

VIETNAMESE MAJORS OF ENGLISH: THE USE OF LEXICAL BUNDLES IN ARGUMENTATIVE ESSAYS

Phuong Dang and Nga Vo

International University, VNU-HCMC

dhphuong@hcmiu.edu.vn; vtnga@hcmiu.edu.vn

Abstract: Lexical bundles (LBs) defined as fixed expressions in spoken and written discourse have been claimed to mark L2 learners' proficiency level and make their language use natural and native-like. A variety of studies have investigated the use of LBs by both native and non-native speakers in different registers and disciplines. However, there is little research on (1) how Vietnamese learners of English use LBs and (2) how LBs are used in a specific type of argumentative essays. The situation is the motivation for the current paper which investigates the use of lexical bundles in the argumentative essays written by Vietnamese majors of English. Both qualitative and quantitative approaches are employed to examine the frequency, structures and functions of the identified LBs. The findings show that these Vietnamese majors use a wide range of 3-, 4- and 5-word LBs and the clausal structures and stance bundles are dominant in the corpus of argumentative essays.

Keywords: *lexical bundles, academic writing, argumentative essays, English majors*

INTRODUCTION

With the advancement in the corpus-based approach (Nekrasova, 2009), lexical bundles (LBs) which are "collocations or word sequences that statistically co-occur in a register" have recently attracted the attention of many researchers (Cortes, 2004, p. 400). The use of these word combinations is believed to be a marker of language proficiency (Qi & Ding, 2011) and fulfill certain discourse functions. Frequent recurrence is the sole feature of lexical bundles (Ädel & Erman, 2012). In a written corpus, a word string is supposed to have a cut-off frequency of 20 times per million words (Cortes, 2004) and 25 times per million words (Chen & Baker, 2010) to be considered an LB. Another feature that differentiates LBs from other formulaic sequences like idioms is that an LB has a transparent meaning constituted by the meanings of its components.

Most LBs are structurally incomplete. They are often incomplete phrasal chunks (e.g. *as a result of, in terms of*) or incomplete clauses (e.g., *it is possible that, there is an opinion that*). Biber et al. (2004, p. 381) arrange the LBs found in their study into three major structural categories: LBs that contain verb phrase fragments, LBs that contain dependent clause fragments and LBs that contain noun phrases and prepositional phrase fragments. Despite being incomplete structures, LBs do perform several functions as "important building blocks in discourse" (Biber & Barbieri, 2007, p. 270). Most researchers (Haliday, 1971, as cited in Cortes, 2004; Biber et al., 2004) agree on the three primary functions of LBs. First, referential bundles or ideational bundles indicate the reference to a specific entity or to a part of the text which helps present the content (e.g., *in case of*). Second, text or discourse organizers show the relation between a sentence and its surrounding discourse (e.g. *on the other hand*). Finally, stance bundles convey epistemic and attitudinal meaning (e.g., *it is unlikely that*). While the broad framework is adopted by many researchers, its subcategories vary across different corpora.

The literature focuses on three research trends, which provides a strong theoretical background for the current paper. The studies in the first trend explore different corpora to investigate the frequency of LBs as well as their functions and structures (Biber & Conrad, 2004; Biber et al., 2004; Nesi & Basturkmen, 2006; Hyland, 2008; Byrd & Coxhead, 2010). Quantitative

and qualitative approaches are employed to investigate a variety of corpora which vary in different registers and disciplines. The findings are fruitful and become a good start for researching, teaching and learning LBs; lists of most frequent LBs are reported, and their functions and structures are systemized. The studies in the second and third trends focus on how L2 learners use LBs. The researchers compare the LB use by native and non-native language users and find the common notion that L2 learners' LBs are less in frequency and different in function and structure in comparison to those by L1 users (Chen & Baker, 2010; Qi & Ding, 2011; Wei & Lei, 2011; Ädel & Erman, 2012). Besides, the third research trend shows promising results that L2 learners can acquire LBs and use them in their spoken and written discourse to reach native-like proficiency. However, there has been little research on (1) how Vietnamese learners, especially those who major in English use LBs and (2) how LBs are used in argumentative essays.

We are motivated to do an empirical study on the use of LBs in argumentative essays written by Vietnamese L2 majors of English in order to contribute to the field, to see the actual pattern of their LB use and to have appropriate adjustments for teaching LBs to non-native students. We specifically find the answers to the following two research questions.

1. What are 3-, 4- and 5-word LBs used in Vietnamese English majors' argumentative essays?
2. What are the structures and functions of the identified LBs?

Both quantitative and qualitative approaches are employed to evaluate the frequency, structure and function of 3-, 4- and 5-word LBs in the students' academic prose; particularly, a small-scale corpus of 86 argumentative essays or approximately 38,000 words is analyzed. The findings are discussed in comparison to those of previous studies.

METHOD

The Corpus

As summarized in Table 1, the corpus consists of 86 argumentative essays written by Vietnamese majors of English. The average word number in each essay is 433 words and ranges from 243 to 630 words. Therefore, the size of the small-scale corpus is 37,263 words.

Table 1. Description of the Corpus

Variable	N	Mean	SE Mean	StDev	Sum	Minimum	Maximum
Word number	86	433.29	9.65	89.53	37263.00	243.00	630.00

Data analysis

The quantitative methodology involves the measurement of the frequency of LBs in students' argumentative essays. 3-, 4- and 5-word LBs in each essay are carefully counted to measure the frequency of each LB and the total number of all LBs. Then, we classify the identified LBs according to their structures and functions in a qualitative manner.

FINDINGS AND DISCUSSION

Frequency

A total of 535 types and 1006 tokens of LBs are identified in the corpus. Specifically, 4-word LBs are most used regarding LB types, 3-word ones are most used regarding LB tokens, and 5-word ones are least used regarding both criteria. These observations are consistent with those by Hyland (2008) who agrees on the lowest frequency of 5-word LBs and with Conrad and Biber (2004) who finds the higher frequency of 3-word LBs than that of 4-word ones. Specifically, Conrad and Biber (2004) examine LBs in research articles and book extracts for technical

background and student audience and find that 3-word LBs take up 18% and 4-word LBs 2% of the corpus.

Table 2. Frequency of 3-, 4- And 5-Word LBs

	Types	Tokens
3-word LBs	206 (38.5%)	512 (50.8%)
4-word LBs	233 (43.5%)	364 (36.2%)
5-word LBs	96 (18%)	130 (13%)
Total	535 (100%)	1006 (100%)

Furthermore, a closer look at the frequency of LBs provides some interesting findings.

One LB is averagely used almost twice in the whole corpus, and its frequency ranges from 1 to 24 repetitions in the examined essays.

Table 3. Frequency of Each LB

Variable	N	Mean	SE Mean	StDev	Sum	Minimum	Maximum
Frequency	535	1.8785	0.0977	2.2590	1006.0000	1.0000	24.0000

Also, nearly 12 LBs are on average used in each essay, ranging from 2 to 29 LBs.

Table 4. LB Use In Each Essay

Variable	N	Mean	SE Mean	StDev	Sum	Minimum	Maximum
LB number	86	11.686	0.697	6.461	1006.000	2.000	29.000

Table 5 presents the 10 most frequent LBs in the corpus and shows that 3-word LBs are frequently employed by Vietnamese majors of English.

Table 5. The Most Frequently Used LBs

LBs	Percentage
be able to	2.4%
in order to	1.8%
one of the	1.7%
the ability to	1.5%
as well as	1.4%
in the future	1.3%
be likely to	1.3%
as a result	1.2%
on the other hand	1.2%
the value of	1.1%

In comparison with other studies, the findings show that Vietnamese English majors seem to overuse LBs in their argumentative essays. This observation is justified by the significantly high ratio of 233 types and 364 tokens of 4-word LBs over 37,263 words. Chen and Baker (2010) investigate the 150,000-word corpora of Chinese L2 learners, English students and expert writers and find 80 types and 507 tokens in the first corpus, 104 types and 667 tokens in the second, and 108 LBs and 704 tokens in the third. Hyland (2008) examines 4 corpora ranging 600,000 to

1,000,000 words and finds 131, 141, 144 and 213 different LBs. Therefore, the overuse of LBs by Vietnamese students may reflect their attempt to express their linguistic competence and produce native-like language via the employment of these fixed expressions.

Structures

Regarding the LB structures, Table 6 shows that the three most frequently used structures are those of prepositional phrases, a noun phrase + a verb phrase, and a noun phrase + a preposition; their frequency takes up around 20% of the corpus. In fact, their frequency is much higher than that of the remaining structures.

Table 6. Structures of the Identified LBs

Categories	Subcategories	Percentage
Phrasal LBs	PP (e.g. <i>in order to, on the other hand</i>)	20.6%
	NP + prep (e.g. <i>one of the, the value of</i>)	19.4%
	VP (e.g. <i>play an important role, have a chance to</i>)	10.1%
	be + complement (e.g. <i>be able to, be likely to</i>)	8.8%
	NP + complement (e.g. <i>the fact that, the reasons why</i>)	1.1%
Clausal LBs	NP + VP (e.g. <i>that is why, I believe that</i>)	20.2%
	Anticipatory 'it' (e.g. <i>it is advisable that, it goes without saying that</i>)	11.9%
	Existential 'there' (e.g. <i>there are some reasons, there is no need to</i>)	2.4%
Others	<i>first of all, last but not least, what is more</i>	5.5%
Total		100%

The pattern of LB structures is similar to the findings in the previous studies in which PP-based LBs are found to be most frequently used by both L1 and L2 writers (Wei & Lei, 2011; Hyland, 2008; Byrd & Coxhead, 2010; Chen & Baker, 2010). However, the high frequency of the NP + VP structure is typical of the LB use in argumentative essays (Yang, 2017). Students tend to use this clausal structure to express their attitudes and opinions.

Functions

The LBs found in students' essays can be classified into three primary functional categories proposed by Biber et al. (2004) including stance expressions, discourse organizers and referential expressions.

Table 7. Overview of LB Functions in the Corpus

Functions	Tokens	Functions	Tokens
Stance expressions	43.6%	Referential expressions	37.5%
Discourse organizers	18.9%	Total	100%

As seen in Table 7, stance expressions take up the largest proportion of LBs (43.6%). The second commonly used LBs fall into the referential expression category with 37.5%. Discourse organizers comprise the smallest percentage with only 18.9%. The pattern in which stance expressions are more frequently used than referential expressions does not match with the findings of Biber et al. (2004) and Chen and Baker (2010) in terms of LB use in written registers but coincides with Yang (2017) 's finding about LB use in argumentative essays. Obviously, in argumentative essays, students have to frequently employ stance bundles to present opinions and attitudes as well as to persuade readers. This backs up the viewpoint that genres play an important role in the use of LBs, which is also supported by Hyland (2008).

Table 8. Details of LB Functions in the Corpus

Categories	Subcategories	Percentage
Stance expressions	Epistemic (<i>e.g. some people think that</i>)	15.1%
	Attitudinal/modality (<i>e.g., it is important to</i>)	28.5%
Discourse organizers	Topic introduction/focus (<i>e.g., this essay will be about</i>)	2.1%
	Topic elaboration/clarification (<i>e.g., on the other hand</i>)	16.8%
Referential expressions	Identification/focus (<i>e.g. one of the</i>)	9.3%
	Framing attributes (<i>e.g., the importance of</i>)	18.5%
	Quantity specification (<i>e.g. a great deal of</i>)	1.5%
	Place/time/text-deixis (<i>e.g. in the modern world</i>)	8.2%
Total		100%

Table 8 provides an insight into the functional subsets of LBs. Most of the LBs are used by students to express attitudes and modality (28.5%), to convey epistemic information (15.1%), to frame attributes of entities (18.5%) and to elaborate the essays (16.8%).

An interesting pattern that shows a relation between structures and functions can also be found. Most of the stance bundles are clausal fragments with the verbs *think, believe, agree, disagree* (*e.g. many people think that, I disagree with*) and anticipatory *it* structures with adjectives that show assessment like *important, necessary, essential* (*e.g. it is important to*). It also implies that students tend to overuse familiar structures for a specific purpose and only change the vocabulary to make a difference.

CONCLUSIONS

This empirical study investigates the question how Vietnamese English majors use LBs in their argumentative essays. It is found that the frequency of the identified LBs is significantly high, which implies that the students may overuse LBs in an attempt to achieve natural and native-like writing. Specifically, 3-word LBs are used most frequently whereas 5-word ones are used least. Furthermore, in terms of structures, the most preferable LBs are PP-based. The two

observations regarding the LB frequency and structures are confirmed by previous studies. However, the clausal structure of NP + VP is typical of the LB use in argumentative essays. In terms of functions, most of the identified LBs fall into the category of stance bundles, which is also a significant feature of argumentative essays. In fact, the results of the study reflect the similar pattern of the LB use by Vietnamese majors of English except for the observation that they tend to overuse LBs in their argumentative essays. Therefore, explicit instruction is recommended for LBs to be appropriately employed in L2 students' academic writings in order that they may achieve academic standards in their written discourse.

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.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26, 263 - 286.
- Biber, D., & Conrad, S. (2004). The frequency and use of lexical bundles in conversation and academic prose, in Teubert, W. & Mahlberg, M. (eds), *The corpus approach to lexicography, Thematischer Teil von Lexicographica'*. *Internationales Jahrbuch für Lexicographie* 20.
- Biber, D., Susan, C., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25/3, 371 - 405.
- Byrd, P., & Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL*, 5, 31 - 64.
- Chen, Y., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology*, 14 (2), 30 - 49.
- Cortes, V. (2008). A comparative analysis of lexical bundles in academic history writing in English and Spanish. *Corpora*, 3(1), 43-57.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23, 397 - 423.
- Crossley, S., & Salisbury, T. L. (2011). The development of lexical bundle accuracy and production in English second language speakers. *IRAL*, 49, 1 - 26.
- Govindasamy, S., & Kahn, M. N. (2006). Malay ESL college students' spoken discourse: The use of formulaic expressions. *Multilingua*, 25, 59 - 76.
- Hinkel, E. (2009). *Teaching academic ESL writing: Practical techniques in vocabulary and grammar*. Routledge: New York.
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27, 4 - 21.
- Liu, D. (2012). The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes*, 31, 25 - 35.
- Nekrasova, T. M. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Language Learning*, 59(3), 647-686.
- Nesi, H., & Basturkmen, H. (2006). Lexical bundles and discourse signalling in academic lectures. *International Journal of Corpus Linguistics*, 11:3, 283 - 304.
- Qi, Y., & Ding, Y. (2011). Use of formulaic sequences in monologues of Chinese EFL learners. *System*, 39, 164 - 174.
- Stengers, H. et al. (2011). Formulaic sequences and L2 oral proficiency: Does the type of target language influence the association? *IRAL*, 49, 321 - 343.

- Ranjbar, N., Pazhakh, A., & Gorjian, B. (2012). The effect of lexical bundles on Iranian EFL Learners Linguistic Production Fluency. *International Education Studies, Vol. 5, No. 4*, 243 - 251.
- Tremblay, A. et al. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning, 61:2*, 569 - 613.
- Wei, Y., & Lei, L. (2011). Lexical bundles in the Academic Writing of Advanced Chinese EFL learners. *RELC Journal, 42 (2)*, 155 - 166.
- Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: an integrated model. *Language and Communication, 20*, 1 - 28.
- Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics, 21(4)*, 463-489.
- Yang, Y. (2017). Lexical bundles in argumentative and narrative writings by Chinese EFL learners. *International Journal of English Linguistics, 7(3)*, 58.