(8) Higher education focuses on promoting the training of autonomous, critical professionals who adapt to the ever-more demanding labor market. To achieve these objectives, ***it is necessary to*** rethink teaching practices in order to allow the student to be the main actor and modeler for their learning process.

(9) The majority of the participants argue that at the beginning of each university course, ***it is necessary to*** formulate its actual and realistic aims and tasks.
(10) When the journal or conference was published with all its editions in some of the selected databases, it means that it is sufficient to apply an automatic search; otherwise, ***it is necessary to*** apply manual searches for each missing edition.
(11) The existing population gradually increases, and better schemes for scheduling are discovered. ***It is necessary to*** note that increasing population leads to increasing calculations and therefore leads to increasing the time of processing.

Although there was a similar use of one bundle like '*it is necessary to'* in all corpora, another one, '*it should be noted that'* revealed an obvious disparity in terms of frequency. IT writers' heavy use of this bundle by L1-Persian writers showed that they used it in a longer clause followed by *that* to represent certain evidence or explanation. This bundle can be of use to convince the learners to express their viewpoints. Regarding '*it is clear that*' AL authors' non-use of it in L1-Persian and less use of other writers made a significant difference. The following sentences are examples from two corpora of this kind:
(12) ***It should be noted*** that the structure of the initial kernel matrix is important, as it should contain the "optimal" kernel.
(13) ***It is clear that*** this measure is less complex and useful in many applications. The degree of each node is calculated as follows:

Attribution markers

In this study, attribution markers refer to those bundles that refer to something already presented, such as tables or figures, and something that is inferred and interpreted from the tables or conclusions already said. The data analysis demonstrated that there were six bundles of this kind: *it can be seen, it can be observed, it can be inferred, it can be said, it is published in, and it is shown in.* Although the frequency of these bundles was nearly the same in three corpora, including L1-English and Persian research papers in AL and L1-English research papers in IT, the use of these phrases was much heavier in L1-Persian research papers in IT. Additionally, these bundles were found after a preposition like *as* in the two L1-English corpora in both AL and IT fields. It seems, therefore, that writers explicitly tried to engage the readers. The following examples can better illustrate this point:
(14) From the reasons that the learners gave to support the statement "I like/do not like learning from transcripts of authentic spoken language" (Question 6), ***it can be seen*** that most of the learners liked this way of learning because it was perceived to be interesting, useful and practical, as shown in the following examples from Classes A and B:**(15)** As ***it can be seen***, the analysis of the students' interview showed that IOCF was more preferred than DOCF by the students, and this was in line with the teachers' preferences in the interviews. **(16)** Table 7 summarizes the results of this research question. As can be seen, in most of the first-level classes, we found at least one definition that might be classified as generic enough to include

other classes and their corresponding definitions.

(17) ***It can be seen*** that the proposed approach is an effective tool for designing multi-microgrid in the smart network within a defined range of tradeoffs.

Epistemic meaning

Epistemic features reflect those bundles that scaffold the ensuing preposition and show certain or uncertain facts and impersonal ones (e.g., *it is suggested that* it was found that) (Hewing and Hewing, 2002). There were only three bundles of this type based on the analysis: '*it was found that*,' '*it was also found'* and *'it was revealed that*.' All of them were used to outline the results and conclusions, referring to the statements that supported their observations. It is important to mention that these bundles were only used in the AL corpus. While L1-Persian research articles were found to have an overall occurrence of 72, L1-English writers used them with a much lower frequency of 25. Although English writers enjoy high language proficiency and have enough exposure to the English language, they were not willing to negotiate their findings as well. The following example shows one instance:

(18) ***It was found that*** the proportion of verb that-clauses decreased, whereas the proportions of noun that-clauses and adjective that-clauses increased.

However, considering the outcome of L1-English and L1-Persian analysis, L1-English writers make less frequent use of *attitude markers, emphatic, attribution*, and *Epistemic* categories and more use of the hedges category only. Such differences were not only in the overall use of these categories but also in frequencies as well. Figure 4 demonstrates the functional differences in the field of AL and IT. Table 8 also represents an overall description of *it*-bundles in both fields.

Figure 4

Functional distribution of it-bundles in AL

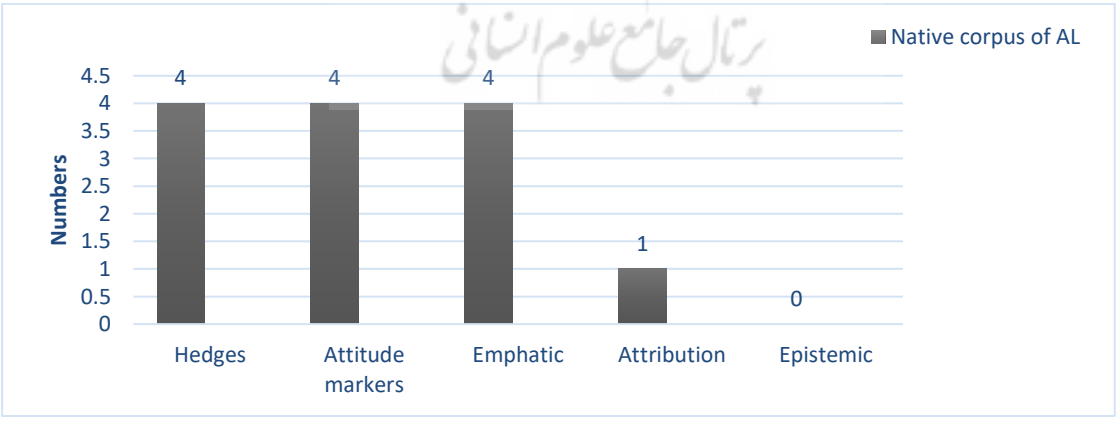
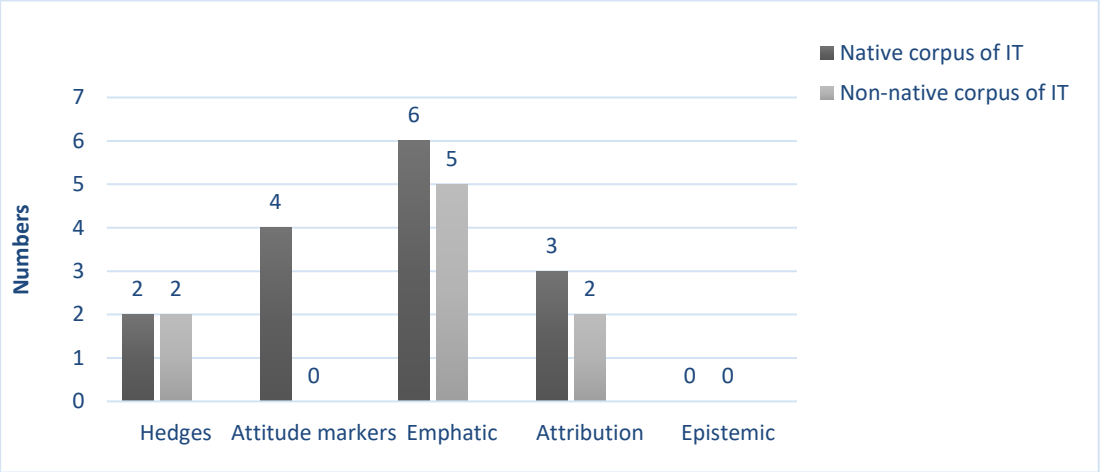


Figure 5

Functional distribution of it-bundles in IT



As displayed in Figure 5, unlike the heavy use of most categories by L1-Persian writers in AL, they make less frequent use of *emphatic and attribution* categories. We can, therefore, note that they didn't understand the *pattern of attitude markers* due to the results. The non-use of the *epistemic* category extends to both writers (L1-English and L1-Persian). However, in the AL corpus, L1-Persian writers tend to use it frequently.

Table 8

Overall functional description of it-bundles in AL and IT research articles

Subcategories	Number	Frequency	Percentage
Hedges**: it is possible to, it is possible that, it is likely that, it can be argued, and it is assumed that**			
1a: likelihood/possibility/certainty/importance/value/necessity etc.	2/1	62/78	50/10
1b: what a writer thinks/assumes to be/will be/was the case	1/1	30/31	10/10
Attitude markers**: it is important to, it is interesting to, it is not surprising, it is difficult to, it is easy to, it is better to, it is worth noting, it is worth mentioning and it is noteworthy that.it depends on the**	3/1	49/15	42.8/25
2a: the writer thinks that something is worthy of note	5/3	153/102	68/75
2b: the writers' evaluation			
Emphatic: **it is necessary to, it should be noted, it should be mentioned, it is clear that, it is not necessary, it is clear to, it is obvious that, it is believed that, it is expected that, it can be concluded, it could be concluded, it was concluded that.**	7/4	166/96	70/40
3a: the writer indicated that a	2/3	28/66	20/30
	1/2	42/67	10/20

Subcategories	Number	Frequency	Percentage
conclusion/deduction should be reached/that a preposition is true.			
3b: the writer strongly draws the readers' attention to a point.	2/2	50/55	50/50
3c: the writer expresses a strong conviction of what is possible/important/necessary etc.	2/1	30/16	50/25
Attribution: **it can be seen.it can be observed, it can be inferred, it can be said, it is shown in**	0	0	0
4a: general attribution (with a reference to the literature)	3/0	128/0	100/0
4b: specific attribution (no reference)	0	0	0
Epistemic: **it was found that, it was also found, it was revealed that**	28/18	738/526	100
5a: certain			
5b: uncertain			
5c: impersonal			
Total			

Discussion

Stance expressions are significant as they reveal the relationships between the reader and the writer. They show the prepositional meaning as well. We could show that different discipline-specific word combinations are the result of studies that evaluate different disciplines (Biber, 2006; Hyland, 2005; Wray, 2002). According to the use of *it-bundles* in the L1-Persian corpus across both AL and IT, some noteworthy differences were found. L1-Persian writers of AL corpus used various *it*-bundles, as compared to their L1-English and L1-Persian peers. Overall, in terms of frequency, there was a slight difference between L1-English writers in AL and IT, with 16 and 14 *it*-bundles, respectively. Moreover, the overall frequency of *it*-bundles was not so different. Only two shared bundles (*it is possible to* and *it is clear that*) were used extremely highly by IT writers. Particularly, the most frequent bundles were not different.

Comparing this study to Adel and Erman (2012), we found significant differences. Adel and Erman concluded that L1-English writers intended to apply various *it*-bundles relative to their L1-Swedish counterparts. Staples et al. (2013), however, claimed that one reason might be that L1-Swedish writers had not attained competence yet. However, some studies have demonstrated that both writers employed approximately the same number of bundles (Shin, 2019).

Regarding the differences, 22 bundles in the AL corpus by L1-Persian and 9 bundles in the IT corpus by L1-Persian were found, respectively. The former contained 36% of the corpora. The less use of bundles by IT writers, thus, revealed the fact that they had not developed the capacity to utilize these particular phrases properly. There was also a difference in bundle tokens, as the top three were between two different disciplines in the L1-Persian corpus. Except for one bundle (*it should be noted*), all other bundles were fully different. More precisely, '*it was found that'* was used by L1-Persian writers in the AL corpus with a frequency of 72. It

seems, therefore, that applied linguistics writing was ahead of information technology since they used bundles more in terms of number and overall frequency.

Regarding functional differences, *attitude markers, emphatic, attribution, and epistemic were used more by L1-Persian writers* in AL, while *Hedges* were found to be more than those of the L1-Persian corpus. L1-Persian writers made a relatively excessive use of the attribution category in the AL corpus, while L1-Persian writers in the IT corpus made less frequent use of this category. To mention the most and least frequent categories in the AL corpus, the order can be like this: *emphatic, attitude markers, attribution, epistemic, and hedges.*

A closer comparison to the studies by Larsson (2017), Lancaster(2016), and Jalali (2017) also shows similar findings. By developing the functional distribution of anticipatory *it*-bundles, the studies could map out the high use of *bundles* by the more proficient students, more specifically, highly-proficient students at the master's and doctoral levels. In contrast, there was also a notable difference regarding the most frequent functional categories. Although epistemic and emphatic were the most highly used functional categories in our research, hedges and emphatic stood out in Larsson's (2017) research. This finding is surprising since emphatic was found to be the most common in both studies due to the importance of its function, which has engaging force, encouraging the readers to pay attention to the points (e.g., *it should be noted*). It must be noted that one possible way to use this functional category is exposure.

Notable differences could also be found when comparing research articles in the IT corpus. Attitude, emphatic, and attribution appeared more in research papers written by L1-English than their L2 counterparts. Although attribution is used less by English writers to refer to something already presented, it was applied relatively rarely in Persian writing, which included only one example of that category. Hedges were used by both writers similarly, and attitude markers did not appear in L1-Persian research articles, thus revealing that such writers did not have much confidence to show their attitude towards the subjects. The epistemic category was the one used only by L1-Persian writers in the field of AL. Interestingly, such bundles appeared to structure the prepositions as certain, uncertain, or impersonal facts (e.g., *it was found that, it was revealed that*). The former was used to express specific observations or results without directly mentioning a single person and the latter was also utilized to express the results using different words and clusters. These findings, thus, suggested that students showed a tendency to present the results of the study. They also learned to narrate the outcomes of their research without expressing themselves in the text. There is also another bundle: *it is suggested* that, which served to express ideas that could be disagreed by others and woud mark the possible differences.

The most and least frequent categories in IT were *emphatic* and *hedges*, which was similar to the AL corpus. The findings were, thus, in line with those obtained by Chen and Liu (2020), in which four disciplines were extensively investigated to shed light on the function of these multi-word expressions. Similar to our study, the category *Hedges* was the least used by the writers, whereas the referential one was the most frequent one in this study. Overall, students

relied on bundles with *emphatic* function rather than *hedges*, thus showing that the instruction regarding these bundles had been apparently effective. However, more instruction and focused attention regarding Hedges could be helpful as they all serve significant and evaluative functions in academic discourse.

Conclusion

This study focused on the sequences of words recognized as lexical bundles that encoded stance expressions. They are fixed phrases that begin with an anticipatory *it*. This study addressed these phrases in the research papers conducted by L1-English and L1-Persian writers in the field of AL and IT, using Hewing and Hewing's functional framework (2002). This research could provide good insights into the use of *it-bundles*. Although they show the heavy use of most categories, including *attitude markers*, *attribution*, and *emphatic* in the field of applied linguistics, compared to their counterparts, the low use of these functional categories was noteworthy. Thus, L1-Persian writers generally used more bundles in the AL corpus, while in the other corpus, L1-English authors drew on more various and frequent bundles. Concerning the differences between both areas of the study, there was not a big difference, but AL writers tended to use bundles slightly more than the IT authors in both L1-English and L1-Persian research papers.

This study, however, involved several limitations that had been uncovered through the research. First, the study was limited to two disciplines. However, the difference between L1 and L2 languages might be more comprehensible if the researchers could evaluate more disciplines. Second, while the corpora used in the present study were larger than those in many previous studies., a corpus containing more articles and different disciplines, as mentioned earlier, would yield more information on it-bundles. Third, the lack of knowledge about the level and age of the article writers may influence the results of the study. Therefore, future research could broaden the evaluation of it-bundles by exploring different disciplines and considering at least the level of the writers. This kind of study would be of great help for teachers in imparting the knowledge of *it-bundles* to their students.

Acknowledgments

We would like to thank the editorial team of TESL Quarterly for granting us the opportunity to submit and publish the current synthesis. We would also like to express our appreciation to the anonymous reviewers for their careful, detailed reading of our manuscript and their many insightful comments and suggestions.

Declaration of conflicting interests

The authors declare no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for this article's research, authorship, and/or publication.

References

Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for specific purposes*, *31*(2), 81-92. https://doi.org/10.1016/j.esp.2011.08.004.

Akeel, E. S. (2024). Exploring the functions of lexical bundles in undergraduate opinion paragraphs for pedagogical use. *Language Learning in Higher Education*, *14*(2), 337-356.https://doi.org/10.1515/cercles-2024-00329(1), 93-124. https://doi.org/10.1515/text.1.1989.9.1.93.

Azadnia, M. (2023). A Corpus-based Study of the Use of Lexical Bundles in EAP Texts by Iranian EFL and ESL Learners. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills)*, *42*(3), 1-27.

Biber, D. (1999). A register perspective on grammar and discourse: variability in the form and use of English complement clauses. *Discourse studies*, *1*(2), 131-150. https://doi.org/10.1177/1461445699001002001.

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

Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for specific purposes*, *26*(3), 263-286. https://doi.org/10.1016/j.esp.2006.08.003.

Brown, P., Levinson, S. C., & Levinson, S. C. (1987). *Politeness: Some universals in language usage* (Vol. 4). Cambridge University Press.

Chafe, W. L., & Nichols, J. (Eds.). (1986). *Evidentiality: The linguistic coding of epistemology* (Vol. 20, pp. 261-312). Norwood, NJ: Ablex Publishing Corporation.

Chen, Y. H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and non-native speakers?. *Applied linguistics*, *29*(1), 72-89. https://doi.org/10.1093/applin/amm022

Cortes, V. (2002). Lexical bundles in freshman composition. *Using corpora to explore linguistic variation*, *9*, 131-145.

Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for specific purposes*, *23*(4), 397-423. https://doi.org/10.1016/j.esp.2003.12.00.

Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and education*, *17*(4), 391-406. https://doi.org/10.1016/j.linged.2007.02.001.

Cortes, V. (2013). The purpose of this study is to: Connecting lexical bundles and moves in research article introductions. *Journal of English for academic purposes*, *12*(1), 33-43.

Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics*, *38*(2), 165-193. https://doi.org/10.1093/applin/amv011.

Farvardin, M. T., Afghari, A., & Koosha, M. (2012). Analysis of four-word lexical bundles in physics research articles. *Advances in Digital Multimedia*, *1*(3), 134-139.

Firth, J. R. (1957). Ethnographic analysis and language with reference to Malinowski's views. *Man and Culture: an evaluation of the work of Bronislaw Malinowski*, 93-118.

Gray, B., & Biber, D. (2012). Current conceptions of stance. *Stance and voice in written academic genres*, 15-33.

Herbel-Eisenmann, B., & Wagner, D. (2010). Appraising lexical bundles in mathematics classroom discourse: Obligation and choice. *Educational Studies in Mathematics*, *75*, 43-63.

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*(4), 367-383. https://doi.org/10.1016/S0889-4906(01)00016-3.

Hunston, S., & Thompson, G. (Eds.). (2000). *Evaluation in text: Authorial stance and the construction of discourse: Authorial stance and the construction of discourse*. Oxford University Press, UK.

Hyland, K. (1996). Writing without conviction? Hedging in science research articles. *Applied linguistics*, *17*(4), 433-454. https://doi.org/10.1093/applin/17.4.433.

Hyland, K. (2001). Bringing in the reader: Addressee features in academic articles. *Written communication*, *18*(4), 549-574. https://doi.org/10.1177/1461445605050365.

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse studies*, *7*(2), 173-192. https://doi.org/10.1177/1461445605050365.

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

Hyland, K., & Guinda, C. S. (Eds.). (2012). *Stance and voice in written academic genres* (pp. 134-150). Basingstoke: Palgrave Macmillan.

Jalali, H. (2017). Reflection of stance through it-bundles in applied linguistics. *Ampersand*, *4*, 30-39. https://doi.org/10.1016/j.amper.2017.06.001

Jhang, S. E., Kim, S., & Qi, Y. (2018). Lexical bundles in ESP writing: Marine accident investigation reports. *Linguistic Research*, *35*.

Jiang, F., Kim, H. D., Na, H. S., Lee, S. Y., Seo, D. W., Choi, J. Y., ... & Chung, M. W. (2015). The influences of CYP2D6 genotypes and drug interactions on the pharmacokinetics of venlafaxine: exploring predictive biomarkers for treatment outcomes. *Psychopharmacology*, *232*, 1899-1909. https://doi.org/10.1007/s00213-014-3825-6.

Jones, M., & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences. *F ormulaic sequences*, 269-300.

Kashiha, H., & Chan, S. H. (2014). Cross-Linguistic and Cross-Disciplinary Investigation of Lexical Bundles in Academic Writing. *Pertanika Journal of Social Sciences & Humanities*, *22*(4).

Labov, W. (1985, June). The several logics of quantification. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 11, pp. 175-195).

Lancaster, Z. (2016). Expressing stance in undergraduate writing: Discipline-specific and general qualities. *Journal of English for Academic Purposes*, *23*, 16-30. https://doi.org/10.1016/j.jeap.2016.05.006.

Larsson, T. (2017). A functional classification of the introductory it pattern: Investigating academic writing by non-native-speaker and native-speaker students. *English for Specific Purposes*, *48*, 57-70. https://doi.org/10.1016/j.esp.2017.06.001.

Liu, C. Y., & Chen, H. J. H. (2020). Analyzing the functions of lexical bundles in undergraduate academic lectures for pedagogical use. *English for Specific Purposes*, *58*, 122-137. . https://doi.org/10.1016/j.esp.2019.12.003.

Pan, F., Reppen, R., & Biber, D. (2016). Comparing patterns of L1 versus L2 English academic professionals: Lexical bundles in Telecommunications research journals. *Journal of English for Academic Purposes*, *21*, 60-71. https://doi.org/10.1016/j.jeap.2015.11.003.

Ruiying, Y., & Allison, D. (2003). Research articles in applied linguistics: Moving from results to conclusions. *English for specific purposes*, *22*(4), 365-385. https://doi.org/10.1016/S0889-4906(02)00026-1.

Shin, Y. K. (2019). Do native writers always have a head start over non-native writers? The use of lexical bundles in college students' essays. *Journal of English for Academic Purposes*, *40*, 1-14. https://doi.org/10.1016/j.jeap.2019.04.004.

Staples, S., Egbert, J., Biber, D., & McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL iBT writing section. *Journal of English for academic purposes*, *12*(3), 214-225. https://doi.org/10.1016/j.jeap.2013.05.002.

Wang, Q. (2015). A conceptual modeling framework for network analytics. *Data & Knowledge Engineering*, *99*, 59-71. https://doi.org/10.1016/j.datak.2015.06.00.

Wei, Y., & Lei, L. (2011). Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELC Journal*, *42*(2), 155-166.

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

Yang, M. (2024). Mapping shared lexical bundles onto rhetorical moves in nursing research articles: A comparative study of paradigmatic variation. *ICAME Journal*, *48*(1), 1-19. https://doi.org/10.2478/icame-2024-0005.