

## Lexical Bundles in Malaysian ESL Learner Corpora

Engku Haliza Engku Ibrahim and Farah Filzah Khalit

International Islamic University Malaysia

To Link this Article: <http://dx.doi.org/10.6007/IJARBSS/v10-i14/7359>

DOI:10.6007/IJARBSS/v10-i14/7359

*Published Date:* 26 June 2020

### Abstract

It is generally agreed that a mastery of lexical bundles is an essential reflection of mastery of the target language. Thus, this paper aims to investigate the use of lexical bundles of Malaysian ESL learners. Frequency-based lexical bundles were retrieved from a corpus of student essays. The data were collected from two types of corpora; 50 graphic-oriented essays and 50 open-ended essays of upper-intermediate English language learners. Findings of this study revealed that learner have the tendency to use more structural classifications in graphic-oriented essays when compared to open-ended essays.

**Keywords:** Lexical Bundles, Formulaic Sequences, Vocabulary Acquisition.

### Introduction

It is a known fact that vocabulary mastery is essential in language acquisition; so much so that according to Wilkins (1972), without vocabulary, nothing could be conveyed. This is even more so when learners need to acquire the language for academic purposes. However, research has shown that the vocabulary of a language does not just consist of individual words. More often than not, these words are frequently multi-word items. In fact, according to Biber and Conrad (1999) multi-word sequences make up a significant portion of academic prose. Learners' ability to use appropriate multi-words is evidence of mastery of the target language, failure to correctly use these multi-words would indicate non-native like writing proficiency. Despite its importance there are few research that focus on the use of multi-words specifically for academic purposes (Omidian, Beliaeva, Todd & Siyanova-Chanturia, 2017). Thus this research intends to fill the knowledge gap. The implication for the learner is that these words need to be learned as multi-word units rather than individual words. As argued by Biber, Cortes and Conrad (2004), corpus study is the best way to study natural use of language which will then allow teachers to specifically address them in the language classroom. Thus the choice of corpus linguistics for the approach of this research, in an attempt to examine our learners choice and range of lexical bundles.

### What are lexical Bundles

Lexical bundles (Biber & Conrad, 1999; Biber and Barbieri, 2007; Biber, Conrad & Cortes, 2004; Chen & Baker, 2010; Cortes, 2004; Hyland 2008) is referred to in the literature

using many terms such as phrasal expressions (Martinez and Schmitt, 2012), formulaic sequences (Schmitt, 2004; Wray, 2008) or recurrent word combinations (De Cock, 1998). They consist of 3 or more words that do not fall in the category of idiomatic expressions, nor are they complete structural units (Biber, Johansson, Leech, Conrad, and Finegan, 1999; Biber, 2006). The characteristic that sets lexical bundles apart is that they are sequences of words that frequently occur together and this can be evident in both written and spoken corpora. Some examples of lexical bundles include expressions like *at the same time*, *to be able to* and *on the other hand*. Being an important feature of academic discourse genre, they perform a variety of discourse functions (Nausa & Ricardo, 2013).

### **Importance of Lexical Bundles in Language Acquisition**

What researchers have agreed on is that learners need to have the ability to correctly use lexical bundles in order for their writing to appear native-like. Specifically, it can also be said that proper use of lexical bundles contribute, to some extent, to discourse coherence and cohesion. According to Basturkmen (2009), certain features of lexical bundles serve similar functions in showing cohesion as that of Halliday and Hassan's (1976) conjunctive relations. As highlighted by Wray (2000), formulaicity is essential to both the writer or speaker and the listener or reader as it facilitates discourse processing which helps to enhance in the perception of discourse coherence. However, research by Adel and Erman (2012) and Chen and Baker (2010) all indicated that writings of non-native speakers contain fewer as well as less varied lexical bundles. This points to the urgent need for learners to master lexical bundles as it is an indication of a mastery of academic discourse. Even though Wray (2002) is sceptical that the lexical bundles found in learner corpus may not in fact be "stored and retrieved whole from memory" (Wray 2002: 9), it is still useful to study learner production so that teachers could address formulaicity as part of their language teaching strategy as it was emphasised by Jones and Haywood (2004) that when learners underuse formulaic expressions in their writings, their writings are often unacceptable.

### **Studies on Lexical Bundles**

According to Biber and Conrad (1999), corpus study is useful for studying actual language use of learners as they enable researchers to analyse actual distribution in natural texts written by learners. Due to the significance of lexical bundles in the writing of ESL learners, several studies have focussed on instigating learner gains in writing performance after being given instructions of their correct use. Jones and Haywood (2004), for instance, gave a 10-week instruction to university students who were non-native speakers of English. His study deduced that his participants had minor gains in the use of formulaic sequences in the writing production, though it was reported that they had high motivation and predisposition towards the use of formulaic expressions. Similarly, Mahmood Kazemi, Sara Katiraei and Abbass Eslami Rasekh (2014) conducted a study on MA TESL students in Iran. Their findings revealed that the teaching lexical bundles significantly helped their students' writing ability. All these studies point to the significance of highlighting lexical bundles in the language classroom

The main objective of this descriptive research is to investigate the use of lexical bundles of Malaysian ESL learners. Specifically, this research aims to compare if there are differences in learners' use of lexical bundles in graphic-oriented essays and open-ended essays. The research questions of this research are:

1. Which structural classification was used the most?

2. Which structural classification was used the most in graphic-oriented essays?
3. Which structural classification was used the most in open-ended essays?

### Methodology

The corpus of this study was taken from essay scripts of learners who took the English Proficiency Test (EPT) at the International Islamic University Malaysia. The responses were from learners at the high intermediate level of English proficiency. The EPT requires learners to produce two types of written responses; a graphic-oriented essay and an open-ended essay. For the purpose of this study 100 essay scripts (18,805 words) were used; 50 graphic-oriented essays (10,319 words) and 50 open-ended essays (17,753 words). Table 1 summarises the corpus of this present study.

Table 1

*Frequency of bundles and word count from corpus of open-ended and graphic-oriented essays*

Type of Essay	Number of Essays	Word Count	Frequency of 3 To 5 Word Bundles
Graphic Oriented Essays	50	10,319	530
Open-Ended Essays	50	17,753	1,052
<b>TOTAL</b>	100	28,072	1,582

The essays were then digitised and inserted into Wordsmith 2.0 (Scott, 1998) for the identification of lexical bundles.

### Findings and Discussion

#### **RQ1: The highest frequency structural classification used in the present corpus**

Table 2 summarises the structural classification that were found in both types of essays investigated in this study. It can be concluded that the highest structural classification was found in both essay types was the noun phrase. This is consistent with the findings of Biber (2006) and Dontcheva-Navratilova (2012) where it was noted that in academic discourse, almost 70 per cent of the most common bundles consist of noun phrase elements or prepositional phrases which serve the function of discourse organizers and facilitate referential expressions. This is unlike in the highly interactional discourse of conversation, where the majority of lexical bundles are stance expressions containing a verbal element.

Table 2

*Percentages of Structural Classification for open-ended essays and graphic oriented essays: a comparison*

	GRAPHIC-ORIENTED ESSAYS	OPEN-ENDED ESSAYS	TOTAL
<b>Noun Phrase + of and other Noun Phrase</b>	68.1%	24.4%	92.5%
<b>Prepositional Phrase</b>	8.1%	23.2%	31.3%
<b>Passive + Prepositional Phrase Fragment</b>	7.4%	16.8%	24.2%
<b>Anticipatory it + Verb/Adjective</b>	4.5%	6%	5.1%
<b>Be + Noun/Adjectival Phrase</b>	3.2%	0.5%	3.7%
<b>Verb/Adjective Clause Fragment</b>	1.1%	3.1%	4.2%
<b>Pronoun Phrase</b>	0.9%	4.6%	5.5%
<b>Adverbial Phrase</b>	0.6%	1.5%	2.1%
<b>1<sup>st</sup> Person Pronoun + Dependent Clause Fragments</b>	1%	5.5%	6.5%
<b>2<sup>nd</sup> Person Pronoun + Dependent Clause Fragments</b>	-	0.2%	0.2%
<b>3<sup>rd</sup> Person Pronoun + Dependent Clause Fragments</b>	-	1.5%	1.5%
<b>Yes/No Question Fragments</b>	-	0.4%	0.4%
<b>Wh-Question Fragments</b>	-	0.8%	0.8%
<b>Others</b>	0.6%	0.6%	1.2%
<b>TOTAL</b>	100%	100%	

### **RQ 2: The most used structural classification in graphic-oriented essays**

Table 3 shows the types of structural classification found in the graphic-oriented essays of this study. The findings revealed that in the graphic-oriented essays, the noun phrase (68.1%) was by far the highest type of structural classification found in our student writings, followed by prepositional phrases (8.1%). This could be contributed to the nature of graphic -oriented essays that require a higher degree of fixed expressions such as in *comparison to, the end of the and the average number*.

Table 3

*Percentages of Structural Classification in open-ended essays*

	<b>Graphic-oriented essays</b>
<b>Noun Phrase + of and Other Noun Phrases</b>	68.1%
<b>Prepositional Phrase</b>	8.1%
<b>Passive + Prepositional Phrase Fragment</b>	7.4%
<b>Anticipatory It + Verb/Adjective</b>	4.5%
<b>Be + Noun/Adjectival Phrase</b>	3.2%
<b>Verb/Adjective Clause Fragment</b>	1.1%
<b>Pronoun Phrase</b>	0.9%
<b>Adverbial Phrase</b>	0.6%
<b>1<sup>ST</sup> Person Pronoun + Dependent Clause Fragments</b>	1%
<b>2<sup>ND</sup> Person Pronoun + Dependent Clause Fragments</b>	-
<b>3<sup>RD</sup> Person Pronoun + Dependent Clause Fragments</b>	-
<b>Yes/No Question Fragments</b>	-
<b>Wh-Question Fragments</b>	-
<b>Others</b>	0.6%
<b>TOTAL</b>	100%

**RQ 3: The most used structural classification in open-ended essays**

Table 4 shows the types of structural classifications in the open-ended essays. The findings revealed that in the open-ended essays, the noun phrase (24.4%) was also the highest type of structural classification used followed by prepositional phrases (23.2%). In comparison to graphic-oriented essays, learners used less noun phrases in open-ended essays. Learners also used 2<sup>nd</sup> person pronoun + dependent clause fragments (0.2%), 3<sup>rd</sup> person pronoun + dependent clause fragments (1.5%), Yes/No question fragments (0.4%) as well as wh-question fragments (0.8%) in the open-ended essays. These 4 classification structures were not used at all in the graphic-oriented essays. What can be seen is that learners used a wider repertoire of lexical bundles in the open-ended essays compared to the graphic-oriented essays.

Table 4

*Percentages of Structural Classification in graphic-oriented essays*

	Open ended essays
<b>Noun Phrase + of and Other Noun Phrases</b>	24.4%
<b>Prepositional Phrase</b>	23.2%
<b>Passive + Prepositional Phrase Fragment</b>	16.8%
<b>Anticipatory It + Verb/Adjective</b>	6%
<b>1<sup>ST</sup> Person Pronoun + Dependent Clause Fragments</b>	5.5%
<b>Pronoun Phrase</b>	4.6%
<b>Verb/Adjective Clause Fragment</b>	3.1%
<b>Adverbial Phrase</b>	1.5%
<b>3<sup>RD</sup> Person Pronoun + Dependent Clause Fragments</b>	1.5%
<b>Wh-Question Fragments</b>	0.8%
<b>Others</b>	0.6%
<b>Be + Noun/Adjectival Phrase</b>	0.5%
<b>Yes/No Question Fragments</b>	0.4%
<b>2<sup>ND</sup> Person Pronoun + Dependent Clause Fragments</b>	0.2%
<b>TOTAL</b>	100%

### Conclusion

The objective of this study was to find out if there different patterns of structural classification found in the two type of essays; graphic-oriented-essays and open-ended essays. Our findings reveled that were differences in the number of structural classifications used, though the noun phrase as well as prepositional phrase were the two most used structural classifications in both essay types. Even though this has been found to be the case across most research findings, previous studies have also shown that non-native learners tend to have a smaller repertoire of lexical bundles (Adel & Erman, 2012 and Chen & Baker, 2010). Our findings also revealed that there was a difference in the repertoire of structural classification between the two types of essays. Our learners displayed a wider repertoire of structural classification types in the open-ended essays. This study further affirms that teachers should explicitly address these fixed forms in order to create noticing and raise learner awareness of their occurrences. This study also emphasizes the significance of explicitly addressing lexical bundles in the language curriculum to ensure that learners acquire communicative competence that could be successfully communicated in their academic writing.

### 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*, 81–92.
- Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. *Language and Computers, 26*, 181-190.
- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. Amsterdam, the Netherlands: John Benjamins.

- Biber, D. (2010). Corpus-based and corpus-driven analyses of language variation and use. In B. Heine & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (pp. 159–191). Oxford: Oxford University Press.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26, 263–286.
- Biber, D., Conrad, S., & Cortes, V. (2004). 'If you look at...': Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25, 371–405.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.
- Cortes, V. (2006). Teaching lexical bundles in the disciplines: an example from a writing intensive history class. *Linguistics and Education*, 17, 391–406.
- Chen and Baker. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology* 14(2): 30–49.
- DeCock, S. (2000). Repetitive phrasal chunkiness and advanced EFL speech and writing. In C. Mair, & M. Hundt (Eds.), *Corpus linguistics and linguistic theory* (pp. 51–68). Amsterdam: Rodopi.
- Dontcheva-Navratilova, O. (2012). Lexical Bundles in Academic Texts by Non-Native Speakers. *Brno Studies in English*, 38(2).
- Hyland, K. (2008). As can be seen: lexical bundles and disciplinary variation. *English for Specific Purposes*, 27, 4–21.
- Lewis, M. (2009). *The idiom principle in L2 English: Assessing elusive formulaic sequences as indicators of idiomaticity, fluency, and proficiency*. Saarbrücken, Germany: VDM Verlag.
- Kazemi, M., Katiraei, S., & Rasekh, A. E. (2014). The impact of teaching lexical bundles on improving Iranian EFL students' writing skill. *Procedia-Social and Behavioral Sciences*, 98, 864-869.
- Nausa T., & Ricardo A. (2013). Lexical Bundles in Academic Presentations Given by PhD level EAP Students: A Literature Review. *Education & Learning Research Journal* Vol. 7, p259-270 from EBSCO.
- Omidian, T., Beliaeva, N., Todd, L., & Siyanova-Chanturia, A. (2017). The use of academic words and formulae in L1 and L2 secondary school writing. *New Zealand Studies in Applied Linguistics*, 23(2), 39.
- Ruan, Z. (2016). Lexical Bundles in Chinese Undergraduate Academic Writing at an English Medium University. *RELC Journal*, 1-14.
- Vo, S. (2016). Use of Lexical Features in Non-native Academic Writing. *Corpus Linguistics Research*, 2, 53-53.
- Wilkins, D. A. (1972). *Linguistics in language teaching*. London: Arnold.
- Wray, A. (2000). *Formulaic sequences in second language teaching: Principle and practice*. *Applied Linguistics* 21 (4), 487–489.
- Wray, A. (2006). Formulaic language. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (pp. 590–597). Oxford: Elsevier.