A Corpus-based Analysis of Conversational Features in *Bahasa Inggris* Textbooks for Junior High Schools in Indonesia

Nurinsani Yaman [a,1,*], Ikmi Nur Oktavianti [b,2]

[a, b] Universitas Ahmad Dahlan, Jl. Ringroad Selatan, Kragilan, Tamanan, Bangunapan, Bantul, Yogyakarta, 55191, Indonesia
[1] nurinsani1900004033@webmail.uad.ac.id *; [2] ikmi.oktavianti@pbi.uad.ac.id
* corresponding author

ARTICLE INFO

ABSTRACT

To support successful language teaching and learning, the teaching materials should represent actual language use. However, some studies have discovered significant mismatches between the language presented in textbooks and the real use of English. This study aims to find out the conversational features (i.e., inserts) and investigate the most frequent features in English textbooks for junior high schools in Indonesia. The data of this study are conversations containing inserts in the textbooks. The data were collected from three *Bahasa Inggris* textbooks for junior high schools compiled as a corpus of textbook conversations with 10.275 tokens. The inserts were identified from the corpus with the assistance of a corpus tool, LancsBox, using KWIC feature. The data were analyzed based on Biber's classification of inserts. However, this study limits the scope in the four most common inserts in conversation, such as interjections, discourse markers, response forms, and hesitators. The results show that conversation texts in *Bahasa Inggris* textbooks for junior high school lack of interjections, less variation of discourse markers, and hesitators are absent. Given the significance of inserts to make the conversation more natural and realistic, these findings are troublesome. Pedagogically, these findings should be taken into account by Indonesian textbook writers and teachers to help the students use the language more naturally in their conversations.

1. Introduction

To support the successful English teaching and learning process, good teaching material is needed. McGrath (2013) stated that teaching materials are an essential component of language programs. One of the most commonly used teaching materials is a textbook. It brings easiness to the classroom due to its comprehensiveness, completeness, and adjustment based on the learning objectives of the syllabus (Nunan, 1991). The use of textbooks is essential in the teaching and learning process, especially in Indonesia where English stays as a foreign language. Most students have a limited amount of exposure to English outside the classroom and few opportunities to practice it in daily life. Therefore, teachers and textbook writers should pay great attention to the language input so that the textbooks can help students to communicate effectively in using the language.

As the main source of students' language input, textbooks must provide good models that are suitable for the teaching and learning process. Unfortunately, some researchers have discovered

significant mismatches between the language presented in textbooks and the real use of English (Nguyen & Ishitobi, 2012; Oktavianti & Fajria, 2021; Oktavianti & Sarage, 2021; Phoocharoensil, 2017). To put it in another way, the language used in the textbooks does not accurately reflect how English is actually used.

In the investigation, some researchers mainly rely on the comparison with corpus to identify the mismatches. Numerous corpus-based studies on textbooks have been conducted. For instance, Cheng and Warren (2007) examined the expression of monitoring and understanding in textbooks and compare it to a corpus of spoken English in Hong Kong. Cullen and Kuo (2007) investigate 24 general EFL textbooks at all levels (from beginner to advanced) published from 2000-2006. Their study highlighted that there are limited features of spoken grammar presented in textbooks. Choi and Chon (2012) analyzed the collocation in 10th-grade high school textbooks from 16 different publishers and compare it to a 2k graded reader corpus. Oktavianti et al. (2020) investigated Indonesian EFL textbooks in comparison with corpus-based English textbooks. Oktavianti & Prayogi (2020) examine the use of future tense markers in Indonesian EFL textbooks and comparing to one of the biggest English corpora, COCA. All of these studies show the mismatches between the language presented in textbooks and the language in the corpora.

In Indonesia, there are English textbooks published by the Ministry of Education and Culture, and they are widely used in many schools. The researchers want to know the types of conversational features used in the textbooks published by the government and to identify the most frequent features distributed in the textbooks.

1.1 Conversation

Conversation is the most common type of spoken text since we live our lives in conversation. Clift (2016) states that people show their existence as a human being through conversation and they spend much of their life through conversation as well. Conversation happens when two people or more communicate by talking to one another. Elizabeth (2003) argues that conversation is a real-time exchange of thoughts and feeling in an informal setting. It is marked by an exchange of comment, information, and feeling, thus each person has an equal opportunity to express themselves.

1.2 Non-clausal Unit: Inserts

Inserts are defined as a class of words, stand-alone words which are characterized in general by their inability to enter into syntactic relations with other structures (Biber et al., 2021). The four most common inserts will be described as follows:

1) Interjections

According to Biber et al., (2021), the term interjections are words that have an exclamatory function. They usually express the speaker's emotions. For example: *oh, ah, ooh, wow, oops, whoops ugh, ow, ouch, aargh,* and *urgh*.

2) Discourse markers

Discourse markers are inserts that tend to occur at the beginning of a turn or utterance. Biber et al. (2021) explain that they signal a transition in the evolving progress of the utterance and signal an interactive relationship between the speaker, hearer, and message. For example: *well, right,* and *now*.

3) Response forms

According to Biber et al. (2021), response forms are short and routinized responses to a previous statement by the interlocutor. For example: *yes, no, okay, yeah, sure, uh huh, mm* and *I see*.

4) Hesitators

Hesitators are pause fillers. Biber et al. (2021) explain that the main function of hesitators is to enable a speaker to pause in the middle of a message while signaling the wish to speak. For example: *uh/er* and *um/erm*.

1.3 Corpus Linguistics

Corpus linguistics is a method for studying language in use through corpus (plural: corpora) and it can be used to analyze any aspects of language (sentence–syntax, word–morphology, dialect–sociolinguistics, etc.). According to Crawford and Csomay (2015) corpus linguistics is a

way to understand linguistics analysis and language, how language is actually used in context and how it can vary from context to context.

A Corpus (pl. corpora) is a large collection of texts stored digitally that can be analyzed using corpus software/tool. In corpus linguistics, the texts are different from those in language pedagogy. The texts in corpus linguistics refers to any instance of language use. They may be written or spoken and may vary in length (Jones & Waller, 2015). The written text can be taken from scanned or downloaded books, newspapers, or magazines. Spoken text, on the other hand, contain transcripts of spoken language such as conversations, phone calls, radio broadcasts, or TV shows.

A corpus can be very useful in the language teaching and learning. It provides information about vocabulary, grammar, formality and informality, the differences between spoken and written language, how we perform such basic functions as requesting, greeting, and apologizing, how people open and close conversations, how we change the subject, how we interrupt one another, how we ask questions, and so on (McCarthy, 2004). Crawford and Csomay (2015) state that instead of using judgement and intuition, corpus uses language that represents the actual of language use. It means, the language in corpus is authentic.

1.4 Previous studies

There are some previous studies that can be used as references when dealing with the naturalness of spoken texts in textbooks.

Many textbooks use made-up conversations rather than authentic materials. IntamAana (2004) investigate the discourse features of seven dialogues published in coursebooks between 1981-1997 and compared them with authentic interactions. The result shows that the dialogues in the coursebooks are very different from their authentic equivalents across the discourse features. The dialogues in the coursebooks are generally smooth and problem free which rarely occurs in real life. Pauses and back-channels are nearly absent from the coursebooks and it obviously does not accurately reflect the nature of the natural conversation.

The characteristics of grammar in spoken language differ from grammar in written language. In their study, Cheng and Warren (2007) figured that conversations in school textbooks in Hong Kong are still influenced by written language features. Many examples of spoken language in some English textbooks for secondary schools were not an accurate reflection of real-world language use.

Another relevant study was carried out by Molavi et al. (2014). Molavi et al. (2014) examined the occurrence of lexical collocations in three series of general English textbooks. They analyze the face-to-face and telephone conversation scripts collected from textbooks. They employ Open American National Corpus (OANC) and AntConc 3.2.1 concordancer program to compare the lexical collocations from those three textbooks to the real use by native speakers. The result shows that the textbooks have limited frequent collocations and the choice of collocations in the textbooks did not have big refers to the actual use of language by native speakers.

Sahragard et al. (2014) investigated interjections' frequency of use in three different EFL course books. All of these course books represented the same level of proficiency, i.e., the intermediate level. The results reveal that the interjections are not evenly distributed among different units of each course books. Several units included more than 100 interjections while others included less than 5 interjections.

A study by Tai (2016) explored the discourse markers (DMs) used in junior high school English textbooks, listening workbooks, and the English test in the Comprehensive Assessment Program (CAP) in Taiwan. The results of this study indicate that the English textbooks have the highest occurrence of DM compared to the listening workbooks and English test of CAP. However, the distributions of DMs were limited to the initial position.

Peksoy and Harmaoğlu (2017) did a study in which they compare the similarities and differences between language learning course books used in high school in Turkey and the 10-million-wors spoken part of the British National Corpus (BNC). In terms of particular grammatical items and frequency of the collocations, it was found that the language learning course books show little resemblance to the actual language use.

Murahata (2018) analyzed both quantitatively and qualitatively how the interjections '*Oh'* and '*Ah'* were used in MEXT (Ministry of Education, Culture, Sports, Science and Technology) authorized English textbooks used in junior high schools in Japan. The main result of this study shows an extreme discrepancy, contrary to the prediction based on a general token tendency in naturalistic discourse, that interjection '*Ah'* was rarely used in the textbooks. it has extremely low token frequency.

Focusing on discourse marker, An (2018) investigated the functions and frequency of discourse marker '*well'* in all the dialogues of five high school practical English conversation textbooks in Korea. According to the finding of this study, discourse markers '*well'* was overwhelmingly frequent in the textbooks. However, there are several functions of well did not appear at all in practical English conversation textbooks.

In another context, in Indonesia, Oktavianti et al. (2020) examined the conversations in Indonesian curriculum-based EFL textbooks for senior high school and compare the textbooks with three corpus-based textbooks. This study proves that there are mismatches between the textbooks. The Indonesian EFL textbooks did not provide sufficient varieties of inserts and the use of hesitators in the textbooks is insignificant.

Similarly, Intamanjana and Sada (2021) also analyzed the conversation texts of an Indonesian EFL textbook. In their study, they examined the lexical collocation in a textbook for tenth-grade students. The result of their study shows that the textbook contains variety of different lexical collocations. However, some of those word combinations do not reflect the naturalness of the language because they are rarely used in real life.

Although there have been numerous studies conducted, the issue of conversational features in Indonesian textbooks has not been much explored. Therefore, the researcher will conduct this study to fill the gap in the current condition.

2. Method

This study is descriptive qualitative research. It focuses on the conversational features used in three English textbooks for junior high schools in Indonesia. The subject of this study is the conversations in three English textbooks for junior high schools in Indonesia. They are *Bahasa Inggris: When English Rings a Bell VII, Bahasa Inggris: When English Rings a Bell VIII, and Bahasa Inggris: Think Globally Act Locally IX.* These textbooks were selected because the researchers want to know whether the textbooks published by the government guarantee that the books are suitable for the teaching and learning process and represent real language use. The data were collected in the following steps: (1) The researchers determined the keywords, (2) The researcher compiled the conversation texts from three English textbooks and make a corpus of textbook conversations with a total size of 10.275 words, (3) The corpus is loaded into LancsBox (Brezina et al., 2020) in the form of zip file, (4) The researchers use the KWIC tool and input the keywords in the search column to find the conversational features distributed in the textbooks, and (5) The result will be shown directly after the researchers click "search" as shown in Figure 1.

Fig. 1. Keyword search in LancsBox

In order to ensure the trustworthiness and minimize bias, this study employed inter-rater agreement. The use of inter-rater agreement to check the validity of a research is common practice especially in corpus-based research (Fajri et al., 2020; Nasrabady et al., 2020; Oktavianti & Prayogi, 2022; Phoocharoensil, 2017; Thao & Khoi, 2022). After collecting the data, the researcher consulted with two experts. In this study, the experts are lecturers in English Education Study Program of Universitas Ahmad Dahlan

The data were analyzed in the following steps. First, the researchers classify the type of inserts based on Biber et al. (2021) because that is the current and complete one. Second, the researchers describe the type of inserts. Third, the researchers calculate the frequency. Fourth, the researchers draw conclusions based on the findings.

3. Findings and Discussion

3.1. Findings

1) Types of conversational features in Bahasa Inggris textbooks for junior high schools in Indonesia

    a) Interjections

The first insert studied is interjection. The usage of interjections is insignificant in all three Bahasa Inggris textbooks for junior high schools in Indonesia. There are only three forms of interjections found in the textbooks. They are *oh*, *ah*, and *wow*.

According to Biber et al. (2021) *Oh* is by far the most common interjection. However, interjection *oh* is only found in *Bahasa Inggris* VII and *Bahasa Inggris* VIII.

(1)

Son:	And what about that building in front of us?
Father:	That is a hospital.
Son:	*Oh* yeah. That building is a hospital?

                                   (Bahasa Inggris: When English Rings a Bell VII, page 67)

(2)

Beni:	Hello, good morning. This is Beni.
Lina:	*Oh*, hi Beni. Good morning. How are you?
Beni:	Fine. I just got your invitation card to your birthday party.

                                  (Bahasa Inggris: When English Rings a Bell VIII, page 47)

Interjection *oh* also frequently combines with other inserts. One of the most common combinations is *oh yeah*. Excerpt (1) shows the example of that combination.

The next two interjections are *ah* and *wow*. Interjection *wow* can be found in all three Bahasa Inggris textbook, meanwhile *ah* is found only in Bahasa Inggris VII and absent in Bahasa Inggris VIII and Bahasa Inggris IX.

(3)

Lina:	But, all the ladies here are wearing batik and black pants.
Dayu:	You are right. *Ah*, there she is. She's the big lady with glasses. She's sitting on the bench in front of the principal's office.

                                  (Bahasa Inggris: When English Rings a Bell VII, page 158)

(4)

Bill:	This awesome gift was given to me by my friend.
Harry:	*Wow*, that looks nice! What is it called? Is that made in Indonesia?

                                  (Bahasa Inggris: Think Globally Act Locally IX, page 149)

From the excerpts above, it can be seen that *oh*, *ah*, and *wow* used in the textbooks already fulfill the exclamatory function of interjection.

b) Discourse markers

The second insert discussed is discourse markers. there are various types of markers used in the conversation section in the textbooks, such as *well, right,* and *now*.

(5)

Dayu: Are you mopping the floor again? You've just finished mopping it, haven't you?

Siti: *Well,* I tripped over the stool and feel and spilled my milk all over the floor.

(Bahasa Inggris: Think Globally Act Locally IX, page 84)

(6)

Siti: We will use English in our English class.

Teacher: *Right*. From now on we will use English in the English class. Everybody, what do you think? Are you ready?

(Bahasa Inggris: When English Rings a Bell VIII, page 5)

(7)

Miss Mutia: That is excellent! How about you, Edo? What do you do?

Edo: I'm a student.

Miss Mutia: That's good. *Now*, do you want to know what your other friends' parents do?

(Bahasa Inggris: When English Rings a Bell VII, page 128)

Excerpts (5)–(7) show various discourse markers used in the textbooks to maintain the flow of conversation. *Well* in excerpt (5) indicate the turn-taking from Dayu to Siti. In excerpt (6), the sentence-initial *right* indicate agreement. And, in excerpt (7), *now* marks a return to a related subject.

Although these textbooks have used a variety of discourse markers, some discourse markers like *you know* and *I mean* only found in Bahasa Inggris: When English Rings a Bell VII.

(8)

Beni: Siti, your house is very clean and tidy.

Siti: It's nice to have a clean and tidy house. It can also keep cockroaches and mice away. They don't like clean and tidy places, *you know*.

(Bahasa Inggris: When English Rings a Bell VII, page 120)

(9)

Edo: Have you seen real monkeys, Udin? *I mean*, not on TV or in a magazine.

Udin: Yes, I have.

(Bahasa Inggris: When English Rings a Bell VII, page 121)

Discourse marker *You know* used in final position in excerpt (8) signals that what follows is an explanation of what has preceded. while *I mean* used in medial position in excerpt (9) signals that a clarification is going to follow.

c) Response forms

Compare to other inserts, *Bahasa Inggris* textbooks unsurprisingly show more varied response forms in conversations (i.e., yes, yeah, no, okay, and sure). This is due to the nature of conversation as interactive communication. The examples of conversation in the textbooks are presented below.

(10)

Edo: And I like the color.

Udin: *Yeah*, I like dark brown, too.

<div align="right">(Bahasa Inggris: When English Rings a Bell VII, page 107)</div>

(11)

Teacher: Hello, excuse me. Listen, everybody. I think you need to read Chapter I, too.

Students: *Yes*, Sir.

<div align="right">(Bahasa Inggris: When English Rings a Bell VIII, page 11)</div>

(12)

Siti: Beni, did you go to Dayu's mother to learn to make stuff ed tofu?

Beni: *No*, she was in Posyandu the whole day yesterday.

<div align="right">(Bahasa Inggris: Think Globally Act Locally IX, page 92)</div>

(13)

Edo: Udin, let's go out. Don't be lazy. Get up and have some exercise outside.

Udin: *Okay*. Wait. Let me put on my shoes.

<div align="right">(Bahasa Inggris: When English Rings a Bell VIII, page 53)</div>

(14)

Lina: Yes, it's my dream. And I have practiced every day for the last two months with my dad.

Siti: Sure, you have to work hard for your dreams. Good luck!

<div align="right">(Bahasa Inggris: Think Globally Act Locally IX, page 4)</div>

Excerpts (10)-(14) show various response forms used in the conversation from Bahasa Inggris VII, VIII, and IX. It is noted that the use of response forms is important because of their role in signaling feedback to the speaker that their utterance is being understood or accepted.

Besides those five forms, there is one response form (i.e., *I see*) found in Bahasa Inggris VII. Response form *I see* is stronger in indicating a high degree of interest in the speaker's utterance. The example of this form is shown in the excerpt below.

(15)

Father: That is a school.

Son: *I see*. So, the school is next to the tax office.

<div align="right">(Bahasa Inggris: When English Rings a Bell VII, page 67)</div>

d) Hesitators

The last insert discussed is hesitator. In contrast to other inserts, hesitator has been neglected by textbook writer. All three Bahasa Inggris textbooks do not exemplify the use of hesitators in the conversations at all. This makes the conversation is unnatural and do not reflect the nature of conversation.

Conversation is typically spontaneous. It takes place in real time and involves participants producing utterances with limited time. As a result, there are time in conversation when a speaker's flow is disrupted by pauses at point where they need to keep talking (Biber et all., 2021). The absence of hesitators in these textbooks is troublesome because hesitators like *uh* and *um* are included in top 50 words in conversation (McCarthy, 2004).

2) The most frequent conversational features in Bahasa Inggris textbooks for junior high schools in Indonesia

The table below shows the detail number of the most frequent inserts found in Bahasa Inggris textbooks for junior high schools in Indonesia based on Biber's theory.

Table 1.   Frequency of inserts in Bahasa Inggris Textbooks

No	Category	Inserts Type	Token Frequency	Percentage
1	Interjections	Wow	11	8,9%
		Oh	4	
		Ah	1	
		Ooh	0	
		Oops	0	
		Whoops	0	
		Ugh	0	
		Ow	0	
		Ouch	0	
		Aargh	0	
		Urgh	0	
	Total		**16**	
2	Discourse Markers	Right	19	17,8%
		Now	6	
		Well	3	
		You know	3	
		I mean	1	
	Total		**32**	
3	Response Forms	Yes	77	73,3%
		Okay	20	
		Sure	16	
		No	12	
		I see	4	
		Yeah	3	
		Mm	0	
		Uh huh	0	
	Total		**132**	
4	Hesitators	Uh	0	0%
		Er	0	
		Um	0	
		Erm	0	
	Total		**0**	
	TOTAL INSERTS		**180**	100%

As shown in Table 1, the researchers found some inserts used in three Bahasa Inggris textbooks for junior high schools in Indonesia. The most frequently used insert is response forms which occurs 132 times. then it is followed by discourse markers with 32 occurrences, and interjections with 16 occurrences. Meanwhile, the least frequency of insert is hesitators which is absent in the conversation in all three Bahasa Inggris textbooks.

Given the nature of conversation as interactive communication, it is plausible to find lots of examples of response forms in the textbooks. These features are pretty basic in conversation to fulfill the need to response to the speaker's utterance (i.e., questions, statements).

3.2. Discussion

Based on the findings above, the researchers discovered three out of four categories of the most commonly used inserts in the conversations in Bahasa Inggris textbooks for junior high schools in Indonesia. They are interjections (oh, ah, wow), discourse markers (well, right, now, you know, I mean), and response forms (yes, yeah, no, okay, sure, I see).

The researchers compare the finding of this study to previous studies including Gilmore (2004), Murahata (2018), Oktavianti et al. (2020), Sahragard et al. (2014) and Tai (2016). The result of this study had some similarities and differences to the other previous studies in the field of conversational features.

This study shows that response forms have the highest number of occurrences in conversations in the textbooks with a total percentage of 73,3%. Meanwhile, all forms of hesitators are absent from three Bahasa Inggris textbooks making hesitators become the least frequent inserts used in the textbooks. The results were in line with the research conducted by Oktavianti et al. (2020), who claim that response forms are copious in Indonesian EFL textbooks for senior high school compared

to the other inserts. Oktavianti et al. (2020) also found that the use of hesitators in Indonesian EFL textbooks is nearly absent. The finding also agreed with Gilmore (2004), who pointed out that the number of instances of hesitators in textbook dialogues is still well below those expected.

Moving on to discourse markers, the three Bahasa Inggris textbooks basically use the same markers. However, the discourse markers were not evenly distributed among these three textbooks. Bahasa Inggris VII included all discourse markers under study while other two textbooks included less variation as there are only two markers found in Bahasa Inggris VIII (i.e., right and now) and three markers in Bahasa Inggris IX (i.e., well, right and I mean). This finding was, in part, contradictory to Tai (2016), who claim that the curriculum-based English textbooks in Taiwan had used various discourse markers with the highest frequency compared to the two corpora under study. The finding of this study also contrasts to An (2018). An (2018) reveals that Korean English conversation textbooks included a lot of examples of discourse marker 'well'.

In terms of interjections, Bahasa Inggris VII, VIII, and IX show limited variation in the conversations as there are only three interjections found (i.e., wow, oh, ah). The textbooks also do not exemplify the use of interjections in conversations that much. As shown in Table 1, interjections only occur 16 times with a total percentage of 8,9%. This small number of occurrences is troublesome given that interjection oh is listed as top 50 words in conversation based on the data from Cambridge International Corpus (McCarthy, 2004). This result was consistent with the findings by Sahragard et al. (2014) and Murahata (2018), both of who pointed out that interjection 'Ah' was rarely used in the textbooks. Sahragard et al. (2014) even claim that interjection 'Ah' repeated only once in the textbooks.

4. Conclusion

This study examines the conversational features (i.e., inserts) in three Bahasa Inggris textbooks for junior high schools in Indonesia. The findings discussed in this study present some significant points to take into account. It is evident that all three textbooks show the mismatches between the language presented in the conversation of textbooks and the actual use of English. In regard to type of inserts used in conversation in the textbooks, there are three out of four most common insert proposed by Biber were identified, such as interjections, discourse markers, and response forms. However, some inserts are not well-presented since the occurrences tend to be monotonous. All three Bahasa Inggris textbooks show limited use of interjections and less various example of discourse markers. What is more surprising is that the conversation in the three textbooks do not even apply hesitators which seems unnatural due to the need for hesitators to keep the flow of conversation. In terms of frequency, response forms (73,3%) take the highest number of occurrences in conversation in the textbooks. Then, it is followed by discourse markers (17,8%) and interjections (8,9%). Meanwhile, hesitators take the lowest number of occurrences because it is absent from the textbooks.

Nevertheless, this study has several weaknesses. First, the corpus of textbook conversations in this study only consists of 10.275 tokens which is still relatively small for corpus data. Second, this study also does not use any general reference corpus or corpus-based textbooks as a comparison. Third, in terms of research object, this study merely examines four common inserts. Hence, further research can be continued by analyzing more inserts and making a comprehensive comparison of inserts by reflecting the use of language in the textbooks to a corpus.

Despite the weaknesses, the findings of this study have potential implications for textbook writer and teacher. Teacher is suggested to not rely solely on textbooks. Teacher can enrich the teaching materials by using a corpus as a valid data source in providing examples of conversations that represent the real use of English. As for textbook writer, it is suggested to be more aware of the importance of conversational features (i.e., inserts) in designing the conversation in the textbooks. The writer is expected to include various inserts in the textbooks to make the conversation more natural and realistic. Furthermore, the writer can consider the use of corpus in selecting the most frequently used inserts in the conversation text.

REFERENCES

An, J. (2018). Analysis of the discourse marker well in high school practical English conversation textbooks. *Modern English Education, 19(1)*, 16-28. https://doi.org/10.18095/meeso.2018.19.1.02

Biber, D., Johansson, S., Leech, G. N., Conrad, S., & Finegan, E. (2021). Grammar of spoken and written English. John Bejamins.

Brezina, V., Weill-Tessier, P., & McEnery, A. (2020). #LancsBox v. 5.x. [software]. Available at: http://corpora.lancs.ac.uk/lancsbox.

Cheng, W., & Warren, M. (2007). Checking understandings: Comparing textbooks and a corpus of spoken english in Hong Kong. Language Awareness, 16(3), 190–207. https://doi.org/10.2167/la455.0

Choi, H.-Y., & Chon, Y. v. (2012). A Corpus-based Analysis of Collocations in Tenth Grade High School English Textbooks. Multimedia-AssistedLanguageLearning, 15(2), 41–73.

Clift, R. (2016). *Conversation analysis*. Cambridge University Press.

Crawford, W. J., & Csomay, E. (2015). Doing corpus linguistics. Routledge.

Cullen, R., & Kuo, I. C. (2007). Spoken grammar and ELT course materials: A missing link? TESOL Quarterly, 41(2), 361–386. https://doi.org/10.1002/j.1545-7249.2007.tb00063.x

Elizabeth, M. (2003). *Painless speaking*. Barrons Educational Serries, Inc.

Fajri, M. S. al, 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.2020.11271

Gilmore, A. (2004). A comparison of textbook and authentic interactions. ELT Journal, 58(4), 363–374. https://doi.org/10.1093/elt/58.4.363

IntamAana, P. D., & Sada, C. (2021). A corpus-based analysis of lexical collocation in conversation texts of English textbooks. Journal of English as a Foreign Language Education (JEFLE), 2(2), 68–75.

Jones, C., & Waller, D. (2015). Corpus linguistics for grammar: A guide for research. Routledge.

McCarthy, M. (2004). Touchstone: From corpus to course book. Cambridge University Press.

McGrath, I. (2013). Teaching materials and the roles of EFL/ESL teachers: Practice and theory. Bloomsbury.

Molavi, A., Koosha, M., & Hosseini, H. (2014). A comparative corpus-based analysis of lexical collocations used in EFL textbooks. Latin American Journal of Content and Language Integrated Learning, 7(1), 66–81. https://doi.org/10.5294/laclil.2014.7.1.4

Murahata, G. (2018). A Critical analysis of the interjections 'Oh'and 'Ah' as pragmatic markers in MEXT authorized English textbooks. *ARELE: Annual Review of English Language Education in Japan*, *29*, 49-64.

Nasrabady, P., Shirvan, M. E., & Golparvar, S. E. (2020). Exploring lexical bundles in recent published papers in the field of applied linguistics. *Journal of World Languages*, *6*(3), 175–197. https://doi.org/10.1080/21698252.2020.1797992

Nguyen, H. T., & Ishitobi, N. (2012). Ordering Fast Food: Service Encounters in Real-Life Interaction and in Textbook Dialogs. JALT Journal, 34(2), 154–185.

Nunan, D. (1991). Language teaching methodology: A textbook for teachers. Prentice hall.

Oktavianti, I. N., & Fajria, A. (2021). Modal Verbs in a Curriculum-Based EFL Textbook of Senior High School in Indonesia: A Corpus-Based Study. TESOL International Journal, 16(5), 171–204.

Oktavianti, I. N., & Prayogi, I. (2020). A Corpus-Based Analysis of Future Tense. In Markers Indonesian Journal of EFL and Linguistics (Vol. 5, Issue 2). www.indonesian-efl-journal.org

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

Oktavianti, I. N., Prayogi, I., Amal, M. A. B., & Pertiwi, R. S. (2020). An Analysis of Conversations in Curriculum-based EFL Textbooks for Senior High School in Indonesia and the Comparison with Corpus-based English Textbooks. Universal Journal of Educational Research, 8(9), 4151–4162. https://doi.org/10.13189/ujer.2020.080941

Oktavianti, I. N., & Sarage, J. (2021). Collocates of 'great' and 'good' in the corpus of contemporary American English and Indonesian efl textbooks. Studies in English Language and Education, 8(2), 457–478. https://doi.org/10.24815/siele.v8i2.18594

Peksoy, E., & Harmaoglu, O. (2017). Corpus based authenicity analysis of language teaching course books. International Journal of Languages' Education and Teaching, 5(4), 287–307. https://doi.org/10.18298/ijlet.2324

Phoocharoensil, S. (2017). Corpus-based exploration of linking adverbials of result: Discovering what ELT writing coursebooks lack. In 3L: Language, Linguistics, Literature (Vol. 23, Issue 1, pp. 150–167). Penerbit Universiti Kebangsaan Malaysia. https://doi.org/10.17576/3L-2017-2301-11

Sahragard, R., Mavaddat, R., Bayani, S. R., & Safavian, M. (2014). Investigating the implementation of interjections in three current EFL course books. *Procedia-Social and Behavioral Sciences*, *98*, 1621-1630. https://doi.org/10.1016/j.sbspro.2014.03.586

Tai, T. Y. (2016). A corpus-based analysis of discourse markers in curriculum-based english textbooks and the english entrance exam in Taiwan. The Journal of Asia TEFL, 13(4), 262–279. https://doi.org/10.18823/asiatefl.2016.13.4.2.262

Thao, T. Q., & Khoi, N. M. (2022). The deployment of English lexical bundles in applied linguistics research articles by Vietnamese researchers. *HCMCOUJS-Social Sciences*, *12*(2), 75–84. https://doi.org/10.46223/HCMCOUJS