

<https://doi.org/10.17323/jle.2023.16305>

# Lexical Bundles in Indonesian EFL Textbooks: A Corpus Analysis

Priyatno Ardi<sup>1</sup>, Yacinta Dinda Oktafiani<sup>1</sup>, Nugraheni Widianingtyas<sup>2</sup>,  
Olga V. Dekhnich<sup>3</sup>, Utami Widiati<sup>4</sup>

<sup>1</sup> Universitas Sanata Dharma, Yogyakarta, Indonesia

<sup>2</sup> Universitas Bunda Mulia, Jakarta, Indonesia

<sup>3</sup> Belgorod State National Research University, Belgorod, Russia

<sup>4</sup> Universitas Negeri Malang, Malang, Indonesia

## ABSTRACT

**Background:** Lexical bundles in textbooks are of paramount importance in foreign language learning. They provide a framework for new vocabulary acquisition and help to build fluency. Despite many studies on lexical bundles, investigations into their usage in EFL textbooks in the Indonesian context are still rare.

**Purpose:** This corpus-based study examines the patterns and structural classifications of lexical bundles in EFL course textbooks for Indonesian senior high school students. As such, it could yield ready-made chunks of English which could be incorporated into students' spoken and written communication.

**Method:** The AntConc software version 3.5.9 was used to extract lexical bundles from five Indonesian Senior High School English textbooks. These books were endorsed by the government to be used across the country. The corpus revealed that the textbooks had 54,009 lexical bundles. In addition, the bundles were categorized into patterns and structural classifications based on Biber et al. (1999).

**Results:** The results showed the patterns included three-word lexical bundles with 32,527 occurrences, four-word with 11,620, five-word with 6,073, and six-word with 3,789. Furthermore, eleven structural classifications of lexical bundles were found in the textbooks: "noun phrase + of phrase fragment" with 173 occurrences; "noun phrase + other post modifier fragment" with 44; "other noun phrases fragment" with 157; "prepositional phrase + of" with 13; "other prepositional phrases" with 243; "anticipatory it + verb phrase/adjective phrase" with 13; "passive verb + prepositional phrase" with 19; "copula be + noun phrase/ adjective phrase" with 30; "(verb phrase +) that- clause" with 59; and "(verb/adjective +) to- clause" with 239.

**Conclusion:** Three-word lexical bundles were the most frequent in the senior high English textbooks. High frequency implies repetition of the bundles. Also, the other prepositional phrase fragment was the most frequent structural classification. Short bundles may have been intended to help students to retain vocabulary and recall the bundles in the usage. This study, therefore, provides valuable insights into the most common groups of words used in the Indonesian EFL textbooks. Pedagogically speaking, repeated bundles in English textbooks can familiarize EFL students with the patterns, and they can use them in spoken and written communication.

## KEYWORDS

lexical bundles, corpus linguistics, senior high school English textbooks

**Citation:** Ardi P., Oktafiani Y. D., Widianingtyas N., Dekhnich O. V., & Widiati U. (2023). Lexical Bundles in Indonesian EFL Textbooks: A Corpus Analysis. *Journal of Language and Education*, 9(2), 25-39. <https://doi.org/10.17323/jle.2023.16305>

## Correspondence:

Priyatno Ardi,  
priyatnoardi@usd.ac.id

**Received:** November 14, 2022

**Accepted:** June 15, 2023

**Published:** June 30, 2023



## INTRODUCTION

Lexical bundles have gained much attention from studies in corpus linguistics (e.g., Apple, 2022; Biber & Barbieri, 2007; Biber et al., 1999; Kim & Kessler, 2022; Liu & Chen, 2020; Panthong & Poonpon, 2020; Sanosi, 2022). This idea pertains to a cluster of words that tend to

occur together with a higher frequency in a particular type of text (Biber et al., 1999). Scott (1997, p. 234) labeled these bundles of words "clusters", since they can be seen as a cohesive unit, similar to treating a set of keywords as a single word. Therefore, the process of grouping the words is called clustering. Stubbs (2015) used the term "N-grams", where

N refers to the sequence of words. It is worth noting that lexical bundles differ from idioms, which tend to be fixed in structure, complete in a phrase, and semantically complete. Unlike idioms, lexical bundles are not formulaic and fixed in their function at pragmatic levels (Biber et al., 2004; Biber et al., 1999; Kurniawan & Haerunisa, 2023; Panthong & Poonpon, 2020), and they frequently appear as strings of three or more words. Since lexical bundles play a significant role in communication as building blocks for discourse, they are considered essential for English language education. They are especially valuable in evaluating students' proficiency in English and supporting their language acquisition (Akbulut, 2020; Hussain et al., 2021; Northbrook & Conklin, 2019; Siricharoen & Wijitsopon, 2020).

The use of bundles of words in EFL textbooks is essential for improving students' vocabulary, since these textbooks place utmost importance on the target language. Firstly, individual words change their meaning when embedded in different contextual environments. Students need to understand the meaning of structured phrases, in order to comprehend the overall meaning of the sentence (Nesi & Basturkmen, 2009). By grasping the meanings of words, students can express themselves beyond their current level of proficiency and expand their vocabulary knowledge. Secondly, although lexical bundles exhibit intricate and incomplete structural patterns, they serve as vital building blocks of typical discourse (Conrad & Biber, 2005; Gil & Caro, 2019). While linguists may not recognize them intuitively, they fulfill an essential function. Thirdly, lexical bundles constitute the basic construction of discourse in academic registers, particularly in spoken and written forms (Conrad & Biber, 2005). These bundles are crucial for English language students, since they comprise simple expressions and are easy to learn within the normal language acquisition process (Biber et al., 1999; Northbrook et al., 2022). Lastly, since lexical bundles are highly prevalent in natural language, mastering them aids students in enhancing their fluency and achieving more natural usage (Hyland & Jiang, 2018).

Studies have shown that lexical bundles are present and significant in English materials, as highlighted by various researchers (e.g., Allan, 2016, 2017; Ganji & Nasrabad, 2021; Hussain et al., 2021; Jablonkai, 2009; Lee, 2020; Northbrook & Conklin, 2018, 2019; Northbrook et al., 2022). Lexical bundles have been proven to be helpful in improving reading comprehension and assisting students in achieving higher levels of language proficiency. This is achieved through the use of graded readers (Allan, 2016) and multi-word combinations (Lee, 2020; Hussain et al., 2021). The use of formulaic and lexically-dense language has also been found to be crucial in enhancing students' proficiency levels (Ganji & Nasrabad, 2021; Jablonkai, 2009). However, Allan (2017) found that the majority of the bundles available were for language instruction, whereas only a small portion was for the target language input. Additionally, the target input bundles were mostly formal, which was not consistent with

English as a Lingua Franca (ELF) daily conversations. Similarly, Northbrook and Conklin (2018) discovered that lexical bundles in Japanese EFL textbooks were unlikely to present and model authentic conversational language, prioritizing high-frequency language instead. Authentic lexical bundles were crucial in sharpening EFL students' communicative competence, as they showed sensitivity towards recurring lexical bundles in their input, primarily from their textbooks (Northbrook & Conklin, 2019; Northbrook et al., 2022).

Since lexical bundles may create simple expressions that accelerate students' vocabulary mastery and target language proficiency, the presence of lexical bundles in English textbooks helps the students to acquire the target language. Therefore, the authors of the books have tried to provide such language inputs in the textbooks. While Indonesian scholars have explored the use of word bundles in writing (e.g., Fajri et al., 2020; Kurniawan & Haerunisa, 2023; Oktavianti & Prayogi, 2022; Putri & Suhardijanto, 2022), as far as we know, no research has yet focused on analyzing lexical bundles in Indonesian ELT textbooks. This study seeks to examine the use of lexical bundles in English textbooks written by Indonesian authors for high school students in Indonesia who are learning English as a foreign language. These textbooks are often used as a resource for the students to improve their language skills. Even though Meunier and Gouverneur (2009) stated that some textbooks are sometimes not worth teaching due to language use, ignorance of practical usage, and infrequent words and phrases, some authors have endeavored to make the target language as natural as possible. This suggests that textbooks should make authentic and natural contexts providing typical English features, such as collocations and vague language (Alasmay, 2022; Alfiandita & Ardi, 2020; Hussain et al., 2021; Meunier, 2012). In this regard, the natural use of the target language is reflected in the patterns and structural classifications of lexical bundles.

By considering the importance of lexical bundles, this paper aims to analyze the frequency of patterns and structural classification which mostly appear in English textbooks for Indonesian senior high school students. In addition, this study is expected to help evaluate the language used in textbooks. The following questions are addressed:

1. Which lexical bundle patterns are most commonly utilized in senior high school EFL textbooks?
2. What is the most common structural classification of lexical bundles in senior high school EFL textbooks?

## THEORETICAL UNDERPINNINGS

Lexical bundles are various in patterns and structural classifications. Therefore, understanding its patterns and structural classifications is vital to frame this study. The following

subsections present patterns and structural classifications of lexical bundles, using the concept and examples from Biber et al. (1999).

## Patterns of Lexical Bundles

Identifying the patterns of lexical bundles gives insight into how words and phrases are combined together in texts. Biber et al. (1999) proposed patterns of word combinations and set different minimum cut-offs for each pattern of lexical bundles. While the minimum cut-off is somewhat arbitrary and dependent on the study's scope, it helps determine what qualifies as a lexical bundle. For instance, a two-word combination such as "I don't" (I do not) can be considered a lexical bundle when it comprises three lexical units. Furthermore, when certain three- or four-word combinations appear at least ten times per million words across a minimum of five texts, they are considered as lexical bundles (cf. Salazar, 2014).

Moreover, studies have established that the shorter the bundles, the more frequent they are. Biber et al. (1999) illustrated this by demonstrating that three-word lexical bundles occur more than 80,000 times per million words in spoken discourse and over 60,000 in academic writing. Four-word lexical bundles occur over 8,500 times per million words in conversational genres and over 5,000 in academic works. However, five- and six-word clusters are infrequent in both conversation and academic writing.

Given that five- and six-word bundles are less frequent than their shorter counterparts, Biber et al. (1999) established a lower cut-off for them. These longer bundles tend to appear across a minimum of five texts to be categorized as lexical bundles. Consequently, the most frequently observed patterns of lexical bundles, as indicated by corpus software, are the three-word and four-word combinations, whereas five- and six-word bundles are relatively less common.

While it is possible to extract lexical bundles using an  $n+1$  algorithm (where  $n$  represents the word number in a bundle) through a corpus program (Kopaczyk, 2013), there is a potential for overlap between these bundles. Such an overlap can occur in two ways, namely syntagmatic and paradigmatic. In the case of syntagmatic overlap, the bundles start and end with a different word. The rest of the words, except the beginning and end, are the same as the previous bundle. Although there is an overlap between two bundles, their counts will vary in the corpus because the items in the beginnings or endings of the bundles are different. In brief, the same-length bundles represent different structural patterns even though there is a syntagmatic overlap between them (Kopaczyk, 2013). On the other hand, paradigmatic overlap occurs like a 'matryoshka doll' effect, where the shorter bundles are included in the longer ones. The next

bundle starts with the same word from the previous ones and ends with a different word. For instance, the first bundle is "a matter of," and the next will be "a matter of fact."

## Structural Classification of Lexical Bundles

In terms of categorizing lexical bundles based on structure, most of them tend to be incomplete. Nevertheless, according to Biber et al. (1999), complete structures of lexical bundles account for 15% in conversational language and 5% in academic texts. Lexical bundles are essentially fragments embedded within phrases or clauses (Salazar, 2014). Despite their incomplete nature, there exists a strong grammatical correlation among lexical bundles, leading to the development of a taxonomy that classifies them into various primary structural categories (Biber et al., 1999; Gil & Caro, 2019).

Biber et al. (1999) categorized the structural framework into twelve classifications. Later, Chen and Baker (2010) reorganized original classifications of lexical bundles of Biber et al. (1999) into three main phrasal categories. These categories are based on the types of phrases they belong to, namely noun phrases, prepositional phrases, and verb phrases. Meanwhile, the clausal categories of lexical bundles consist of adverbial clauses, and bundles formed by that-clauses and to-clauses. Adverbial clauses modify the verb in the independent clause, while that-clauses and to-clauses function as nominal elements within the sentence. The following subsections offer a more detailed overview of the twelve classifications by Biber et al. (1999).

### ***Noun Phrase with of- Phrase Fragment***

These lexical bundles possess a wide range of meanings and are used for the tangible attributes, including place, size, and quantity (e.g., "*the size of the sample makes ...*"). Moreover, they are used for existence or presence, and abstract qualities (e.g., "*... due to the presence of the ...*" and "*the use of a diameter...*"). They also explain processes or events which extend throughout a duration (e.g., "*... contributed to the development of an additional depletion ...*").

### ***Noun Phrase with Other Post-Modifier Fragments***

Recurrent expressions rarely follow a noun phrase with a post-modifier other than of-phrase. There are two types of these bundles, namely noun phrase with a post-nominal clause fragment and noun phrase with a prepositional phrase fragment. These bundles describe the process of how it occurs using "*the way in which*" and "*to extent to which*" (e.g., "*... concerning the way in which electrons were...*"). It is used to identify the relationship among entities. This also consists of the noun + complement clause combination (e.g., "*the fact that the ... is difficult to ...*").

**Prepositional Phrase with Embedded of- Phrase Fragment**

This category functions as the post-modifier of the noun. The majority of this category's lexical bundles express abstract and logical relationships (e.g., "as a result of these ...").

**Other Prepositional Phrase (fragment)**

Most lexical bundles which begin with a preposition "in" are frequent in this category (e.g., "in the next chapter ..."). Another expression is "at the same time," which is used to differentiate between two statements or occurrences considered to be compatible (e.g., "In this way, ..., yet at the same time provided ...").

**Anticipatory it + Verb Phrase/ Adjective Phrase**

The predominance of lexical bundles which initiate an extraposed structure contain predicative adjectives governing a complement phrase. This is typically a to-clause (e.g., "... it is possible to recognize ..."). The majority of these extraposed bundles with verb predicates are of the passive that-clause (e.g., "it should be noted that ...").

**Passive Verb + Prepositional Phrase Fragment**

In this category, the lexical bundles are composed of a passive voice verb and a prepositional phrase. It describes logical relationships rather than the agent in a by-phrase (e.g., "... is shown in figure ...") and also identifies the basis of findings (e.g., "... is based on the fact that ...").

**Copula be + Noun Phrase/ Adjective Phrase**

Each lexical unit in this category starts with the copula be (or may be). There are two major subgroups dependent on the subject predicative, namely noun and adjective phrases. These bundles serve as the subject of the copula be (e.g., "the project is one of the most..."). In addition, lexical bundles with adjective phrases are used to show causative/comparative relations (e.g., "it may be due to ...").

**(Verb Phrase +) that-Clause Fragment**

The majority of lexical bundles containing a verb followed by a that-clause in the main sentence are extraposition structures. These bundles frequently appear in unmarked declarative structures as a complement clause (e.g., "... have found that there is a positive correlation between ...").

**(Verb/ Adjective +) to-Clause Fragment**

Lexical bundles with predicative adjectives which control a to-clause indicate a possibility or ability (e.g., "... work is likely to be..."). Meanwhile, those with verb predicates that con-

trol a to-clause identify previous findings (e.g., "... has been shown to operate in ...").

**Adverbial Clause Fragment**

This structural classification is used to relate references to other discourse segments. An example is "each individual is, as we have seen, ..."

**Pronoun/ Noun Phrase + be (+...)**

It is rare to find lexical bundles which begin with clauses. They have copula be as the verb, and the subject can be the pronouns "this" and "there". The bundles with "this" as a subject are used to connect the information followed by the preceding discourse. An example is "this is not to say that..." Meanwhile, those which begin with existential "there" are used for informational packaging purposes (e.g., "... showed that there was no significant difference between ...").

**Other Expressions**

This category refers to the bundles which do not fit with the other classifications. The examples are "as well as the", "may or not may not", and "the presence or absence."

## MATERIALS AND METHODS

### Materials

This study examined five English textbooks written by Indonesian authors and published by the government. These textbooks were chosen as the data source because they were primarily used in schools and distributed for free. Even though the authors were non-native English speakers, they were able to provide the target language inputs by considering the proficiency levels of Indonesian EFL students. Since the books were developed based on the existing national curriculum, the Indonesian government also supervised the writing process to ensure that the learning outcomes set in the curriculum were met. The corpus consists of 300,859 tokens, as shown in Table 1.

### Method

The research team employed a corpus-based study. According to McEnery and Wilson (2019), corpus linguistics studies language based on actual examples of usage. Furthermore, the AntConc corpus software version 3.5.9<sup>1</sup> was used to extract the data. This is a multiplatform practical tool that can be used in corpus studies and language classrooms for investigating data-driven language learning (Yusuf, 2020). The software was also used to convert the English textbooks, including instructions and teaching materials, from pdf to txt.

<sup>1</sup> Anthony, L. (2020). AntConc (Version 3.5.9) [Computer Software]. Waseda University. <https://www.laurenceanthony.net/software>

**Table 1**  
*List of the Books*

No	Title	Author(s)	Publisher(s)	Tokens
1.	Interlanguage: English for Senior High School Students X	Joko Priyana, Arnys Rahayu Irijayanti, and Virga Renitasari	Pusat Perbukuan Departemen Pendidikan Nasional	38,871
2.	Interlanguage: English for Senior High School Students XI Science and Social Study Program	Joko Priyana, Riandi, and Anita Setyo Mumpuni	Pusat Perbukuan Departemen Pendidikan Nasional	54,404
3.	Interlanguage: English for Senior High School Students XI Language Study Program	Joko Priyana, Zayin Adib Muhammad, and Eka Dennis Machfutra	Pusat Perbukuan Departemen Pendidikan Nasional	79,171
4.	Interlanguage: English for Senior High School Students XII Science and Social Study Program	Joko Priyana, Triyani Retno Putri Saridewi, and Yuliyanti Rahayu	Pusat Perbukuan Departemen Pendidikan Nasional	52,381
5.	Interlanguage: English for Senior High School Students XII Language Study Program	Joko Priyana, Zayin Adib Muhammad, and Eka Dennis Machfutra	Pusat Perbukuan Departemen Pendidikan Nasional	76,032
<b>Total Tokens in Corpus</b>				<b>300,859</b>

The instructions in the textbooks were deliberately included in the corpus, since they provided authenticity for communicative practice (Northbrook & Conklin, 2018). We treated both spoken and written language in the books equally, since they could help students improve their communication skills. However, we chose to exclude the table of contents, acknowledgments, references, authors' biographies, and appendices in the front and back of the books.

## Procedures and Data Analysis

The AntConc software has seven main features, namely concordance, concordance plot, file view, clusters/N-grams, collocates, word list, and keyword list. The features used to examine lexical bundles include clusters/N-grams, concordance, and file view. The "N-gram" button was used to identify lexical bundles in the texts, and the "concordance" feature helps identify the concordance line of the bundles. Meanwhile, the "file view" feature functions to see the detailed text of the bundles. In order to provide a broader range of lexical bundles for analysis, we incorporated all possible combinations of words.

The textbooks were converted and stored in the software. Then the research team set the minimum cut-off for all lexical bundle patterns ten times per million words, in order to generate their frequency of patterns (cf. Biber, 2006; Biber et al., 1999). Nevertheless, in the case of 5- and 6- word bundles, we modified the cutoff frequency requirement to include only instances which appeared at least five times within the corpus. The dispersion values utilized in this procedure ranged from zero to one. When a lexical bundle had a dispersion value of zero, it only appeared in one part of the corpus. In contrast, higher dispersion values indicated that the lexical bundle appeared in multiple sections of the cor-

pus (Burch et al., 2017). Subsequently, a list of lexical bundles was generated along with frequencies and dispersion by clicking the start button of the software. There was no treatment of bundle overlapping because either the beginnings or the endings of the item differed. Accordingly, their counts in the corpus were different, and the overlapping bundles were included in the analysis and then categorized into actual patterns. In order to analyze the structural classification of lexical bundles, the research team divided the overlapping bundles into two broad categories, namely phrasal and clausal (Chen & Baker, 2010). We manually double-checked the overlapping bundles to ensure that they did not affect the analysis. Lastly, the frequency for each pattern and structural classification was calculated.

## RESULTS

The result section is divided into two subsections. The first section presents the answer to the first research questions, i.e., patterns of lexical bundles in Indonesian EFL textbooks. Meanwhile, the second section describes the structural classifications of the lexical bundles in the textbooks.

### Patterns of Lexical Bundles

54,009 lexical bundles were obtained from an analysis of the English textbooks. The patterns of the bundles which mainly occurred in the textbooks were: 3-word with 32,527 frequencies; 4-word with 11,620 frequencies; 5-word with 6,073 frequencies; and the smallest pattern was 6-word with 3,789 frequencies. The authors of the books tend to use three-word lexical bundles. The detailed results and examples of the patterns in the sentences are presented in the following subsections.

### Three-Word Lexical Bundles

The list of three-word lexical bundles extracted from the textbooks is displayed in Table 2. The most frequently used was "in this unit", followed by "what is the" and "answer the questions".

This study showed that the bundles appeared helpful as building blocks, even though they were less formulaic in capturing the whole meaning of chunks. The examples of three-word bundles are shown in the following sentences.

- (1) Learn those and more **in this unit** through challenging tasks.
- (2) **What is the** difference among a folktale, myth, and legend?
- (3) Listen to the conversation and then **answer the questions**.
- (4) But **I do not** remember where I left the book.

### Four-Word Lexical Bundles

Four-word lexical bundles serve as a suitable illustration of a given aspect (Kopaczyk, 2013). However, this pattern can be considered a more formulaic construction, since they create overlapping bundles and reappear in the form of longer ones.

**Table 2**

*Three-Word Lexical Bundles*

Rank	Lexical Bundles	Freq
1	in this unit	367
2	what is the	237
3	answer the questions	218
4	I do not	213
5	study the following	195
6	based on the	191
7	is in the	157
8	in the appendix	149
9	the listening script	142
10	listening script is	140
11	script is in	137
12	do you think	126
13	what do you	125
14	let us make	122
15	of the story	121

The four-word lexical bundles consist of 11,620-word combinations. According to Table 5, the highest frequency of four-word lexical bundles was "the listening script is", with 140 frequencies. Therefore, all the bundles included in this pattern were the most formulaic in understanding the whole meaning of the chunks. The examples of four-word lexical bundles in the sentence are as follows.

- (5) **The listening script is** in the Appendix.
- (6) The listening script **is in the Appendix**.
- (7) The **listening script is in** the Appendix.
- (8) The listening **script is in the** Appendix.
- (9) **Let us make a** reflection.

As seen from the examples, the meaning of the whole chunks in this pattern can be fully understood. However, the four-word lexical bundles tend to overlap, since the meaning of the three-word was difficult to capture. Even though examples (5), (6), (7), and (8) showed that one sentence contained several four-word lexical bundles, their beginnings or endings were different. In example (7), the bundle was "listening script is in", and was "script is in the" in example (8), where the word "script is" was from the previous bundle. These examples showed syntagmatic overlaps to the right of the following bundle as the consequences of linear arrangement in the texts.

**Table 3**

*Four-Word Lexical Bundles*

Rank	Lexical Bundles	Freq
1	the listening script is	140
2	is in the appendix	138
3	listening script is in	137
4	script is in the	137
5	let us make a	118
6	then answer the questions	108
7	based on the following	88
8	in pairs study the	86
9	and answer the questions	78
10	in this unit you	73
11	and then answer the	67
12	I would like to	65
13	what do you think	64
14	I do not like	62
15	the words in the	60

### Five-Word Lexical Bundles

Five-word lexical bundles appear less frequently in conversation and academic prose (Biber et al., 1999). For this reason, Biber et al. (1999) set the lower cut-off for five- and six-word lexical bundles. However, the research teams set five times per million words for the five-word bundles, since the frequency cut-offs were somewhat arbitrary from one study to another, depending on the scope of the research.

Five-word lexical bundles were rarely found in the textbooks. Their examples are shown in the following sentences.

- (10) The **listening script is in the** Appendix.
- (11) The listening **script is in the Appendix**.
- (12) **The listening script is in** the Appendix.
- (13) Read the text again **and then answer the questions** below.

Some lexical bundles of a specific length are merged into longer ones by appending «+» symbols before or after them (Biber et al., 1999). For example, the four-word "*listening script is in*" is merged into the five-word "*listening script is in **the***". This means that the longer such strings, the more fixed and formulaic they are. The five- and four-word bundles also tend to overlap together. Nevertheless, the same-length bundles still represent different structural patterns (Kopaczyk, 2013).

**Table 4**

*Five-Word Lexical Bundles*

Rank	Lexical Bundles	Freq
1	listening script is in the	137
2	script is in the appendix	137
3	the listening script is in	137
4	and then answer the questions	63
5	in this unit you learn	58
6	after learning English in this	56
7	do not like in this	56
8	I do not like in	56
9	I like best in this	56
10	in this unit and write	56
11	learning in this unit and	56
12	let us check your competence	56
13	let us make a summary	56
14	not like in this unit	56
15	of the story	121

### Six-Word Lexical Bundles

Six-word bundles are structurally more complex, and often contain two or more embedded phrases and syntagmatic overlaps either to the left or right of a particular bundle (Kopaczyk, 2013). This pattern was the least common among the others. There were only 3,789 six-word bundles found in this study.

Based on Table 5, the highest frequencies of six-word bundles were "*listening script is in the appendix*" and "*the listening script is in the*", repeated around 137 times in five different English textbooks. The examples of this bundle are in the following sentences.

- (14) The **listening script is in the Appendix**.
- (15) **The listening script is in the** Appendix.
- (16) The improvement I have made **after learning English in this unit**.
- (17) What I **do not like in this unit**.

Even though this pattern rarely appeared in the English textbooks, it was more formulaic than the short ones, because the whole meaning of the chunks can be easily understood.

### Structural Classification of Lexical Bundles

The results of the structural classification of all lexical bundles found in senior high school English textbooks are shown

**Table 5**

*Six-Word Lexical Bundles*

Rank	Lexical Bundles	Freq
1	listening script is in the appendix	137
2	the listening script is in the	137
3	after learning English in this unit	56
4	do not like in this unit	56
5	I do not like in this	56
6	I like best in this unit	56
7	in this unit and write down	56
8	learning English in this unit	56
9	learning in this unit and write	56
10	like best in this unit	56
11	on your learning in this unit	56
12	this unit and write down your	56
13	this unit let us make a	56
14	this unit the improvement I have	56
15	what I do not like in	56

in Table 6. The examples of each structural classification are presented in the following subsections.

### ***Noun Phrase with of- Phrase Fragment***

The noun phrase with of- phrase fragment had the most significant number of lexical bundles among the noun structures - 173. The bundles in this classification indicate quantities, qualities, or measurements and also describe events (Salazar, 2014). Some examples found in English textbooks are as follows:

- (18) Therefore, by buying local groceries we are helping reduce **the amount of** carbon dioxide produced.
- (19) **The beginning of** the story has the function of introducing the characters of the story, where and when the story occurred.
- (20) You advise him to reduce **the use of** pesticides and start using organic farming method as it produces healthier products.

These examples show that this structure always requires a complement (Kopaczyk, 2013). The structure is built from Det + N+ of+ xNP. The xNP part can be filled by various complements. The first example of the bundle "the amount of

" indicates the quantities, the second "the beginning of " represents the description of events, and the last "the use of " describes qualities.

### ***Noun Phrase with Other Post Modifier Phrase Fragment***

There were two types of this bundle structure, namely noun phrases with post-nominal clause fragments and noun phrases with prepositional phrase fragments. Examples (21) and (22) show a noun phrase with a post-nominal clause fragment, used to identify the relationship among entities. Meanwhile, examples (23) and (24) indicate a noun phrase with a prepositional fragment.

- (21) Listen carefully to **the dialogue between** Denias and Retno.
- (22) Study the words below before you listen to **the conversation between** Anita and a librarian.
- (23) Let us make **a summary in** this unit.
- (24) The purpose of hortatory exposition texts is to argue **a case for** or against a particular position or point of view and it proposes a suggestion at the end of the argumentation.

**Table 6**

*Structural Classification of Lexical Bundles in the Textbooks*

Category	Structural Classification of Lexical Bundles	Frequency	Percentage
<b>A Phrasal</b>			
<b><i>NP- based</i></b>			
1	Noun Phrase+ of	173	17,35%
2	Noun Phrase+ other post modifier	44	4,41%
3	Other Noun Phrase	157	15,75%
<b><i>PP- based</i></b>			
4	Prepositional Phrase+ of	13	1,30%
5	Other Prepositional Phrase	243	24,37%
<b><i>VP- based</i></b>			
6	Anticipatory <i>it</i> + Verb Phrase/ Adjective Phrase	13	1,30%
7	Passive Verb+ Prepositional Phrase	19	1,91%
8	Copula <i>be</i> + Noun Phrase/Adjective Phrase	30	3,01%
9	Pronoun/ Noun Phrase+ <i>be</i>	7	0,70%
<b>B Clausal</b>			
10	(Verb Phrase+) <i>that</i> - clause	59	5,92%
11	(Verb/ Adjective+) <i>to</i> - clause	239	23,97%
12	Adverbial Clause	0	0,00%
<b>TOTAL</b>		<b>997</b>	<b>100,00%</b>

### Other Noun Phrases

This classification was only for the remaining lexical bundles which did not fit with other noun structure classifications (Biber et al., 1999). The results showed that 157 bundles of other noun phrase structures were found in the English textbooks. The examples are as follows.

- (25) The purpose of today's show is to raise students' and youths' awareness to participate in saving our earth from **the global warming**.
- (26) **The little boy** pressed his face against the chain link fence.
- (27) In a single mouthful, the wolf swallowed **the old lady**.

This bundle structure was built from determiner + noun. However, there was no specific function for this classification, since the remaining bundle is not included in the first and second noun structures.

### Prepositional Phrase with Embedded of- Phrase Fragment

Prepositional phrase fragments were the primary group of lexical bundles that were mainly formulaic. The results revealed there were only 13 bundles found in the textbooks. Furthermore, the bundles, consisting of prepositional phrases with embedded of- phrase fragments, serve as the post-modifier of the noun (Biber et al., 1999). Therefore, the bundles in these classifications commonly signified the abstract or logical relationship between prepositional elements. The examples are as follows.

- (28) **As a matter of fact**, the government has planned some strategies to do that.
- (29) The purpose of hortatory exposition text is to argue a case for or against a particular position or point of view and it proposes a suggestion **at the end of** the argumentation.
- (30) **In groups of four**, read the following discussion and then summarize it.

Example (28) showed that the bundle "as a matter of fact" served to mark the fact. The bundles beginning with a preposition "at," as in Example (29), signify temporal relations. Other bundles with a preposition "in," as in Example (30), indicate quantities.

### Other Prepositional Phrases (Fragment)

Fragment was the most frequently found in the English textbooks. This classification had the highest number, with a

total frequency of 243 lexical bundles. The examples are as follows.

- (31) Find the equivalents in Indonesia **according to the** context.
- (32) The old king's birthday arrived, and the two oldest daughters brought him presents that were very necessary, but **at the same time** extremely expensive.
- (33) After that, use the words **on the left** column to complete the sentences.
- (34) Some countries have already banned smoking **in public places**, like Italy and New Zealand.
- (35) **On the other hand**, not many students have the opportunity to compete winning the scholarships.

The lexical bundle "according to the" was very common in this classification as it occurred 61 times. In example (31), several bundles refer to something or the text itself (Salazar, 2014). Another very common example in this classification is "at the same time". This bundle contrasts two propositions or events considered compatible (Biber et al., 1999). Example (32) showed the contrast between the very necessary and the costly presents. They were considered compatible with each other. Examples (33) and (34) were used to identify orientation as "on the left" and place as "in public places". Furthermore, example (35) showed a comparison between students who had the opportunity to win scholarships and those who did not. This indicates that other prepositional phrase fragments are also used to compare or establish temporal and logical associations between different concepts (Salazar, 2014). Since this structural classification is the most frequent, the authors of the book often use this structure in the teaching materials. This can raise students' awareness of lexical bundles and their use in spoken and written texts. Therefore, the students can improve their fluency in learning a second language.

### Anticipatory it + Verb Phrase/ Adjective Phrase

An adjective controls the beginning of lexical bundles with an anticipatory it pattern, and few of them are controlled by a verb phrase (Biber et al., 1999). The results showed there were 13 lexical bundles for this classification. The bundles beginning this pattern mostly feature a predicative adjective followed by to- or that- clause (Salazar, 2014). The examples are as follows.

- (36) **It is nice to** meet you.
- (37) When "Dateline CNBC" catches enough of these freaks to have a weekly show on it, **it is time to**

start showing people that we will not tolerate perverts touching our kids.

- (38) Based on this, **it can be assumed that** the persona in a poem is not necessarily the poet himself.

It was found that most of the bundles in this classification were “anticipatory it” with a predicative adjective followed by a to-clause. However, there were no bundles with “anticipatory it” and a predicative adjective followed by that-clause in the English textbooks. Examples (36) and (37) contain “anticipatory it” with a predicative adjective, as in the word “nice” and “time” followed by a to-clause. In contrast, a few numbers of lexical bundles in this classification were included in the verb predicative followed by that-clause as in Example (38).

### **Passive Verb + Prepositional Phrase Fragment**

The prepositional phrase shows a logical connection instead of specifying the agent in a by-phrase (Biber et al., 1999). The result showed 19 lexical bundles were in this classification and the examples are as follows.

- (39) Questions 12 - 16 **are based on the** following text.
- (40) Usually, it **is based on the** criteria of weak and strong arguments.

The bundles “is/are based on the” in Examples (39) and (40) occurred most frequently in the textbooks. The bundle “are based on the” occurred 42 times in five different English textbooks, while “is based on the” occurred ten times. Both examples showed the relation between questions and texts (39) as well as arguments and criteria (40).

### **Copula be + Noun Phrase/ Adjective Phrase**

All lexical bundles in this classification are initiated with the copula “be”. The two types of this classification depend on the subject’s predicative, whether a noun phrase or an adjective phrase (Biber et al., 1999). There were only 30 bundles for this classification, most of which were followed by a noun phrase.

- (41) Washington is home to numerous national landmarks and **is one of the** most popular tourist destinations in the United States.
- (42) You **are the one** who gave me the idea about the story I told the class.
- (43) Spoof **is a type** of story which has a twist (funny part in the end of the story)

The most frequent lexical bundle in the textbooks was “is one of the” as in Example (41). It occurred 19 times, followed

by “are the one” in example (42), which appeared 18 times. The other examples are as follows.

- (44) I **am sorry to** hear the news today.
- (45) Really? I **am very happy to** hear you say so.
- (46) Well, I think those **are the main** goals, but you know, we can add more goals later on.

These examples show the use of a copula followed by an adjective phrase, such as “sorry” (44), “very happy” (45), or “are the main” (46). They are used to express emotions.

### **Pronoun/ Noun Phrase+ be**

The bundles in this classification were less frequently found in the textbooks and the results showed only seven lexical bundles.

- (47) **There are some** mistakes on the spelling.
- (48) Total language death occurs when **there are no** speakers of a given language idiom remaining in a population where the idiom was previously used (i.e., when all native speakers die).

These bundles had the copula “be” as the verb, and the subject pronoun can be “this” or “there”. In this case, there were no bundles with the subject pronoun “this”. Lexical bundles with “there” as the subject pronouns were used for informational packaging purposes, as seen in the two examples above. The bundle “there are some”, as in (47), was the most frequent as it appeared 20 times per million words. It was followed by “there are no” which occurred 11 times. Example (47) showed that the bundle “there are some” is used to give information about mistakes in spelling. Meanwhile, Example (48) showed “there are no” was used to inform that when there were no speakers, it would lead to language death.

### **(Verb Phrase+) that- Clause**

Lexical bundles consisting of a that-clause had a verb phrase in the main clause. There are two main types of this structure, namely those containing the main clause verb and those beginning with the that-clause (Biber et al., 1999). The results showed there were 59 lexical bundles for this classification, and all were included in the two classes. The examples are as follows.

- (49) **I think that** people who ride without wearing helmets are suicidal, deluded, or just dumb.
- (50) **I must say that** the movie was awesome, but the dialogue was sometimes rude.

The two examples used the verb phrase followed by that-clause, in order to emphasize propositional statements. In Example (49), the person wanted to emphasize their statement about people who ride without wearing helmets as well as their statement about the movie in Example (50). Other examples of this structure which only consist of that-clause are as follows.

- (51) Many parents also think **that it is** better for their children to work and earn money than to go to school.
- (52) The current problem is **that there is** still a possible danger from smoke since smokers can light cigar everywhere.
- (53) Does Ayu think **that there is a** possibility for the drug addicts to recover?

Typically, lexical bundles with that-clause often appear as a complement clause in a declarative structure. There are two types of this structure. The first was followed by "it" in (51) as the subject and the copula "is" as a verb. This bundle was also the most frequent, occurring 47 times per million words. The second type was followed by "there" in (52) and (53) as the subject with the present tense "is" as a verb.

#### ***(Verb/ Adjective+) to- Clause***

Lexical bundles in this structural classification consist of to-clauses or to-clauses that come after a predicative adjective or a verb phrase (Salazar, 2014). It was found that this classification appear in second place with 239 lexical bundles. The examples of lexical bundles used in the sentences are as follows.

- (54) Department for Transport that drivers are four times **more likely to** crash if they are holding a mobile or sending a text while at the wheel.
- (55) The shoes brought such a good price that this time he **was able to** buy enough leather for four pair of shoes.
- (56) The expressions above **are used to** accept an invitation and an offer.
- (57) It is your chance **to say that** you love her very much.
- (58) **To show that** he was the most powerful man in the kingdom.

There were three main classification types, namely predicative adjective+ to-clause, (passive) verb phrase+ to-clause, and simple to-clause. The bundles containing predicative adjective+ to-clause were commonly used to indicate likeli-

hood as in (54) and possibility/ ability as in (55). In contrast, those with verb phrases before to-clause were used to refer to something (56). The phrase was typically in the passive voice. Finally, the simple to-clause was often used to indicate purposes/aims, as in (57) and (58). It was also usually found at the beginning of the sentence (Salazar, 2014).

## **DISCUSSION**

The results showed that the shorter the bundles, the more frequent they were. The shortest or three-word bundle with 32,527 frequencies, exceeded the others in the five Senior High School English textbooks written by Indonesian authors. Furthermore, the least used pattern was the longest or the six-word bundles with only 3,789 frequencies. However, it was unexpectedly found that some word combinations, such as "I do not", were in the most frequent pattern of three-word bundles. According to Biber et al. (1999), it is commonly found in non-academic texts rather than academic textbooks. In a language learning context, such word combinations can be easily found in textbook dialogue samples. The topics of dialogue vary from formal to non-formal. They are a normal speaking exercise in textbooks used by EFL students, in this case, in Indonesia. Kopaczyk (2013) as well as Northbrook and Conklin (2018) also revealed that this pattern tended to appear in conversations, since they were too short to capture the whole meaning of a chunk. It was common, therefore, to come across the bundle "I do not" in an everyday conversation between peers at school: for example, in the line "I do not remember where I left the book." As the frequency of vocabulary usage in textbooks tends to increase as the grade level progresses (e.g., Cao et al., 2022; Paraschiv et al., 2023), the authors of the books might have also been aware of students' English proficiency levels across the country. In this way the use of simple expressions in the textbooks could catalyze their next levels of language proficiency.

Written dialogues can provide a practically linguistic explanation of using particular terms, vocabulary, or grammar. Since it has become commonplace for short lexical bundles to be found in EFL textbooks, particularly in the dialogue sections, four-word bundles, such as "I don't like," "I would like to," and "what do you think," are considered important for understanding particular grammar points, thereby making them high-frequency bundles relevant to study. These bundles can encourage students to express themselves in different ways, such as stating their dislike for swimming, offering assistance, or asking for opinions on an idea. Northbrook and Conklin (2018) argued that the word combinations in textbooks should be more versatile and open to interpretation, especially considering the topics that students are likely to discuss. Their study in junior high school English textbooks in Japan revealed that the phrase "do you have any pets" had little impact. Students were more likely to be interested in discussing their plans for the weekend with

friends, rather than someone's pets. In order to make textbooks more relatable, it was suggested that the word "ideas" be used instead of "pets" in this context (Northbrook & Conklin, 2018, p. 328). By using language which reflects real-life conversations, the students could easily apply it in both informal and academic settings (Ganji & Nasrabad, 2021). Thus, a corpus can provide valuable insights to make sure that the language used in textbooks is as natural as possible.

The research findings indicate that the lexical bundles were primarily derived from the instructional materials found in the textbooks, as shown in Table 2. This supports Allan's (2017) study which found that these bundles were often used in providing instructions. Moreover, as the textbooks featured a variety of text types, such as narratives, news, and speeches, the target bundles were also identified in these pieces, as in (18), (27), and (28). The presence of lexical bundles in instructions, reading materials, and exercises helped students acquaint themselves with and use these combinations of words in spoken and written communication (Abu-Rabiah, 2023; Cortes, 2004, 2006; Northbrook & Conklin, 2019; Northbrook et al., 2022). Furthermore, repeated exposure to three or more words in a specific order assisted students in retaining vocabulary and recalling the bundles more easily. This finding is consistent with Nekrasova's (2009) study which suggested that EFL students remembered word sequences better in short-term memory and recalled them with minimal modification. Learning such a common formulaic language is also useful in EFL writing, since students can quickly choose the appropriate words and phrases to express their ideas. In this regard, teachers can use corpus linguistics, among other techniques, to enhance students' writing skills (Birhan, 2018; Durrant & Mathews-Aydinli, 2011; Wijaya & Ardi, 2022). During teaching and learning activities, drawing explicit attention to lexical bundles, such as asking students to read instructions and introducing new chapters where lexical bundles usually appear, can raise students' awareness of these essential elements of academic language. They can enhance their ability to recognize, store, and process information over numerous exposures (Northbrook et al., 2022; Richter et al., 2022).

The findings of this study are consistent with those of Putri and Suhardijanto (2022) who examined bundles of words in legal documents in the Indonesian context. Both studies revealed that the most common types of bundles in senior high school English textbooks and legal documents were composed of three or four words. Three-word bundles were more prevalent in legal and academic writing because they utilized simpler expressions. They were less formulaic, and could be more easily learned through natural language acquisition. Conversely, the longest bundles, consisting of six or seven words, were the least frequently used. These findings suggested that Indonesians had a preference for three-word bundles in academic and legal writing, despite the differences in the specific bundles utilized in each corpus. The

textbooks and legal documents both favored three-word bundles, but textbooks used simpler language to facilitate comprehension for students. Legal documents employed more complex and formulaic bundles, in order to convey precise legal information. Therefore, teachers can assist advanced students in improving their writing skills in academic and legal domains by providing activities and materials that focus on teaching and practicing these lexical bundles.

With regards to structural analysis, both the current study and Putri and Suhardijanto (2022) divided lexical bundles into two categories: phrasal (dependent) and clausal (independent). However, a contrasting difference between the two studies was identified. Clause-based bundles were more dominant in Indonesian legal documents, in order to avoid multi-interpretations in legal documents. This preference for complete structure stems from the belief that they can convey more comprehensive and accurate meaning. Nevertheless, this finding contradicts the idea put forth by Biber and Barbieri (2007) that most lexical bundles do not have a complete sentence structure. In addition, the current study found that phrase-based lexical bundles, particularly prepositional phrases, such as "according to", were the most frequently used word combinations in the five English textbooks used in Indonesia. The high frequency of prepositional phrases can be attributed to the common writing instructions templates and the authors' levels of English language proficiency. Besides prepositional phrases, the authors of English textbooks exhibited a tendency to use more noun phrases with *of*-, such as "the beginning of" and "the amount of". These research-oriented bundles indicate quantities, qualities, or measurements and also describe events. This is consistent with Allen (2009) as well as Beng and Keong (2015), who reported that the most prevalent lexical bundles in their database were those associated with the noun phrases used in academic work.

Although this study sheds light on the lexical bundles of English textbooks in Indonesia, it has some limitations that may impact data collection and analysis. First, the number of data used was relatively limited, with 300,859 tokens from the five English textbooks used by Indonesian schools. Moreover, the textbooks used in this case were all written by Indonesian authors, with the same primary authors, and published by the same publishers. This narrows the possibility of obtaining a variety of data. Since the first authors were the same, it was most likely that the choice of words would be monotonous and less varied. Future studies are expected to employ a more significant number of English textbooks with different authors for more comprehensive findings. Secondly, the research team did not conduct qualitative data gathering, such as interviews with the textbook users, in this case, students and teachers, to justify the pedagogical implications of lexical bundles. Future researchers are expected to further examine how Indonesian students process the common word combinations they come across in their textbooks (cf. Northbrook & Conklin, 2019; North-

brook et al., 2022). Finally, it should be noted that this study did not compare the lexical bundles found in the textbooks to those used by native English speakers. In order to further investigate this issue, future studies can compare the lexical bundles in the textbooks to those found in established corpora, such as the Corpus of Contemporary American English (COCA) and the British National Corpus (BNC). The comparison can shed light on the use of English by non-native English speakers, since multilingual speakers now use it in various linguistically and culturally diverse communities of practices worldwide (Iswandari & Ardi, 2022; Nindya et al., 2022; Sugiharto, 2023).

## CONCLUSION

This study highlighted that the most frequent patterns of lexical bundles were three-word and four-word bundles. The least frequent pattern was the 6-word. In terms of structural classifications, eleven out of twelve were used in English textbooks. Furthermore, other prepositional phrase fragments structure include the most frequent classification which occurred 243 times. The authors tended to use more noun phrases with of-, such as "the beginning of". This means they predominantly used prepositions in the teaching materials. The pronoun/Noun phrase + be structure was the least systematic classification, and no adverbial clause fragment structure was found in the textbooks. The use of lexical bundles in English textbooks gives exposure to L2 students. The short bundles enable students to retain English vocabulary and make them easily recall short expressions in usage. Also, the frequent structural patterns found in textbooks, like "the beginning of" and "the use of," can prompt students to further expand various noun phrases in their writing and speaking. Thus, repeated exposure to the word combinations familiarizes students with them and allows their usage in oral and written communication.

## REFERENCES

- Abu-Rabiah, E. (2023). Evaluating L2 vocabulary development features using lexical density and lexical diversity measures. *LLT Journal: A Journal on Language and Language Teaching*, 26(1), 168-182. <https://doi.org/10.24071/llt.v26i1.5841>
- Akbulut, F. D. (2020). A bibliometric analysis of lexical bundles usage in native and non-native academic writing. *Journal of Language and Linguistic Studies*, 16(3), 1146-1166. <https://doi.org/10.17263/jlls.803583>
- Alasmary, A. A. (2022). Comparing lexical bundles across three advanced mathematical text types: A corpus-based genre-focused investigation. *SAGE Open*, 12(3), 21582440221113824. <https://doi.org/10.1177/21582440221113>
- Alfiandita, O. L., & Ardi, P. (2020). Collocations in the English «LKS» books for senior high school students. *Indonesian Journal of English Language Teaching and Applied Linguistics*, 4(2), 213-227. <https://dx.doi.org/10.21093/ijeltal.v4i2.463>
- Allan, R. (2016). Lexical bundles in graded readers: To what extent does language restriction affect lexical patterning? *System*, 59, 61-72. <https://doi.org/10.1016/j.system.2016.04.005>
- Allan, R. (2017). From do you know to I don't know: An analysis of the frequency and usefulness of lexical bundles in five English language self-study books. *Corpus Pragmatics*, 1(4), 351-372. <https://doi.org/10.1007/s41701-017-0016-9>
- Allen, D. (2009). Lexical bundles in learner writing: An analysis of formulaic language in the ALESS Learner Corpus. *Komaba Journal of English Education*, 1, 105-127.

## ACKNOWLEDGEMENTS

We would like to thank two anonymous reviewers for their constructive feedback. The first author also expresses gratitude to Lembaga Pengelola Dana Pendidikan (Indonesia Endowment Fund for Education) for awarding a scholarship to pursue a doctorate degree in English Language Education at Universitas Negeri Malang, Indonesia.

## DECLARATION OF COMPETING INTEREST

None declared.

## AUTHORS' CONTRIBUTION

**Priyatno Ardi:** Conceptualization; Formal analysis; Methodology; Project administration; Resources; Writing – original draft; Writing – review & editing.

**Yacinta Dinda Oktafiani:** Conceptualization; Investigation; Methodology; Writing – original draft; Writing – review & editing.

**Nugraheni Widianingtyas:** Methodology; Writing – original draft; Writing – review & editing.

**Olga V. Dekhnich:** Validation; Writing – review & editing.

**Utami Widiati:** Validation; Writing – review & editing.

- Apple, R. (2022). Lexical bundles in L2 English academic texts: Relationships with holistic assessments of writing quality. *System*, 110, 102899. <https://doi.org/10.1016/j.system.2022.102899>
- Beng, C., & Keong, Y. (2015). Functional types of lexical bundles in reading texts of Malaysian University English Test: A corpus study. *GEMA Online Journal of Language Studies*, 15(1), 77-90. <https://doi.org/10.17576/GEMA-2015-1501-05>
- Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. John Benjamins.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26, 263-286. <https://doi.org/10.1016/j.esp.2006.08.003>
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405. <https://doi.org/10.1093/applin/25.3.371>
- Biber, D., Conrad, S., & Finegan, E. (1999). *The Longman grammar of spoken and written English*. Pearson Education.
- Birhan, A. (2018). Effects of mastery learning instruction on engineering students' writing skills development and motivation. *Journal of Language and Education*, 4(4), 20-30. <https://doi.org/10.17323/2411-7390-2018-4-4-20-30>
- Burch, B., Egbert, J., & Biber, D. (2017). Measuring and interpreting lexical dispersion in corpus linguistics. *Journal of Research Design and Statistics in Linguistics and Communication Science*, 3(2), 189-216. <https://doi.org/10.1558/jrds.33066>
- Cao, D. T. P., Pho, P. D., & Chi, D. N. A. (2022). The treatment of collocations in English textbooks for Vietnamese students. *TEFLIN Journal*, 33(2), 220-239. <http://dx.doi.org/10.15639/teflinjournal.v33i2/220-239>
- Chen, Y.-H., & Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology*, 14(2), 30-49. <http://dx.doi.org/10.125/44213>
- Conrad, S., & Biber, D. (2005). The frequency and use of lexical bundles in conversation and academic prose. *Lexicographica*, 20, 56-71. <https://doi.org/10.1515/9783484604674.56>
- 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.2003.12.001>
- 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>
- Durrant, P., & Mathews-Aydinli, J. (2011). A function-first approach to identify formulaic language in academic writing. *English For Specific Purposes*, 30, 58-72. <https://doi.org/10.1016/j.esp.2010.05.002>
- Fajri, M. S. A., Kirana, A. W., & Putri, C. I. K. (2020). Lexical bundles of L1 and L2 English professional scholars: A contrastive corpus-driven study on applied linguistics research articles. *Journal of Language and Education*, 6(4), 76-89. <https://doi.org/10.17323/jle.2020.10719>
- Ganji, M., & Nasrabad, P. (2021). An analysis of the multiword units presented in IELTS speaking preparation books published by Iranian authors. *Applied Research on English Language*, 10(3), 137-160. <https://doi.org/10.22108/are.2021.126390.1666>
- Gil, N. N., & Caro, E. M. (2019). Lexical bundles in learner and expert academic writing. *Bellaterra Journal of Teaching & Learning Language and Literature*, 12 (1), 65-90. <https://doi.org/10.5565/rev/jtl3.794>
- Hussain, G., Zahra, T., & Abbas, A. (2021). Discourse functions of lexical bundles in Pakistani chemistry and physics textbooks. *GEMA Online Journal of Language Studies*, 21(1), 221-238. <http://doi.org/10.17576/gema-2021-2101-13>
- Hyland, K., & Jiang, K. F. (2018). Academic lexical bundles: How are they changing? *International Journal of Corpus Linguistics*, 23 (4), 383-407. <https://doi.org/10.1075/ijcl.17080.hyl>
- Iswandari, Y. A., & Ardi, P. (2022). Intercultural communicative competence in EFL setting: A systematic review. *rEFLections*, 29(2), 361-380. Retrieved from <https://so05.tci-thaijo.org/index.php/reflections/article/view/260249>
- Jablonkai, R. (2009). "In the light of": A corpus-based analysis of lexical bundles in two EU-related registers. *WoPaLP*, 3(3), 1-27.
- Kim, S., & Kessler, M. (2022). Examining L2 English University students' uses of lexical bundles and their relationship to writing quality. *Assessing Writing*, 51, 100589. <https://doi.org/10.1016/j.asw.2021.100589>
- Kopaczky, J. (2013). *The legal language of Scottish burghs*. Oxford University Press.
- Kurniawan, E., & Haerunisa, Z. F. (2023). A comparative study of lexical bundles in accepted and rejected applied linguistic research article introductions. *Studies in English Language and Education*, 10(2), 628-648. <https://doi.org/10.24815/siele.v10i2.28119>
- Lee, H. K. (2020). Lexical bundles in linguistics textbooks. *Linguistic Research*, 37(1), 121-145. <https://doi.org/10.17250/khisli.37.1.202003.005>
- 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>

- McEnery, A. M., & Wilson, A. (2019). *Corpus linguistics: An introduction* (2<sup>nd</sup> ed.). Edinburgh University Press.
- Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics*, 32, 111–129. <https://doi.org/10.1017/S0267190512000128>
- Meunier, F., & Gouverneur, C. (2009). New types of corpora for new educational challenges. In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 179-201). John Benjamins Publishing.
- Nekrasova, T. M. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Language Learning*, 59(3), 647-686. <https://doi.org/10.1111/j.1467-9922.2009.00520.x>
- Nesi, H., & Basturkmen, H. (2009). Lexical bundles and discourse signalling in academic lectures. In J. Flowerdew & M. Mahlberg (Eds.), *Lexical cohesion and corpus linguistics* (pp. 23-44). John Benjamins.
- Nindya, M. A., Widiati, U., & El Khoiri, N. (2022). Understanding and practices of intercultural communicative competence elements perceived by English pre-service teachers. *International Journal of Language Education*, 6(3), 297-313. <https://doi.org/10.26858/ijole.v6i3.22747>
- Northbrook, J., & Conklin, K. (2018). "What are you talking about?" An analysis of lexical bundles in Japanese junior. *International Journal of Corpus Linguistics*, 23(3), 311–334. <https://doi.org/https://doi.org/10.1075/ijcl.16024.nor>
- Northbrook, J., & Conklin, K. (2019). Is what you put in what you get out? —Textbook-derived lexical bundle processing in beginner English learners. *Applied Linguistics*, 40(5), 816-833. <https://doi.org/10.1093/applin/amy027>
- Northbrook, J., Allen, D., & Conklin, K. (2022). 'Did you see that?'—The role of repetition and enhancement on lexical bundle processing in English learning materials. *Applied Linguistics*, 43(3), 453-472. <https://doi.org/10.1093/applin/amab063>
- Oktavianti, I. N., & Prayogi, I. (2022). Discourse functions of lexical bundles in Indonesian EFL learners' argumentative essays: A corpus study. *Studies in English Language and Education*, 9(2), 761-783. <https://doi.org/10.24815/siele.v9i2.23995>
- Panthong, P., & Poonpon, K. (2020). Functional analysis of lexical bundles in doctor talks in the medical TV series Grey's Anatomy. *LEARN Journal: Language Education and Acquisition Research Network*, 13(2), 335-353.
- Paraschiv, A., Dascalu, M., & Solnyshkina, M. I. (2023). Classification of Russian textbooks by grade level and topic using ReaderBench. *Research Result. Theoretical and Applied Linguistics*, 9(1), 50-63. <https://doi.org/10.18413/2313-8912-2023-9-1-0-4>
- Putri, M. T. D., & Suhardijanto, T. (2022). Struktur dan fungsi gugus leksikal dalam teks peraturan perundang-undangan, surat resmi, dan surat bisnis (Structure and function of lexical bundles in legal texts, formal letters, and business letters). *Kandai*, 18(1), 1-21. <https://doi.org/10.26499/jk.v18i1.3322>
- Richter, K. G., Gaskaree, B. L., & Mirzai, M. (2022). A functional analysis of lexical bundles in the discussion sections of applied linguistics research articles: A cross-paradigm study. *Russian Journal of Linguistics*, 26(3), 625-644. <https://doi.org/10.22363/2687-0088-27752>
- Salazar, D. (2014). *Lexical bundles in native and non-native scientific writing: Applying a corpus-based study to language teaching*. John Benjamins Publishing Company.
- Sanosi, A. B. (2022). The use and development of lexical bundles in Arab EFL writing: A corpus-driven study. *Journal of Language and Education*, 8(2), 106-121. <https://doi.org/10.17323/jle.2022.10826>
- Scott, M. (1997). PC analysis of key words - And key key words. *System*, 25(2), 233–245. [https://doi.org/10.1016/S0346-251X\(97\)00011-0](https://doi.org/10.1016/S0346-251X(97)00011-0)
- Siricharoen, A., & Wijitsopon, R. (2020). A corpus-based comparative study of lexical bundles in authentic and textbook English business emails. *LEARN Journal: Language Education and Acquisition Research Network*, 13(2), 41-63.
- Stubbs, M. (2015). Computer-assisted methods of analyzing textual and intertextual competence. In D. Schiffrin, D. Tannen & H. E. Hamilton (Eds.), *The handbook of discourse analysis* (2<sup>nd</sup> ed.) (pp. 486-504). John Wiley & Sons.
- Sugiharto, S. (2023). From unequal Englishes to the praxis of decolonial fissure: Englishes in the Indonesian periphery. *Journal of Multicultural Discourses*, 0, 1-13. <https://doi.org/10.1080/17447143.2023.2210098>
- Wijaya, L. E., & Ardi, P. (2022). Rhetorical moves of research article abstracts: A comparative study of national and international journals. *Research Result. Theoretical and Applied Linguistics*, 8(1), 135-150. <https://doi.org/10.18413/2313-8912-2022-8-1-0-9>.
- Yusuf, K. (2020). Data driven learning by discovering lexical bundles using corpus resources. *Advances in Social Science, Education and Humanities Research*, 434, 47-50. <https://doi.org/10.2991/assehr.k.200427.011>