EXPLORING COLLOCATIONS IN BAHASA INGGRIS TEXTBOOK: A CORPUS STUDY

Raden Indrajid Lukman Sarjono¹, Ayuni Kabobu Heda² & Barli Bram³
¹ ² ³ Sanata Dharma University, Yogyakarta, Indonesia

DOI: 10.23917/humaniora.v23i2.14799

Received: June 7th, 2021. Revised: August 20th, 2022. Accepted: August 27th, 2022
Available Online: July 29th, 2021. Published: August 29th, 2021

Keywords: collocation, Bahasa Inggris textbook, grammatical, lexical

Abstract

This paper explored the types of collocations used in an English textbook entitled Bahasa Inggris (SMA/MA/SMK/MAK, Kelas X, Semester 1). The primary source of data was the textbook, entitled Bahasa Inggris, which was published by the Indonesian government for the first-year or class 10 students of senior high schools (SMA/MA/SMK/MAK). This study was a corpus analysis. The researchers utilized AntCont 3.5.8 to extract and discover the required data and then explored the types of collocations used in the Bahasa Inggris textbook. The results showed that the textbook comprised of Grammatical G1 (1807), Grammatical G3 (178), Grammatical G4 (1770), Grammatical G5 (1), Lexical L1 (594), Lexical L3 (125), Lexical L5 (21), and Lexical L6 (1). Grammatical G1 collocation was the most frequent type in the textbook. The results indicated that it would be beneficial for English teachers to provide collocation materials for the students to accommodate their writing and speaking skills.

Abstrak

Makalah ini mengeksplorasi jenis-jenis kolokasi yang digunakan dalam buku teks bahasa Inggris berjudul Bahasa Inggris (SMA/MA/SMK/MAK, Kelas X, Semester 1). Sumber data utama adalah buku teks berjudul Bahasa Inggris, yang diterbitkan oleh pemerintah Indonesia untuk siswa tahun pertama atau kelas 10 Sekolah Menengah Atas (SMA/MA/SMK/MAK). Penelitian ini merupakan penelitian analisis korpus. Peneliti menggunakan AntCont 3.5.8 untuk mengekstrak dan menemukan data yang diperlukan dan kemudian mengeksplorasi jenis kolokasi yang digunakan dalam buku teks Bahasa Inggris. Hasil penelitian menunjukkan bahwa buku teks terdiri dari Grammatical G1 (1807), Grammatical G3 (178), Grammatical G4 (1770), Grammatical G5 (1), Lexical L1 (594), Lexical L3 (125), Lexical L5 (21), dan Lexikal L6 (1). Kolokasi G1 gramatikal adalah jenis yang paling sering ditemukan dalam buku teks. Hasil analisis menunjukkan bahwa akan bermanfaat bagi guru bahasa Inggris untuk menyediakan materi kolokasi bagi siswa untuk mengakomodasi keterampilan menulis dan berbicara mereka.
INTRODUCTION

Since textbooks play a crucial role in most Indonesian classrooms, many teachers use textbooks as guidance in the teaching process. Richards (2014) defines a textbook as one of teaching sources used to support learning through stimulating cognitive processes and providing structure and progression for learners to follow. Textbooks were utilized for a long time in the history of education in Indonesia. Teachers and students use textbooks to ease the learning and teaching process. McCutcheon (1982) argues in his article that “we have statistics pointing to an extensive use of textbook, and some indication that this has been the case for at least 50 years, but we know little about the nature, character, and qualities of teachers’ and students’ use of textbooks” (p. 3). There are three benefits of the use of textbooks for teachers, namely: (1) inspections of the textbook as a subject authority, (2) principles about what should be taught, and (3) information of the subject substance. (Freeman & Schmidt, 1982; Talmage, 1972; Barr, 1988; McCutcheon, 1982). As an international language, Indonesian students also learn English as a foreign language. The English textbook provided by the Indonesian government is one of the sources for the teacher to guide the learning process. Weninger and Kiss (2013) state that English textbooks play an essential role in achieving social transformative and cultures. One of the items that are often found in an English textbook is collocation.

The theory of collocations was foremost recognized by Palmer (1933, as quoted in Nation, 2002, p. 317) as a sequence of words that “must be learned, or best conveniently learned as an integral whole or independent entity, rather than by the process of piecing together their parts.” In this study, the researchers took on Benson, Akella, and Maltz (2010) and Lewis (2000) as the theoretical framework. They divided collocations into two elements: grammatical and lexical collocation. Grammatical collocation encompasses one lexical word or dominant word, such as noun, verb, or adjective alongside grammatical structure like infinitive or clause (see Table 1). On the other hand, lexical collocation is the mixture of two lexical elements that usually do not include some preposition, infinitives, or clauses. It can be seen in Table 2.

Table 1. Grammatical Collocations

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Structure</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>noun + preposition combination</td>
<td>Give away personal information on online forums and discussion boards.</td>
</tr>
<tr>
<td>G2</td>
<td>noun followed by to + infinitive</td>
<td>She makes an apple pie to give to her birthday friend.</td>
</tr>
<tr>
<td>G3</td>
<td>noun followed by that clause</td>
<td>The vacation that they had been planning is canceled.</td>
</tr>
<tr>
<td>G4</td>
<td>preposition + noun combination</td>
<td>I am not interested in games they play.</td>
</tr>
<tr>
<td>G5</td>
<td>adjective + preposition combination</td>
<td>I am afraid of the darkness.</td>
</tr>
<tr>
<td>G6</td>
<td>adjective + to + infinitive</td>
<td>It is difficult to get into Ms. Ayu’s class.</td>
</tr>
<tr>
<td>G7</td>
<td>adjective + that clause.</td>
<td>It is hard that she should leave him.</td>
</tr>
</tbody>
</table>

Sources: Benson, Akella and Maltz (2010); Lewis (2000)

Table 2. Lexical Collocations

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Structure</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>verb (transitive) +</td>
<td>The students speak English fluently</td>
</tr>
</tbody>
</table>
Yule (2017) says that collocation is a cluster of words which appear mutually. Collocation has a crucial role in the English language learning process. It is needed to be learned by the students in Indonesia because they can distinguish the word combinations suitable for a phrase. Lewis (2000) and Woolard (2000) mentioned that “by learning collocation, the students can increase their vocabulary knowledge and be reactivate in the half-known words in the form of word combinations”. The students can construct genuine English when they have a high level of collocation competency (Alfiandita & Ardi, 2020). Many students are not interested in learning English because they do not have basic knowledge of English structure. Nesselhauf and Tschichold (2002) say that students tend to compose a grammatical inaccuracy because of the lack of collocation knowledge. They cannot recognize the proper word combinations to express their ideas. Henriksen (2013) adds, “Producing inappropriate collocations will eventually lead to a misunderstanding between the addressee and the addressees”.

Since English textbooks and worksheets are one of the resources for discovering English collocations, some studies have explored collocations in such sources. Alfiandita and Ardi (2020) studied the types of variations of collocations used in English LKS (Lembar Kerja Siswa) or Student Worksheets. They found out that the variety of collocations in the English LKS was limited. Roohani (2011) also compared the pre-university and high school textbooks from the Iranian government with the New Interchange book series. The findings revealed that the New Interchange books had a larger amount of collocations than the pre-university and high school textbooks. Koya (2004b) also linked the UK history books with the English textbooks in Japan and revealed that the UK history books have a small number of collocations than the Japanese English books caused by the point that the UK history books merely absorbed the history instead of the grammatical issues.

Previous studies reported collocations as scarce and even absent in some/many English textbooks. Based on the gap, this study aimed to explore the use of collocations in an English textbook provided by the Indonesian government. In the light of the above review, the study aimed to explore the types of collocations used in the textbook Bahasa Inggris (SMA/MA/SMK/MAK, Kelas X, Semester 1) for senior high school students.

METHOD

The study was corpus-based research. A research based on corpus linguistics. Corpus linguistics is the study of language that emphasizes the importance of studying patterns of daily language or real-life language use in linguistics research (Koteyko, 2006; Hasko, 2012; McEnery & Wilson, 2001). Moreover, corpus linguistics data are collected from computer files in the form of the written text or the huge collection of more than one written text from daily life (Baker, 2010; McEnery & Wilson, 2001). This study was employed to analyze patterns of collocations in an English textbook.
entitled *Bahasa Inggris*. In order to extract the data, AntCont 3.5.8 was used to discover the token and types of words (Anthony, 2004, 2005).

**Instruments and Data Collection**

The data were collected from an English textbook entitled *Bahasa Inggris* provided by the Indonesian government for the first grade of senior high schools (SMA/MA/SMK/MAK). The book *Bahasa Inggris* was written by Widyati, Furaidah and Rohmah (2014) and published by Pusat Kurikulum dan Perbukuan, Balitbang Kemdikbud. There are two steps undergoing in obtaining the data. First, the English textbook was converted into a Word file using an online PDF converter, namely www.pdf2go.com, and then moved to Notepad software to obtain (.txt) document. Second, utilizing AntCont 3.5.8 to extract and discover the required data. The process of extracting data was started with opening the AntCont 3.5.8 software, selecting “file”, clicking “open file(s)”, choosing the (.txt) document from the computer, and selecting “word list” and “start” at the bottom of the page. Furthermore, the collocation menu in AntCont 3.5.8 was used to analyze the data further.

**Data analysis**

The data analysis was started with reading the collected data to fully comprehend ones. The data gathered were the words and letters in an English textbook. Not all of them were included in the analysis of collocation types. Thus, the data were classified based on fourteen types of colocation by Lewis (2000) and Benson et al. (2010). The words used in this study were taken from the parts of speech in English, namely the noun, preposition, adjective, verb, and adverb. Then, AntCont 3.5.8 was utilized to obtain the first ten top function and content words from the list. Next, the twenty gathered words were the basis for analyzing the types of collocation.

The fourteen types of collocation were further simplified into two categories: lexical and grammatical collocation. The categories were labelled based on the collocation types, namely G1, G2, G3, G4, G5, G6, and G7 for grammatical collocation and L1, L2, L3, L4, L5, and L7 for lexical collocation (Alfiandita & Ardi, 2020). G1 consisted of noun + preposition combination; G2 consisted of noun + to + infinitive; G3 noun + that clause; G4 preposition + noun combination; G5 adjective + preposition combination; G6 predicate adjectives followed by to + infinitive; and G7 adjective + that clause.

The lexical collocation was formed by L1 consisted of verb + noun (CA collocation); L2 verb + noun (EN collocation); L3 adjective + noun; L4 noun + verb; L5 the unit that is associated with a noun; L6 adverb + adjective; and L7 verb + adverb. Lastly, the data of 14 types of collocation were scrutinized to discover the most frequent type of collocation that existed in an English textbook entitled *Bahasa Inggris* provided by the Indonesian government.

**RESULTS AND DISCUSSION**

The data from the textbook *Bahasa Inggris* were extracted to obtain the words and their numbers by utilizing AntConc 3.5.8 to explore the types of collocation in *Bahasa Inggris* textbook. The ten top function words are shown in Table 3.

**Table 3. Top Ten Function Words in *Bahasa Inggris* Textbook**

<table>
<thead>
<tr>
<th>No.</th>
<th>Function Word</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>the</td>
<td>1818</td>
</tr>
<tr>
<td>2</td>
<td>to</td>
<td>935</td>
</tr>
</tbody>
</table>
The findings indicated that ten function words commonly appeared in the textbook. Function words were words that had lexical meaning and expressed grammatical relationships between other words in a sentence. The function words played an essential role in sentence structure. According to Corver and van Riemsdijk (2013), function words include a preposition, article, auxiliary verb, conjunction, and pronoun. In table 2, the word “to”, “and”, “in”, and “of” are the preposition. In this textbook, the words “a” and “the” are the article. The next, the word “I”, “you”, and “your” is the part of pronoun words. The word “is” in the textbook is used as an auxiliary verb.

The finding was about content words in the English textbook. Content words contain nouns, verbs, adjectives, and adverbs (Hasselmath, 2004). The ten top content words are shown in Table 4.

**Table 4. Top Ten Content Words in Bahasa Inggris Textbook**

<table>
<thead>
<tr>
<th>No.</th>
<th>Content Word</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>have</td>
<td>174</td>
</tr>
<tr>
<td>2</td>
<td>do</td>
<td>162</td>
</tr>
<tr>
<td>3</td>
<td>task</td>
<td>151</td>
</tr>
<tr>
<td>4</td>
<td>when</td>
<td>138</td>
</tr>
<tr>
<td>5</td>
<td>picture</td>
<td>118</td>
</tr>
<tr>
<td>6</td>
<td>semester</td>
<td>115</td>
</tr>
<tr>
<td>7</td>
<td>word</td>
<td>114</td>
</tr>
<tr>
<td>8</td>
<td>text</td>
<td>109</td>
</tr>
<tr>
<td>9</td>
<td>question</td>
<td>102</td>
</tr>
<tr>
<td>10</td>
<td>like</td>
<td>91</td>
</tr>
</tbody>
</table>

Content words provided in the data showed that the high frequency of those words were nouns and verbs. The words “picture”, “text”, “questions” could be identified as a noun or verb. The words “do”, “have”, and “like” could be considered as verbs. However, the word “like” could be categorized into noun, verb, adjective, and adverb since its flexible range of uses. Then, the words “task”, “semester”, and “word” were nouns. Moreover, the word “when” was classified into adverbs.

The data of function and content words were further combined to obtain the majority of collocations used those twenty words as presented in Figure 1. Figure 1 shows the distribution of collocations utilizing function and content words. The types of collocations were divided into two categories, namely, G1, G2, G3, G4, G5, G6, and G7 collocation for grammatical. L1, L2, L3, L4, L5, L6, and L7 collocation for lexical. The findings displayed that there were eight (8) out of fourteen (14) collocations that appeared in the English textbook. G2, G6, G7, L2, L4, and L7 collocations were not found in the textbook. It also indicated that G1 collocation, whose pattern noun + preposition combination, was the most productive collocation. Then, G4 collocation,
whose pattern preposition + noun combination was the second after G1 collocation. Meanwhile, the lexical L1 collocation was the highest among other lexical collocations.

Figure 1. Collocation Types

1.1 Grammatical G1 Collocation
Grammatical G1 collocation was in the first position among other collocations with frequencies of 1807. Benson et al. (2010) mention that collocation subsists of nouns followed by preposition combination (namely books in a huge library). As Grammatical collocation 1 consists of a preposition; the dominant findings of G1 are indicated by the word “in” with the total number of 967 and “of” with the total number 840. The following are some examples of grammatical collocation 1:

Datum 1: Medical corps deployed for a peace mission in a conflict area.

Sentences 1 and 2 show the noun followed by the preposition alignment transpires in “a peace mission in” and “scuba diving in”. Moreover, the preposition “of” appeared in G1 collocation, which was in a second position after “in”. It can be concluded that the preposition “in” and “of” were paramount in this type of collocation. There were seven words considered as the noun in G1 collocation, namely “task”, “picture”, “word”, “semester”, “text”, “questions,” and “like,” which the most frequent noun is ‘words’ as presented in the following example:

Datum 3: Complete the following sentences using the words in the box.
Datum 4: Try to catch the words in the song.

Sentences 3 and 4 refer to G1 collocation in the English textbook. The word “words” was the noun followed by the preposition “in” to form the G1 collocation in both sentences.

1.2 Grammatical G3 Collocation
Grammatical G3 collocation posited the fourth of all of the collocation types. As provided in figure 1 that this collocation was under L1 collocation. G3 collocation was formed from a noun + that clause, which frequency was 178. The examples of G3 collocation were shown in the following sentences.
Datum 5: I give an example that Malin Kundang was a healthy and diligent man.
Datum 6: My grandparent’s village that I visited last year.

Sentences (5) and (6) indicate that “an example” was a noun and “that Malin Kundang was a healthy and diligent man” was that clause. Moreover, “village” was a noun, and “that I visited last year” was the clause. Both sentences were the evident G3 collocation whose noun pattern followed by the clause existed in the English textbook. It also appeared that the most frequent noun was the word ‘words’, whose frequency was 30 in the textbook. It was indicated in the following sentences.

Datum 7: Fill in the gaps with the words that you hear in this home.
Datum 8: You have to guess what words that your teacher has described.

Sentences (7) and (8) contained the word “words” as the noun and followed by that clause from both clauses “the words that you hear in this home” and “words that your teacher has described”.

1.3 Grammatical G4 Collocation

Collocation of grammatical G4 contains preposition + noun combination (Benson et al., 2010). This type of collocation was in the second position after G1 collocation based on Figure 1. The frequency for this collocation was 1770. The words “of” and “in” were the most frequent words that appeared in the textbook. There were 887 G4 collocation patterns formed by the word “of”.

Datum 9: Now write an entry of a diary from the point of view.
Datum 10: Give some examples of other ecotourism destinations that you know.

Sentences 9 and 10 indicated the G4 collocation patterns. It was proven in sentence 9 that the word “of” was the preposition followed by “a diary from the point of view” as the noun combination. Meanwhile, sentence 10 had the word “of” as the preposition and “other ecotourism destinations that you know” as the noun combination. Moreover, other words appeared in G4 collocation, namely “to”, “question”, “when”, “like”, “picture”, “text”, and “word”. The only absent noun from the G4 collocation in the textbook was “semester”. The following sentence was the example of the word “to” and “question” that appeared together.

Datum 11: Discuss the answer to the questions with your classmates.

In sentence 11, the word “to” was functioned as a preposition and the word “the question” was the noun. It was evident that sentence 11 followed the G4 collocation pattern and showed the possibility of noun combinations from the function and content words.

1.4 Grammatical G5 Collocation

Collocation of grammatical G5 is formed by combining adjective + preposition combination. The English textbook provided this kind of collocation which the frequency was only 1. The following sentence was an example of grammatical G5 collocation.

Datum 12: However, it is infrequent to find noun phrases with more than 3 modifiers like in the examples in the table above.
Sentence 12 indicated that the word “like” was functioned as an adjective, and “in the examples” was a preposition combination with “in” as the preposition.

1.5 Lexical L1 Collocation

Lexical L1 collocation was the highest among other collocations. The collocation consists of verbs followed by nouns/pronouns. The verb has to be a transitive verb which requires an object. Based on the findings, 594 lexical L1 collocations were identified in the English textbook for senior and vocational high school students. (explain based on the chart). Moreover, there were some words not included in the findings from the top twenty words, namely, “the”, “to”, “a”, “in”, “and”, “of”, and “is”. Those words belong to the function words. The words “questions”, “picture” and “text” were excluded since they played a role as nouns. Then, the word “have” was the most frequent word appearing in the English textbook with 594 frequencies. The following example showed the collocation of “have” in the sentence.

Datum 13: Once upon a time, there was an old couple who didn’t have a child.
Sentence 13 displayed that the word “have” shows a verb and the word “a child” as the noun. The word “have” was a transitive verb since it needed a direct object. On the other hand, the word “do” was the least frequent lexical L1 collocation in the textbook. The examples are as follows.

Datum 14: I am glad you do it for me.
Datum 15: Let’s do something together this long weekend.

The word “do” was found to be 461 in total. The provided examples indicated that the word “do” was a verb, and the words “it” and “something” were pronouns and nouns. Further analysis revealed that the function of the word “do” was primarily used in interrogative sentences.

1.6 Lexical L3 Collocation

Collocation of type Lexical L3 is formed by combining adjective + noun. The findings suggested that this type of collocation was mostly used as instructions to perform tasks. The word “text”, “questions”, “words” commonly existed in the textbook with 125 frequencies. Moreover, the most frequent word was “text”, as presented in the following sentence.

Datum 16: Read the following text.

Sentence 16 showed that the word “following” was the adjective and “text” was the noun. Meanwhile, the least frequent word was “words” in lexical L3 collocation. The example was provided in datum 17.

Datum 17: Jono has harsh words for him.
Sentence 17 consisted of “harsh” as an adjective and “words” as a noun. The word “words” occurred 17 times in an English textbook.

1.7 Lexical L5 Collocation

This lexical collocation was found to appear 21 times in the English textbook in which the words were “words” and “text”. L5 collocation consisted of the unit that is associated with a noun. The examples were provided in the following sentences.

Datum 18: Below are lists of words related to tourist destinations.
Datum 19: Does your friend use a suitable choice of words?
Sentence 18 indicated that the word “lists” was associated with the unit and the word “words” was a noun. Sentence 19 also showed that the word “suitable choice” was found as a unit followed by “words” as a noun. Thus, this type of collocation can be classified as a combination of first noun and second noun.

1.8 Lexical L6 Collocation

The way to build this collocation was by combining adverbs + adjectives. The findings indicated that there was one sentence that existed in L6 collocation. The example was provided below.

Datum 20: When strong wind was coming, his sister asked.

Sentence 20 stated that the word “when” was the adverb and followed by “strong” as the adjective to modify “wind”. Lexical L6 collocation was infrequent in the English textbook.

DISCUSSION

Based on the results of analysis, the findings indicated that there were eight collocations for both grammatical and lexical that appeared in the textbook Bahasa Inggris. The analysis revealed that the total frequency of grammatical collocation was 3756. Meanwhile, the total number of lexical collocations was 741. It can be concluded that grammatical collocation, specifically G1 and G4 were the most frequent collocation among nine collocations.

Moreover, grammatical collocations in the English textbook dominated the collocation types because of the use of the prepositions “of” and “in” in G1 and G4 collocations. Both of the function words were the basis to build the majority of sentences in the textbook Bahasa Inggris. Alfiandita and Ardi (2020) agreed with the notion that grammatical collocations provided significant numbers in English LKS books which further indicated the minimum attention to the use of collocation types in the book. Further, Martynska (2004) showed the importance of collocational competence and its role in second language learning. Thus, the ability to combine words in the target language is imperative for non-native speakers of English in achieving a satisfying level of proficiency (Martynska, 2004).

The collocation study results imply that facilitating students to learn the target language could be initiated by exposing the concept of collocation in English textbooks. This is in line with Jaén (2007), who states that it is an accepted idea that collocations are a significant part of the knowledge of English language learning in enhancing students’ speaking or writing skills. In other words, it would be advantageous for English teachers to provide their students with frequent collocation examples to enable to use collocations appropriately and improve their language skills, particularly writing and speaking skills.

CONCLUSION

The study analyzed 14 types of collocations, which were categorized into two, namely grammatical and lexical collocations. The findings indicated that there were eight collocations for both grammatical and lexical that appeared in the textbook Bahasa Inggris. The analysis revealed that the total frequency of grammatical collocation was 3756. Meanwhile, the total number of lexical collocations was 741. It can be concluded that grammatical collocation, specifically G1 and G4 were the most frequent collocation
among nine collocations. Moreover, grammatical collocations in the English textbook dominated the collocation types because of the use of the prepositions “of” and “in” in G1 and G4 collocations. Both of the function words were the basis to build the majority of sentences in the textbook Bahasa Inggris.

The analysis proved that the English textbook’s authors had not provided the variations of collocations to assist students’ language process and learning. The textbook requires collocations to enrich the students’ comprehending of English. Furthermore, it will be beneficial for students if English teachers provide collocation materials in their teaching-learning process. As supported by Jaén (2007) that it is an accepted idea that collocations are a significant part of the knowledge of English language learning in enhancing students’ speaking or writing skills. Exposures to collocations will accommodate students to write and speak fluently and accurately.

REFERENCES


