A Corpus-Based Analysis of Future Tense Markers in Indonesian EFL Textbooks for Senior High School

By Ikmi Nur Oktavianti

A Corpus-Based Analysis of Future Tense Markers in Indonesian EFL Textbooks for Senior High School

Ikmi Nur Oktavianti
Universitas Ahmad Dahlan
e-mail: ikmi.oktavianti@pbi.uad.ac.id

Icuk Prayogi
Universitas PGRI Semarang
e-mail: icukprayogi@gmail.com

Abstract:
A well-designed textbook can enhance the teaching and learning of a foreign language and it should also consider the authenticity aspect of the target language. This paper aims at studying the use of future tense markers (i.e. will, be going to) in three EFL textbooks for senior high school in Indonesia and comparing to one of the biggest English corpora, Corpus of Contemporary American English, to investigate the authenticity of the textbooks. This study employs corpus-based analysis as it focuses on the frequency and collocates of future tense markers. This study, however, delimits the scope in conversations of the textbooks and the spoken sub-corpus. It shows that, although 'will' and 'be going to' are used in the textbooks and the corpus, there are some distinctions observed, be they the frequency of use or the collocates preceding the future tense markers. As for the semantic of the verb collocates, there seem to have many differences. These differences then need to be revisited and re-evaluated to improve Indonesian EFL textbooks materials to equip the English learners in a foreign language context with actual language use.
Keywords: *authenticity, corpus linguistics, future tense, textbooks*

1. INTRODUCTION

In teaching and learning, textbooks hold crucial role as they provide learning sources beneficial and useful to the learners, especially in teaching English as a foreign language (EFL) (Charalambous, 2011; Radić-Bojanić & Topalov, 2016). Römer (2005) states that a well-designed EFL textbook will enhance the teaching and learning process. Hence several studies on textbooks, in this case is Indonesian EFL textbooks, have been conducted, focusing on cultural aspects (Mayangsari et al., 2018; Rahmah et al., 2018; Stockton, 2018), evaluation of tasks (e.g. Ayu & Indrawati, 2018) and values depicted in the textbooks (e.g. Widodo, 2018).

In learning a language, the materials in the textbooks should consider the authenticity and richness of the target language. The textbook, thus, has to be in line with the nature of the learning materials. Tomlinson (1998) states that language materials in a textbook should be rich and authentic. The term 'authenticity' itself, however, remains debatable. Widdowson (1998) argues that learners are the outsiders of the communities using the language 12 it is unnecessary to use authentic language in learning process. On the other hand, McEnery and Wilson (2001) argue that students who learn English using traditional textbooks (with artificial language use examples) often find difficulties in analyzing longer sentences in real life. Intriguingly, Gilmore (2004) proves that textbook dialogues differ from the equivalent language use in real life in terms of the length, turn-taking pattern, lexical density, false start, repetition, and so on. Cheng and Warren (2007) examine the expression of understanding between speakers in textbook conversation and compare it to *Hongkong Corpus of Spoken English* and find out that the language used in conversation in EFL textbooks is rarely found in daily use of English in Hongkong. It seems that the language in the textbook is strongly influenced by the characteristics of an academic context. It shows that although the term 'authenticity' is debatable, yet the findings prove it to be one of the important factors in designing and developing teaching materials, including textbooks.

Concerning the prominence of textbooks, Collins (2006) says that textbooks do not always provide accurate information about English use, and this is likely harm the accuracy of both the teachers and the students' knowledge of English structure and use. In relation to EFL textbooks in Indonesia, Collins (2006) identifies some shortcomings, including accuracy, factual errors, prescriptive bias, unnatural English, and so on. Given the findings, it is known that English textbooks analyzed by Collins do not correspond to the 'real' English (i.e. English language used in real communication), thus it is not an authentic English.

One of the linguistic features to be analyzed is the future tense marker. As one of the Indo-European languages, English is categorized into language that distinguishes past and non-past tense in relation to time (Comrie, 2000). Therefore, to express future time, there are *will, be going to, shall, can* in which they basically and naturally belong to modal verbs. Unlike future tense, English just attaches inflections to the end of the

212

verb to indicate past tense. It is interesting then to study the use of future tense markers. In addition, future tense markers are prominent due to the expression of modality. According to some corpora, such as *Corpus of Contemporary American English* (COCA), *British National Corpus* (BNC), *A Representative Corpus of Historical English Registers* (ARCHER), the expression of prediction or future-related using the modal verb *will* is the most frequent one. Below is the normalized frequency (*nf*) per million words of *will* in those corpora.

Table 1: The most frequently used modal verbs in corpora

ARCHER	BNC	COCA
will (2,486)	would (2,421)	would (2,303)
would (1,824)	will (2,398)	can (2,174)
may (1,538)	can (2,284)	will (1,979)

Thus, it is obvious to say that the modal verb *will* expressing future/prediction is an ultimately salient unit in English grammar due to its high frequency of use. The teaching of future tense markers—to cover future tense (tense) and prediction (modality)—becomes significant and the assistance of textbooks discussing the relevant materials properly is unavoidable. As studied by Ojanen (2008), the use of future tense markers in two English textbooks and the comparison with BNC and ICLE prove that there is a slight difference in the use of the markers in the textbooks and the corpora as the representative of 'real' English. In BNC, for instance, *be going to* is frequently used, but it is rarely found in textbooks. Moreover, following Khadim (2015), teaching future time to EFL learners is challenging as English and the L1 have different ways of perceiving time. Thus, the study of future tense markers used in textbooks is compelling.

This paper studies the use of future tense markers in Indonesia EFL Textbooks for senior high schools (grade 10, 11, and 12). This research, however, focuses on the spoken part of the textbooks or in this case is the dialogues or conversations presented in the textbooks due to the high requirement of speaking to be as natural as possible and as authentic as possible. In daily language use, the spoken variety is of importance because it is the primary means of communication. The teaching of English to the extent of conversation should be as realistic as possible, portraying the actual language use. Hence, to be able to describe the authenticity of the language, this study compares the use of future tense markers *will* and *be going to* with one of the biggest and most recent English corpora, COCA (Davies, 2008). Language use compiled in COCA is expected to represent 'real' English usage due to its huge size (more than 1 billion words), long period of compilation (from 1990 to 2019), and various sources (e.g., fiction, academic, news, etc.).

Previous studies examining linguistic feature in textbooks and the actual use have been conducted, e.g., by comparing modal verbs used in EFL textbooks with the ones produced by EFL teachers (Al-Jaboori, 2008). Relevant to this study, there are several previous studies on the teaching of future tense and/or other English tense systems (Handayani et al., 2013; Khadim, 2015) and the investigation of future tense and/or other tense systems in textbooks (Pounds, 2011). Other studies addressing the corpus-based analysis of linguistic features have been carried out, particularly investigating textbook in comparison with a certain and a self-compiled corpus (Arellano, 2018; Hsieh et al., 2011; Zambrana, 2017), analyzing linguistic features of textbook compared to the online general reference corpora (Cheng & Warren, 2007; Choi & Chon, 2012; Khijasteh & Kafipour, 2012; Orlando, 2009), using the corpus-based to describe lexis of the textbooks (Nordberg & Nordlund, 2018), using the corpus-based analysis to investigate the use of modality in EFL textbooks (Nordberg, 2010). Nonetheless, little is known about the study of Indonesian EFL textbooks using corpus-based analysis. Most researchers analyzing Indonesian EFL textbooks still focused on cultural values (e.g., (Mayangsari et al., 2018; Rahmah et al., 2018; Stockton, 2018). This paper, thus, describes the future tense markers used in Indonesian EFL textbooks and the comparison with COCA as the representative of 'real' English, regarding the frequency of use and the collocates, to identify the authenticity of the textbooks.

2. CORPUS AND LANGUAGE TEACHING

Corpus was first invented to assist language research, studying the structure and use of linguistic units who formerly relied on the intuition of the researchers. Due to the advancement of science, intuition is no longer relevant to some extent and linguistic research seeks empirical evidence and one of them is from a corpus. McEnery and Xiao (2013) state that corpus, the digital storage of language data, is no longer exclusive for linguistics and it has gained its position among teachers and ELT researchers. They make use of corpora to help enhance the teaching learning process, directly and indirectly, including the design of the materials. Mindt (1996) states that teaching syllabuses should be based on empirical evidence rather than intuition and tradition. The frequency of use, thus, can be one of the guides to prioritize the contents of teaching materials. According to Mindt (1996), the information about frequency can assist to make the learning process more effective. Moreover, Römer (2004) claims that frequencies are the keys to show us words or structures that are central in a certain language. By knowing which words or structures that are vital, it is unexacting to decide what should be included in teaching materials and which one should be prioritized or emphasized due to its higher occurrence in 'real' English. The role of corpora in language teaching thus is unavoidable and crucial. Baker (2010) says that frequency is important because it shows markedness or what's important among many other things or concepts.

As a concrete example of using corpus in designing materials is the books written by

McCarthy et al. (2014) known as *Touchstone* book series, which is based on the *Cambridge International Corpus*. This corpus-based book aims to present the vocabulary, grammar and functions students encounter most often in real life. Another example is *In Focus* series (Browne et al., 2013) published by Cambridge University Press is a recent corpus-based English textbook, designed by using *Cambridge Corpus*. Therefore, a corpus-based textbook is necessarily salient to provide concrete language use to the learners. Besides, Römer (in Sinclair, 2004) claims that empirical evidence served from corpus can contribute to improve teaching materials and it is important to pay more attention to linguistic units and patterns that are typically used in the language in pedagogical context.

3. RESEARCH METHODOLOGY

The data can be classified into qualitative and quantitative data. The qualitative data is the utterances containing modal verbs *will* and *be going to* expressing future tense. As for the quantitative data, it is related to the frequency of use (raw frequency or *rf*) of *will* and *be going to*. The data were compiled from the conversation sections in three EFL textbooks for senior high school in Indonesia; they are *Bahasa Inggris Kelas X, Bahasa Inggris Kelas XI,* and *Bahasa Inggris Kelas XII* (henceforth BI X, BI XI, BI XII respectively). As a comparison, this study also collects the data from spoken sub-corpus of COCA so that it is comparable with the textbook data. In addition, to complete the analysis, two other corpora were used, namely BNC and ARCHER.

The qualitative data were analyzed by identifying and classifying the collocates of *will* and *be going to*. The right collocates or the verb collocates of *will* and *be going to* were analyzed semantically by classifying the verb collocates using Halliday's verb classification combined with Dixon's verb classification as described in Scheibman (Scheibman, 2001).

Table 2: Verb types (from Scheibman in Bybee and Hopper, 2001: 67)

Verb Type	Description
Cognition	cognitive activity
Corporeal	bodily gesture, bodily interaction
Existential	exist, happen
Feeling	emotion, wanting
Material	concrete and abstract doings and happenings
Perception	perception, attention
Perception-relational	perception (subject not senser)
Possessive	possession
Relational	process of being
Verbal	saying, symbolic exchange of meaning

Meanwhile, the quantitative data, raw frequency (*rf*) of the corpus, is normalized into normalized frequency (*nf*). Below is the formula for the normalization (Brezina, 2018; McEnery & Hardie, 2012) with the base of normalization is per one million words.

Below is the formula for the normalization.

$$nf = \frac{tokens}{size\ of\ copus} \times base\ of\ normalization$$

Both qualitative and quantitative data are prominent to solve the problems in investigating the use of future tense markers in EFL Textbooks.

4. FINDINGS

The analysis of the data concerns three main points, namely frequency of use, collocates, and the semantic of collocates. The following subsections are the elaboration for each part.

4.1 Frequency of Use

Observing the three textbooks, there found the use of future tense *will* and *be going to*. The frequency of use of *will* and *be going to* in BI X, XI, and XII compare to COCA can be seen in the following table. The frequency for the textbooks is displayed in *rf* and COCA in *nf*.

Table 3: Frequency of *will* and *be going to*

Future tense markers	Textbooks			COCA (*nf*)
	BI X (*rf*)	**BI XI** (*rf*)	**BI XII** (*rf*)	
will	11	8	10	2,211
be going to	9	0	1	1,903

As seen in the table, the frequency of use of *will* is higher than *be going to* among all textbooks (BI X, BI XI, and BI XII). Most of the future tense expressions are marked by using *will* instead of the periphrastic *be going to*. This corresponds to the frequency of *will* and *be going to* in spoken sub-corpus of COCA which reflects the same result. The raw and normalized frequencies can be displayed in percentage as in figure 1.
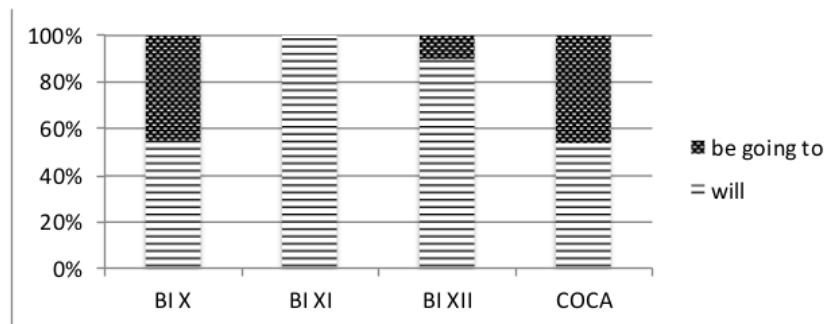

Figure 1: Usage percentage of *will* and *be going to*

Interestingly, in BI XI, there is no modal verb *be going to* used in the spoken section, while 100% of all future tense expressions are marked through *will*. Meanwhile, BI X shows 55 % for *will* and 45% for *be going to*; BI XII exhibits 90% use of *will* and 10% use of *be going to* in future tense expressions. As for COCA, the use of *will* is 54% and *be going to* is 46%. Of all three textbooks, it is obvious that the use of future tense markers in BI X is the closest to their use in COCA.

Concerning the findings of frequency above, it can be seen that two out of three textbooks analyzed do not provide the accurate information of the frequency of *will* and *be going to* since they use lesser number of *be going to* or do not use *be going to* at all in the conversations of the textbooks. While in the spoken sub-corpus of COCA, *be going to* occupies the third-highest modal verb, but not in the written sub-corpus (Oktavianti, 2019). Comparing spoken and written variety of English, the use of *will* and *be going to* is higher in spoken than in written one, as displayed in table 4.

Table 4: Frequency of *will* and *be going to* in spoken and written English (COCA)

Future tense markers	Spoken (*nf*)	Written (*nf*)
will	2,211	1,746
be going to	1,903	267

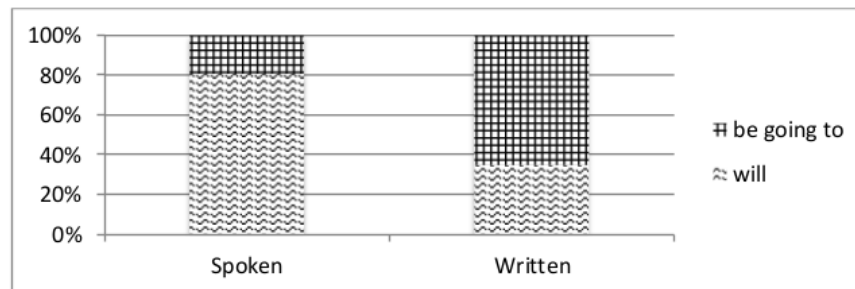Table 4 can be illustrated into percentages as in figure 2.

Figure 2: Percentage of use of *will* and *be going to* in spoken vs. written English (COCA)

Figure 2 shows the different characteristics of spoken and written English in relation to the use of *will* and *be going to*. This empirical data corresponds to Leech et al. (2009) explaining that the use of semi-modal (*be going to* belong to semi-modal) might indicate colloquialization or the shift of language use into more speech-like. Therefore, seeing the frequency of *will* and *be going to* in the spoken section in BI XI and XII, it is not unusual to refer to the frequency of *will* and *be going to* in written English. Meanwhile, the use of future tense markers in the spoken section in BI X is in line with empirical data in the corpus. Nevertheless, the latter point might occur because there is

indeed a particular discussion about using *be going to* in BI X that we cannot find in BI XI and BI XII. It is no wonder if the use of *be going to* is higher in BI X.

To get a more comprehensive analysis of *will* and *be going to,* this study compiles the use from three different corpora ranging over periods of English. Taken from ARCHER (data from 1600—1900), BNC (data from 1980—1993), and COCA (data from 1990—2017) below is the raw illustration of usage frequency of *will* and *be going to* overtime (from Early Modern English to Present-day English).
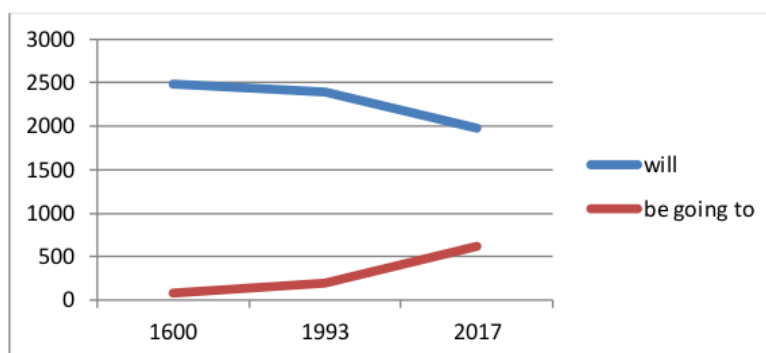

Figure 3: The use of *will* and *be going to* overtime (ARCHER, BNC, COCA)

From figure 3 we can notice that the use of *will* is insignificantly declining, although it is still more frequent compared to other linguistic units of the same category. As for *be going to,* its use is inclining gradually overtime. In short, figure 3 presents the contrasting usage trend of *will* and *be going to,* yet in general, *will* is still more frequent than *be going to.* Even though the frequency of *be going to* is still under the frequency of *will,* but it is plausible that *be going to* can outnumber the usage frequency of *will* in a certain context of language use. Hence the use of *be going to* should reflect the actual trend of language use to equip the learners with what is happening in the language they are learning.

4.2 Collocates of Future Tense Markers

From the analysis of frequency, we can notice that there is a slight difference but crucial regarding the use of *will* and *be going to.* Frequency alone, however, is not enough to justify the authenticity of future tense marker use. This subsection analyzes the collocates of those two markers, concerning the first left collocate or L1 (i.e. units precede the markers) and first right collocate or R1 (i.e. units follow the markers). Below are the collocates of *will* and *be going to* in BI X (the numbers in the brackets present the raw frequencies).

Table 5: Collocates of *will* and *be going to* in BI X (*rf*)

Future tense markers	L1	R1
will	it (4)	be (3)
	I (3)	do, try, go, feel, forget, help, continue (1)
	you, we (1)	
be going to	you, we (3)	do (3)
	I (2)	go (2)
		take, practice, bake (1)

Table 5 shows that, in the left collocates of BI X, the future marker *will* mainly precedes the third-person singular *it* and first-person singular *I*, different from *be going to* that mainly precedes the second-person singular/plural *you* and the first-person plural *we*. In the right collocates, *will* is predominantly followed by *be* in which it is rather unusual for *be going to* since it already comprises *be* auxiliary. Along with *be* as the right collocate of *will*, there are some other lexical verbs, such as *do, try, go, feel, forget, help, continue* occur once in the spoken section of BI X. Meanwhile, *be going to* is mostly followed by *do* and *go*, along with some other less frequent verbs, such as *take, practice*, and *bake*.

Unlike BI X, there is no *be going to* found in BI XI, causing in no description of the collocates of the future tense marker. However, *will* is still used and thus it is preceded and followed by other linguistic units as described in the following table.

Table 6: Collocates of *will* and *be going to* in BI XI (*rf*)

Future tense markers	L1	R1
will	I (6)	be (7)
	there, you (1)	start (1)
be going to	-	-

Based on table 6, it is seen that the future marker *will* mainly precedes the first-person singular *I* and is followed by *be* auxiliary. Compared to BI X, the L1 and R1 of *will* in BI XI is slightly different from BI X; in BI XI, there is no third-person singular *it* found. Nonetheless, they share something in common in which *will* is followed predominantly by *be* auxiliary in BI X and BI XI. As for the collocates of *will* and *be going to* in BI XII is listed below.

Table 7: Collocates of *will* and *be going to* in BI XII (*rf*)

Future tense markers	L1	R1
will	I (5)	get (2)
	she, we, you, it (1)	be, edit, help, have, feel, do, take (1)
be going to	they (1)	hold (1)

Table 7 exhibits the L1 of *will* in BI XI is in accordance with L1 of *will* in BI XII with the first-person singular *I* being mostly found. Interestingly, it turns out that auxiliary *be* is not the most frequent collocate of *will* in BI XII. As for pronoun subject of *be going to,* the only pronoun found is *they* as well as the only verb follows (*hold*). In short, comparing the collocates of *will* and *be going to* in three EFL textbooks, it is found that there are some similarities and differences, including the same pronoun (*I*) as the L1 collocate of *will* and the same verb *be* as the R1 collocate of *will*. To classify, *will* is mainly preceded with the first-person singular and the third-person singular and followed by auxiliary *be*. For *be going to,* however, there is insufficient data to depict the collocates so the description will not do any justice.

To investigate the authenticity of future tense markers used in BI X, BI XI, and BI XII, it is necessary to contrast the results found with those in COCA. The following table displays the collocates of *will* and *be going to* in COCA—using *nf* per million words.

Table 8: Collocates of *will* and *be going to* in COCA (*nf*)

Future tense markers	L1		R1	
will	we (40)		be (109)	do (7)
	it (35)		have (19)	continue (7)
	I (27)		go (7)	come (6)
	they (24)		get (7)	see (6)
	you (20)		take (7)	say (6)
	L1		**R1**	
be going to	we (93)		take (0,08)	happen (0,04)
	I (52)		get (0,07)	make (0,04)
	you (52)		see (0,06)	come (0,04)
	it (42)		go (0,05)	find (0,03)
	they (38)		say (0,05)	try (0,02)

Table 8 presents the most frequent left collocate of *will* is *we,* followed by *it, I, they,* and *you* respectively. In other words, the dominant pronoun precedes *will* in COCA is third-person singular and plural (i.e. *we, it, they*) as well as the first-person singular (*I*). Furthermore, the right collocates of *will* is auxiliary *be,* followed by *have* and other lexical verbs, such as *go, get, take, do continue,* and so on. Overall, the R1 or the verb

collocates of *will* in COCA are dominated with auxiliary verb *be* (52%), while the lexical verbs achieve 48%; the verb collocates of *be going to* in COCA are all lexical verbs. Verb collocates of *will* in BI X, however, show distinct characteristics with auxiliary *be* occupy 30% and the other 70% is lexical verbs. The extreme result can be seen from the verb collocates of *will* in BI XI and in BI XII in the way it shows the use of auxiliary verb *be* for 87,5% and lexical verbs for 12,5%. On the contrary, verb collocates of *will* in BI XII show 11% of auxiliary verb *be* and 89% of lexical verbs. Nonetheless, verb collocates of *be going to* in BI XI and BI XII do not provide significant statistic.

Comparing the results from textbooks and COCA regarding the collocates of *will*, we notice that the L1 collocates of *will* in textbooks are not totally in line with those in COCA as the use of the pronoun *we* is less common in textbooks. On the other hand, the R1 collocates of *will* in textbooks do correspond to those in COCA since auxiliary *be* is the most frequent R1 collocate—and some other verbs are the same (e.g., *have, do, take, continue*).

In the meantime, the description of *be going to* is not quite comprehensive, but reflecting the results in textbooks, it is found that the most frequent use of the pronoun *we* in BI X is in accordance with those in COCA, but not in the other textbooks. Alike the left collocates, the right collocates of *be going to* in BI X also share the same verbs as in COCA (i.e. *take, go*). However, the other textbooks, BI XI and BI XII, do not share the same collocates.

4.3 Semantic of the Collocates

In addition to the description of collocates, it is necessary to describe the semantic classification of the verb collocates—as the right collocates—of both *will* and *be going to* in the textbooks and COCA. This is primarily to map the semantic aspect as well as to find out the similarities and differences of semantic aspects in textbooks and COCA. The top 20 verb collocates in COCA are classified and compared to those in textbooks. In COCA, the verb collocates of *will* can be classified into *relational, material, possession, verbal, perception*, and *existential*. Meanwhile, the verb collocates of *be going to* in COCA are classified into *material, possession, perception, existential*, and *verbal*.

Table 9: Semantic of verb collocates in COCA

will		*be going to*	
Relational (be)	53%	Material (take, go, make, come, find, try, let, put, kill, live, give, continue)	62,5%
Material (go, take, do,	25%	Possession (get)	12%

continue, come, make, help, give, find, work, bring, try)			
Possession *(have, get)*	12,5%	Perception *(see, look)*	12%
Verbal *(say, tell)*	11,5%	Existential *(happen)*	8,5%
Perception *(look, see)*	3%	Verbal *(say, ask)*	5%
Existential *(happen)*	1,5%		

Verb collocates of *will* in BI X denote *relational, material, feeling*, and *cognition* and the verb collocates of *be going to* denote *material* type only. In BI XI, verb collocates of *will* can be classified into *relational* and *material* and there is no verb collocate of *be going to* in this textbook due to its zero usage. BI XII is closer to BI X as its verbs collocate of *will* also denote *relational, material, possession*, and *feeling* (but there is no *cognition* type) and the verb collocate of *be going to* in BI XII belongs to the *material* type.

Table 10: Type of verb collocates of *will*

Rank	COCA	BI X	BI XI	BI XII
1	Relational (53%)	Material (50%)	Relational (87,5%)	Material (45%)
2	Material (25%)	Relational (30%)	Material (12,5%)	Possession (33%)
3	Possession (12,5%)	Feeling (10%)		Relational (11%)
4	Verbal (5%)	Cognition (10%)		Feeling (11%)
5	Perception (3%)			
6	Existential (1,5%)			

From the comparison, verb collocates of *will* in BI X and BI XII are more likely to follow those in COCA. Although there are *feeling* (in BI X and BI XII) and *cognition* (in BI X) type that cannot be found in COCA. Furthermore, there are *verbal* and *existential* types in COCA in which they are absent in all textbooks. All in all, COCA and BI X and BI XII share something in common with some exceptions. It means that they don't resemble 100%. However, it is too premature to say that the choice of verb collocates in textbooks is not authentic, but there is some suggestion to reconsider by ELT practitioners.

In terms of percentage, the percentage composition of verb collocates in COCA and textbooks seem to show a distinct number. Even though BI XI shows *relational* and *material* types in the first and second rank respectively (87,5% and 12,5%) similar to COCA (*relational* 53% and *material* 25%), but there is no more verb collocate type found in BI XI. In other words, the number of future tense markers used in BI XI is too small to be discussed any further. Meanwhile, the types of verb collocate of *be going to* are described below.

Table 11: Types of verb collocates of *be going to*

Rank	COCA	BI X	BI XI	BI XII
1	Material (62,5%)	Material (100%)	-	Material (100%)
2	Possession (12%)			
3	Perception (12%)			
4	Existential (8,5%)			
5	Verbal (5%)			

Table 11 presents the limited number of use of *be going to* as well as the limited number of collocates being described. If the types of verb collocate in COCA are *material, possession, perception, existential*, and *verbal*, there is only one type, i.e. *material,* in the textbooks. The gap of verb collocates found in COCA and textbook in this case is much more notable. The small number of *be going to* as future tense marker in conversation parts of the textbooks shows that future tense marker is apparently of unimportance.

5. DISCUSSION

This study shows that there are some discrepancies in the use of adjectives between the textbooks and COCA. The discrepancies include the different frequency of use of *will* and *be going to* in which textbooks seem to ignore the latter and focus on the former. In textbooks, *will* is very dominant over *be going to* and only textbook BI X that resembles those in COCA. As for the collocates, the left collocates (L1) of *will* are not in line with those recorded in COCA. In COCA, the pronoun *we* is the most frequent, but it is less frequent in the textbooks. The right collocates (R1), however, shows no notable problem in comparison with COCA. Finally, aside from some exceptions, the semantic of the verb collocates (the right collocates) are most likely to follow those in COCA, especially in BI X and BI XII. The results of this study confirm those of the previous work showing that there are some mismatches between future tense markers in textbooks and corpora (Ojanen, 2008). Ojanen (2008) also finds out that the use of future tense reference in the textbooks differs from those in corpora.

In regard to conversations in textbooks, some mismatches have been identified between the language presented in the conversations of the textbooks and the authentic conversations recorded in corpora. Cheng and Warren (2007) examine the expression of understanding in textbook conversations and a spoken corpus and figure out that the use of the expression in the textbooks differs from those in the spoken corpus. Siegel (2014) claims that there are some differences in the topics of textbook conversations compared to real-life conversations among students in Japan. Students mainly talk about their school life in real conversations, while this topic on school life is less frequent in textbook conversations. Similarly, Setiaji (2016) also identifies the differences between conversations in Indonesian EFL textbooks in comparison with authentic conversations. In line with those studies, Oktavianti et al. (2020) point out

that textbook conversations do not correspond to real conversations in terms of the conversational features (e.g. discourse markers, interjections, hesitation devices, etc.). These features are used less frequently and monotonously in the textbooks.

The results of this study, in general, also support the results of other studies comparing textbooks and corpora investigating different linguistic units. Related to the presentation of adjectives in textbooks, Biber and Reppen (2002) explain that adjectival nouns are less frequently used in the textbooks, while this sort of adjective is frequent in the *Longman Corpus of Spoken and Written English* (LCSWE). Having similar results as the other work, Biber and Reppen (2002) set out the fundamental basis of language discrepancies in ELT materials, which is also underpinned by the results of this research. Another discrepancy is also found in the use of modal verbs. Orlando (2009) shows that the frequency of modal verbs in textbooks differs from the frequency of modal verbs in the corpora. Similar to Orlando, Hsieh et al. (2011) investigate the vocabulary input in textbooks and compare it to the storybooks and the requirement of the Ministry of Education. The study shows that the vocabulary in the textbooks does not meet the requirement of vocabulary designed by the ministry. Phoocharoensil (2017) claims that the linking adverbials used in textbooks differ from all the common patterns produced by the English native speakers in the academic context. One of the obvious examples is the use of linking adverbial *hence*. Apart from the fact that *hence* is the third most frequent linking adverbial in the corpus, its use is limited in the textbooks. The identical results mentioning the mismatches of linguistic constructions used in the textbooks and in the corpora to represent natural use lead to the questionable textbook designs.

From the elaboration, it is evident that many textbook writers do not refer to corpora in the design of the textbooks. Despite the increased use of corpus in ELT, Dongkwang and Chon (2011) and Burton (2012) argue that many textbook writers are hesitant to use corpus in the process of writing textbooks. Burton (2012) states that this fact happens due to two different points of view, such as the publishers and textbook writers. In the perspective of the publisher, it is important to minimize revolutionary side to maintain the marketing (Burton, 2012). In the meantime, the textbook writers are not familiar with corpus and have lack of sufficient knowledge on how to utilize corpus for designing textbooks (Burton, 2012). In a wider context of ELT, corpus itself is still less acknowledged because of enormous factors, such as technology use, unfamiliar terms, and limited access to the Internet (Kim, 2019; Kızıl & Savran, 2018; Lai, 2015; Leńko-Szymańska, 2015; Lin, 2016). Therefore, this study attempts to elaborate one of the fundamental reasons of revisiting the language content of textbooks and consult the content with corpora. As for textbook writers, this study tries to suggest considering corpora as the native sources for the improvement of language quality of the textbooks. By consulting with a corpus, it is expected that the language presented in the textbooks are in accordance with the real use of English.

6. CONCLUSION

Investigating the use of future tense markers in textbooks and COCA, it is found that there are some similarities in using *will* and *be going to*. The textbooks and COCA use the same right collocates for *will* and *be going to* and the first-person singular *I*. All of them are frequently found in the textbooks and the corpus. Apart from the similarities, some differences should be taken into account. Among the three textbooks, only BI X that is closer to the empirical data of usage frequency based on COCA by presenting the perfect balance of *will* and *be going to*. Meanwhile, BI XI and BI XII use an unjust portion of *will* and *be going to* in the conversations (*will* being extremely dominant or the only future tense marker as in BI XI). In 'real' English as represented in COCA, *be going to* occupy the third-highest modal verb expressing future/prediction and it is one of the indicators of colloquialization which is supposed to occur in spoken context. Therefore, it is unusual to find *be going to* being less commonly used in the conversation section. As for the collocates, there is a slight difference in which the collocates of *will* in textbooks do not correspond to those in COCA. In terms of the semantic aspect, BI X and BI XII are more likely to resemble the semantic of verb collocates in COCA and there is insufficient evidence for BI XII. In general, the small number of *be going to* used in textbooks, especially in the conversation parts, is rather unnatural. To sum up, even though some parts of future tense marker usage are in accordance with that in COCA, but there are some dissimilarities need to be revisited and re-evaluated. Moreover, among the three textbooks, BI X is more likely to reflect the use of *will* and *be going to* in COCA, but it is because there is a specific discussion about it in the book. Therefore, it is rather problematic to state that BI X is authentic enough and the other two textbooks are not. Briefly, this study finds out that, apart from BI X containing *be going to* discussion, these textbooks need to be re-evaluated in terms of future tense marker usage. This study, however, is a preliminary one and further observation is hence mandatory.

7. REFERENCES

Al-Jaboori, A. T. (2008). *A comparative study between modals in ESL/EFL tertiary textbooks and actual used by EFL teachers* [Ph.D. Thesis]. American University of Sharjah.

Arellano, R. (2018). A corpus linguistics application in the analysis of textbook as national teaching instruments of English as second language in Chile. *Actualidades Investigativas e Educacion, 18*(1), 1–19.

Ayu, M., & Indrawati, R. (2018). EFL textbook evaluation: The analysis of tasks presented in English textbook. *Teknosastik, 16*(1), 21–25.

Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Edinburgh University Press.

Biber, D., & Reppen, R. (2002). What does frequency have to do with grammar teaching? *Studies in Second Language Acquisition, 24*(2), 199–208. https://doi.org/10.1017/S0272263102002048

Brezina, V. (2018). *Statistics in corpus linguistics: A practical guide* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781316410899

Browne, C., Culligan, B., & Phillips, J. (2013). *In focus*. Cambridge University Press.

Burton, G. (2012). Corpora and coursebooks: Destined to be strangers forever? *Corpora*, *7*(1), 91–108. https://doi.org/10.3366/corp.2012.0019

Charalambous, A. C. (2011). *The role and use of coursebooks in EFL* [M.A. Thesis]. http://files.eric.ed.gov/fulltext/ED524247.pdf

Cheng, W., & Warren, M. (2007). Checking understandings: Comparing textbooks and a corpus of spoken English in Hongkong. *Language Awareness*, *16*(3), 190–207.

Choi, H.-Y., & Chon, Y. V. (2012). A corpus-based analysis of collocations in tenth grade high school English textbooks. *Multimedia-Assisted Language Learning*, *15*(2), 41–73.

Collins, P. (2006). Grammar in TEFL: A critique of Indonesian high school textbooks. *TEFLIN Journal*, *17*(1), 1–10.

Comrie, B. (2000). *Tense*. Cambridge University Press.

Davies, M. (2008). *The Corpus of Contemporary American English (COCA): 560 million words, 1990-present*. https://corpus.byu.edu/coca/.

Dongkwang, S., & Chon, Y. V. (2011). A corpus-based analysis of curriculum-based elementary and secondary English textbooks. *Multimedia- Assisted Language Learning*, *14*(1), 149–175.

Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT Journal*, *58*(4), 363–374.

Handayani, N., Isyam, A., & Fitrawati. (2013). Teaching-learning processes: Simple past tense and simple future tense and their perception. *Journal of English Language Teaching*, *1*(2), 291–300.

Hsieh, M., Wang, F., & Lee, S. (2011). A corpus-based analysis comparing vocabulary input from storybooks and textbooks. *The International Journal of Foreign Language Teaching*, 25–33.

Khadim, N. O. A. R. (2015). On teaching future time to EFL learners: Problems and solutions. *Journal of Education College for Women*, *9*(7), 7–23.

Khijasteh, L., & Kafipour, R. (2012). Are modal auxiliaries in Malaysian English language textbooks in line with their usage in real language? *English Language Teaching*, *5*(2), 68–77.

Kim, H. (2019). The perception of teachers and learners towards an exploratory corpus-based grammar instruction in a Korean EFL primary school context. *The Korea Association of Primary English Education*, *25*(1), 123–152. https://doi.org/10.25231/pee.2019.25.1.123

Kızıl, A. Ş., & Savran, Z. (2018). The integration of corpus into EFL speaking instruction:A study of learner perceptions. *International Online Journal of Education and Teaching*, *5*(2), 376–389.

Lai, S.-L. (2015). EFL students' perceptions of corpus-tools as writing references. *Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy*, 336–341. https://doi.org/10.14705/rpnet.2015.000355

Leech, G., Hundt, M., Mair, C., & Smith, N. (2009). *Change in Contemporary English: A grammatical study*. Cambridge University Press.

Leńko-Szymańska, A. (2015). A teacher-training course on the use of corpora in language education: Perspectives of the students. In A. Turula, B. Mikołajewska, & D. Stanulewicz (Eds.), *Insights into Technology Enhanced Language Pedagogy*. Peter Lang.

Lin, M. H. (2016). Effects of corpus-aided language learning in the EFL grammar classroom: A case study of students' learning attitudes and teachers' perceptions in Taiwan. *TESOL Quarterly*, *50*(4), 871–893. https://doi.org/10.1002/tesq.250

Mayangsari, L., Nurkamto, J., & Supriyadi, S. (2018). Cultural content: An analysis of EFL textbook in Indonesia. *International Journal of Scientific and Research Publications (IJSRP)*, *8*(11), 192–199. https://doi.org/10.29322/IJSRP.8.11.2018.p8325

McCarthy, M., McCarten, J., & Sandiford, H. (2014). *Touchstone*. Cambridge University Press. https://www.cambridge.org/gb/cambridgeenglish/catalog/adult-courses/touchstone/methodology-and-research

McEnery, T., & Hardie, A. (2012). *Corpus linguistics*. Cambridge University Press.

McEnery, T., & Wilson, A. (2001). *Corpus linguistics*. Edinburgh University Press.

McEnery, T., & Xiao, R. (2013). What corpora can offer in language teaching and learning. In *Handbook of research in second language teaching and learning*. Routledge. https://doi.org/10.4324/9780203836507.ch22

Mindt, D. (1996). A corpus-based empirical grammar of English modal verbs. *Papers from the Sixteenth International Conference on English Language Research on Computerized Corpora (ICAME 16)*, *16*, 133–141.

Nordberg, C., & Nordlund, M. (2018). A corpus-based study of lexis in L2 English textbooks. *Journal of Language Teaching and Research*, *9*(3), 463. https://doi.org/10.17507/jltr.0903.03

Nordberg, T. (2010). *Modality as portrayed in Finish upper secondary school EFL textbooks: A corpus-based approach*. University of Helsinki.

Ojanen, E. (2008). *Authenticity of future time reference in textbook English: A corpus study*. University of Tampere.

Oktavianti, I. N. (2019). *Verba bantu modal bahasa Inggris: Karakteristik, pemakaian dan perubahan*. Universitas Gadjah Mada.

Oktavianti, I. N., Prayogi, I., Amal, M. A., & 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

Orlando, M. E. (2009). *The frequency and collocation of modal verbs in English as a second language textbooks as compared to Standard English corpora*. University of Québec.

Phoocharoensil, S. (2017). Corpus-based exploration of linking adverbials of result: Discovering what ELT writing coursebooks lack. *3L The Southeast Asian Journal of English Language Studies*, *23*(1), 150–167. https://doi.org/10.17576/3L-2017-2301-11

Pounds, R. (2011). *Analysis of approaches to the present perfect tense in English textbooks published in Japan and in English-speaking countries* [Ph.D. Thesis]. Ball State University.

Radić-Bojanić, B. B., & Topalov, J. (2016). Textbooks in the EFL classroom: Defining assessing and analyzing. *Collection of Papers of the Faculty of Philosophy*, *XLVI*(3), 137–153.

Rahmah, A., Kasim, U., & Fitriani, S. S. (2018). Cultural values analysis in English textbook "Bahasa Inggris." *English Education Journal*, *9*(4), 614–631.

Römer, U. (2004). A corpus-driven approach to modal auxiliaries and their didactics. In J. Sinclair (Ed.), *How to Use Corpora in Language Teaching*. John Benjamins Publishing Company.

Römer, U. (2005). *Progressives, patterns. pedagogy: A corpus-driven approach to English progressive forms, functions, contexts, and didactics*. J. Benjamins Pub. Co.

Scheibman, J. (2001). Local patterns of subjectivity in person and verb type in American English conversation. In *Frequency and the emergence of linguistic structure* (pp. 61–89). John Benjamins Publishing Company.

Setiaji, S. (2016). *An authenticity analysis of conversational texts in the Indonesian EFL textbooks* [Thesis]. Universitas Pendidikan Indonesia.

Siegel, A. (2014). What should we talk about? The authenticity of textbook topics. *ELT Journal*, *68*(4), 363–375. https://doi.org/10.1093/elt/ccu012

Stockton, R. J. (2018). Reculturing language in Indonesian English language teaching. *Indonesian Journal of English Language Teaching*, *13*(2), 131–153.

Tomlinson, B. (1998). *Materials development in language teaching*. Cambridge University Press.

Widdowson, H. G. (1998). Context, community and authentic language. *TESOL Quarterly*, *32*(4), 705–716.

Widodo, H. P. (2018). A critical micro-semiotic analysis of values depicted in the Indonesian Ministry of National Education-endorsed secondary school English textbook. In H. P. Widodo, M. R. Perfecto, L. Van Canh, & A. Buripakdi (Eds.), *Situating Moral and Cultural Values in ELT Materials* (Vol. 9, pp. 131–152). Springer International Publishing. https://doi.org/10.1007/978-3-319-63677-1_8

Zambrana, M. (2017). Corpus analysis of phraseology in an A1 level of German as a foreign language. *Quaderns de Filologia: Estudies Linguistics*, *22*, 13–22.

A Corpus-Based Analysis of Future Tense Markers in Indonesian EFL Textbooks for Senior High School

ORIGINALITY REPORT

10%

SIMILARITY INDEX

PRIMARY SOURCES

1   www.jurnal.unsyiah.ac.id
    Internet
    197 words — 3%

2   Rian Abdus Salamudin, Efransyah Efransyah.
    "ANALYZING THE DEIXIS OF SONG LYRICS IN ADELE
    ENTITLED ALL I ASK", PROJECT (Professional Journal of English
    Education), 2021
    Crossref
    66 words — 1%

3   ejournal.unib.ac.id
    Internet
    63 words — 1%

4   Laura Aull. "First-Year University Writing", Springer
    Science and Business Media LLC, 2015
    Crossref
    48 words — 1%

5   www.lancaster.ac.uk
    Internet
    40 words — 1%

6   documents.mx
    Internet
    35 words — 1%

7   Laleh Khojasteh. "Are Modal Auxiliaries in
    Malaysian English Language Textbooks in Line
    with Their Usage in Real Language?", English Language
    Teaching, 01/29/2012
    Crossref
    28 words — < 1%