The Vocabulary Input of Indonesia's English Textbooks and National Examination Texts for Junior and Senior High Schools

Furqanul Aziez
University of Muhammadiyah Purwokerto, Indonesia

Feisal Aziez
Muhammadiyah University of Purwokerto, Indonesia

Abstract

This study was designed to firstly develop a corpus of Indonesia's junior high school (JHS) and senior high school (SHS) English textbooks (TB) and national examination (NE) texts. By means of corpus analysis, this study then attempted (1) to find out the vocabulary levels (i.e. distribution among the K1 – K20, where K refers to 1000-word band in the BNC word frequency list) of JHS and SHS TBs and JHS and SHS NE texts; (2) to compare the lexical variety of JHS and SHS TBs and NE texts, (3) the number of interdisciplinary academic words, and (4) the number of words beyond the 2,000 high frequency English words the texts contained. Research results reveal that (1) JHS TBs belonged to K-4, K-5, and K-4 levels (i.e. knowing 4,000-5,000 words is necessary for the 95% comprehension of these texts) for Grades 7-9 respectively and SHS TBs belonged to K-4, K6, and K5 levels for Grades 10, 11, and 12 respectively; while the vocabulary levels of JHS NE texts belonged to K4-K5 and SHS NE texts belonged to K3-K5; on average the vocabulary levels of JHS TBs were lower than those of SHS TBs, whereas JHS NE texts levels were slightly higher than those of SHS NE texts; (2) the lexical variety of JHS TBs was the same as that of SHS TBs with the average index of 0,23; while the lexical variety of JHS NE texts was 0,27, which was lower than the SHS NE texts with 0,38; (3) for the words beyond the 2,000 basic words, the JHS and SHS NE texts covered 11.20% and 15.10% of the entire words, or higher than those of JHS and SHS TBs, which covered only 7.89% and 11.87% respectively; and (4) the same profile appeared in the interdisciplinary academic words (Coxhead's AWL), where JHS and SHS TBs contained 1.75% and 3.56% respectively, or lower than JHS and SHS NE texts, which contained only 3.26% and 5.65% respectively of all the tokens in the texts.

Keywords: vocabulary levels, examination texts, textbooks, academic words

Introduction

In the current Indonesian curriculum for junior and senior high schools, teachers are freed from the work of preparing teaching materials. The Ministry of Education provides textbook, which includes instructional materials and the broad lesson programs. This follows that textbook becomes the main tool in the teaching of any subjects including English. Textbook plays a facilitating and even central tool for both teachers and students. It provides not only well-structured texts, but also exercises. Teachers do not have to prepare framework for discussions and homework either. In short, as Jimenez and Mancebo (2008) call, these textbooks are "containers of information and guides to the study of the target language". What Indonesian teachers need to do then is just to plan the lesson on the basis of the textbook, which suit his or her actual context. So influential is the centrality of the textbook in Indonesia that even the ever growing use of electronic device for instructional purposes cannot take over the role the textbooks have so far played in the classroom.

Therefore the quality of an instruction relies very much on the the quality of the textbooks, where the quality of textbooks themselves is among other things affected by the quantity and the quality of the language input contained in them (Orio, 2014). A study (Donzeli, 2007) on the language input and its effect on learners' learning achievements, for example, showed that there was a close relationship between vocabulary input of a textbook and learners' vocabulary uptake. Flogenfeldt and Lund (2016) further stress the relationship between textbooks and vocabulary when saying that what teachers want through the textbooks is that learners are able to express themselves as effectively as possible and to get their messages across to their listeners or readers; and to accomplish this the textbooks need to

provide them with variety of English words.

Even though vocabulary research indicate that vocabulary is fundamental in language instruction, most classroom activities and textbooks do not reflect the significance. The vocabulary and grammar input contained in textbooks is far from being sufficient. In an Indonesian English textbook, *Developing English Competencies for Senior High School* (Doddy, Sugeng, Effendi, 2008), for instance, vocabulary activities appear only in reading section, in which it is given in one of the nine activities in the chapter. The activity is to find synonyms of words (around ten items) taken from one of the texts discussed in the chapter. This type of activity is apparently traditional since it comprises of closed or open tasks only and emphasizes semantic sets (Jimenez, 2014). Besides, no vocabulary learning strategies are present and no recycling is found in all the chapters.

This lack of attention to the vocabulary component of textbooks is even more worse when it comes to the selection and grading of the target words to be taught. Factors such as the number of words, which words, proficiency level, and semantic fields do not seem to be considered well in the selection. Research on vocabulary input in Indonesian high school textbooks is if any still very rare. However, a study by Aziez (2011) on the vocabulary levels of texts used in English national exams, which were administered during the period of 2005-2009 in Indonesia, may help understand this. Aziez's study revealed that the vocabulary of the junior and senior high school English NE texts was at the 4,000-word level, which was of course too high for students at those education levels. However, a more surprising finding was that the 4,000-word level covered 95.96% of the running words in senior high school (SHS) national exam (NE) texts, but only 95.80% of the junior high school (JHS) national exam (NE) texts. This finding is especially surprising because JHS leavers are believed to know less English words than the SHS leavers.

Information regarding the vocabulary input in textbooks is especially important because without prudent consideration of the selection and grading of the words they contain the textbooks would have readability problems. As assured by Nation (2001), if a text contains unknown words more than five percent of the entire running vocabulary items, then it is very likely that there will not be anymore meaning-focused reading activity because too much attention is given to handling language features. As a matter of fact, meaningful reading involves more than just being able to derive meaning from words in a text, but a lack of knowledge of more than 5% of the words in a text can cause reading to be a daunting task (Laufer, 1989). In other words, if learners know less than 95% of the words in a textbook, they would be unable to read the textbook independently. This 95% 'coverage' as the threshold for independent reading was confirmed by some researchers (Hatori, 1979; Laufer, 1989; and Tono *et al*, 1997). Hence, it is quite clear that it is impreative to examine the vocabulary input of the textbooks the learners have to work on and then to determine whether they have the 95% of the running words in the texts (Schmitt and McCarthy 1997, Read 2000, and Hayashi 2002). Without taking this into consideration, learners would have to encounter a discouraging amount of dictionary work (Chujo, 2004).

For the reason, the aim of this study was to generate a corpus of Indonesia's junior high school (JHS) and senior high school (SHS) English textbooks (TB) and national examination (NE) texts used in the 2011-2015 administration as the basis of analysis, and to examine the vocabulary input of the TBs and NE texts, namely the vocabulary levels, the lexical variety, the number of interdisciplinary academic words, and the amount of words which are beyond the first 2,000-word band the texts contain.

Method

This study investigated the vocabulary input of Indonesia's JHS and SHS English TBs and NE texts. Through analysis of those texts, this research attempted to answer the following questions:

1. What are the vocabulary levels of JHS and SHS English textbooks and NE texts?
2. What is the lexical variety of JHS and SHS English textbooks and NE texts?
3. What is the number of interdisciplinary academic words the JHS and SHS English textbooks and NE texts contained?
4. What is the number of words beyond the 2,000-word bands that the JHS and SHS English textbooks and NE texts contained?

To approach the research questions, three main steps were taken. First, English NE texts from administration years of 2011- 2015 and six textbooks representing six grades were selected as the basis for corpus creation. Second, corpus was created following criteria set in advance. Third, the vocabulary input of the JHS and SHS English NE texts was explored using Web VocabProfilers program developed by Cobb (2009) from UQAM.

The instrument

This study used Web VocabProfiler program developed by Cobb (2001) to analyse the texts. Cobb's program was inspired and based on RANGE program, which was created by Heatly, Nation, and Coxhead. RANGE program incorporates the General Service List (GSL) of English Words, the Academic Word List (AWL), and the British National Corpus High Frequency Word List (BNC HFWL). Web VocabProfilers has some sub-programs and this study used two of them.

The first sub-program is VocabProfile BNC, which examines and compares targeted texts with 20 1000-word bands (K1-K20). The output of this sub-program comprises the number of word families, types, tokens, text coverage, cummulative coverage, type-token ration, tokens per type, tokens per family and types per family.

The second sub-program involved is VocabProfile Classic, which examines and compares targeted texts with the 2,000 basic English words (K1-K2) and Academic Word List (AWL). This sub-program produces the number of word families, types, tokens, and percentage that the K1 and K-2 levels and AWL cover.

Creating corpus

Before creating the corpus as the basis of analysis, the first step to take was selecting which textbooks and national examination administration which were used at the same time and encompassing the transition period of the previous and current curriculums. Taking this into consideration, the exam documents taken as samples were those from administration years of 2011 until 2015, while the textbooks which were still in use were as in the following table.

Table 1 Textbooks and examination texts used during the transition period

Levels	**Examination texts**	**Textbooks**
JHS	2011 – 2015	*English in Focus*
SHS	2011 – 2015	*Developing English Competencies*

Since the previous curriculum started in 2006 and the current curriculum started in 2013, the last two years of adminsitration were taken from the previous (2011-2012) and the first three years of administration were taken from the current (2013-2015). In fact, the difference between the previous and the current curriculum lies principally on the teaching method, not on the teaching material, so the selection of the two textbook series were still relevant with both curriculums. Therefore, there were ten exam documents and six textbooks used as the basis for creating the corpus.

In developing corpus of examination texts, the questions (test items) and listening section were exluded and the passages were scanned into ten electronic files and proofread for completeness. The same steps were taken for textbooks, except that the texts taken from the books were not only reading passages but also exercises. After creating the raw corpus, deletion of some words was then carried out (see Table 2). This was done so because otherwise the vocabulary size might inflate and text coverage on the other hand would shrink.

Table 2. Words excluded from the raw corpus

Part of speech	Examples
1. Proper nouns	Shinta, East Java
2. Numerals	2010, 11
3. Interjections	Wow, Oh
4. Unclassified	Ehm, err., GMT,
5. Alphabetical symbols	Kg, cm. Etc., eg.
6. Units	

The exlusion criteria of uncounted words from the raw corpus was rather different from that of Chujo's (2004), where in this study days of the week, months of the year, numerals in words, and prepositional phrases were not excluded. The inclusion of those words in the corpus was merely based on the fact that the words were parts of the English instructional goals at the JHS and SHS curriculums. After the exclusion of the words the number of tokens and types of the JHS and SHS textbooks corpus was as in Table 3 below.

Table 3. Types and tokens in JHS and SHS English Textbook corpus

Texts		Tokens	Types
JHS TBs	Grade VII	4415	998
	Grade VIII	8385	1800
	Grade IX	5259	1646
SHS TBs	Grade X	7561	1640
	Grade XI	8134	1901
	Grade XII	7081	1770

Meanwhile the number of tokens and types of the JHS and SHS NE texts through the five years of administration was as shown in the following table.

Table 4. Types and tokens in JHS and SHS English NE text corpus

Texts		Tokens	Types
JHS NE	2011	3551	1142
	2012	3999	1119
	2013	7932	1917
	2014	4669	1203
	2015	4484	1251
SHS NE	2011	1473	631
	2012	2080	855
	2013	1826	780
	2014	9633	2527
	2015	2790	1095

As shown in Table 4 above there were more words in JHS NE texts than in SHS NE texts. This was because there were more texts in JHS NE texts than in SHS NE texts, not because there were more words in each text.

Profiling vocabulary input

After creating the corpus, the next step was to profile the vocabulary input of both NE texts and textbooks. This step was done by comparing each text with the 20 1000-word bands using Cobb's Web VocabProfilers program, which were partly taken from the BNC HFWL. There were four profiles of the vocabulary input measured through the program: the vocabulary level, the lexical variety, the interdisciplinary academic words, and the number of words beyond the 2000-word band to reach the 95% comprehension.

To measure the vocabulary levels, the percentage level of comprehension coverage was established first. The comprehension level targeted followed some research results in this field. It was confirmed that for the readers to achieve sufficient comprehension and to successfully guess meaning from context they would have to know at least 95% of the running words covered in the text (Nation 2001, Laufer 1989, Schmitt & McCarthy 1997, and Read 2000). To measure the 95% coverage of the words contained in the texts the 20 1000-word bands in the BNC HFWL were used. This was done by measuring how many words from the top of the bands that a reader would need to know in order to achieve a roughly 95% coverage of the targeted words (Aziez, 2011). That is to say, the vocabulary level of each target text was determined by how many words in the text counted from the top of the BNC HFWL account for more or less 95% of the running words in the text. The coverage of each 1000-word band over the target text was decided by counting how many word bands needed until the coverage reached approximately 95%.

To put it in other words, the vocabulary level of each target text can be measured following these steps: (1) count how many percent of the words in the text can be found in the first 1000-word band, (2) then count how many percent of the remaining words can be found in the second 1000-word band, (3) then how many percent of the remaining words outside the first and the second 1000-word band can be found in the third 1000-word band, (4) and so forth until the cummulative percentage reached more or less 95%. To make it clear look at the following table.

Table 4. Coverage count of a junior high school text

Vocabulary levels	**Coverage**	**Cummulative Coverage**
K1 Words (First 1000-word band)	85.62%	85.62%
K2 Words (Second 1000-word band)	7.33%	92.95%
K3 Words (Third 1000-word band)	2.21%	95.16%

As seen at the table, the first 1000-word band cover 85.62% of the words in the passage or in other words 85.62% of the words in the text can be found the first 1000-word band. And then, 7.33% of the words in the text can be found in the second word band; and then 2.21% of the words in the text are found in the third band. Up to this level (K3 or Level-3) the cummulative coverage has already reached 95.16%, which means that the knowledge of the 3000 words (or Level3) is large enough to enable the learners to comprehend the text independently.

In fact, the text coverage cannot be determined precisely because the number of words within one band may cover more than 1% of the words in the text. For instance, a text may reach 94.10% comprehension coverage at 3,000-word level, but at level 4,000-word level its comprehension coverage has touched 97.20%. The same steps were taken to measure the words beyond the 2000-word level and AWL coverage. As for the lexical variety index was obtained by dividing the types by the tokens of each text. The VocabProfilers program generated the index as well.

Results and discussions

Vocabulary levels of JHS and SHS textbooks and NE texts

According to Hsu (2009) vocabulary level was defined as the number of words counted from the top of BNC HFWL which accounted for 95% of the running words in a text, assuming that for a successful independent reading, where he could guess meanings of unknown words from context and gain reasonable comprehension, a learner needed to know at least 95% of the words in a text. The text level of each text was measured by calculating the number 1000-word bands

needed to reach the total coverage of approximately 95%.

Table 5. Vocabulary Levels of JHS and SHS TBs

Texts		**Vocab Levels**	**Coverage (%)**
JHS TB	Grade 7	K-4	95.26%
	Grade 8	K-5	95.43%
	Grade 9	K-4	96.16%
Average		**K-4.33**	**95.62%**
SHS TB	Grade 10	K-4	95.73%
		K-6	94.78%
	Grade 11	K-5	95.38%
	Grade 12		
Average		**K-5**	**95.30%**

After profiling all the textbook corpus one by one on the base words (the BNC HFWL 20), the results indicated that the vocabulary levels of English textbooks for JHS varied at each grade, where grade VII reached level 4,000 (K-4), grade VIII raised at level 5,000, and grade IX dropped at level 4,000. However, eventhough grade VII and grade IX belonged to the same level (K-4), the cummulated text coverage for grade IX already reached 96.16%, while grade VII only 95.26%. That means, by knowing the same amount of words (4,000) a learner would know 96.16% of the total words in grade IX textbook but only 95.26% in grade VII textbook. For textbook grade VIII, the total words needed to know 95.43% of the entire words in the book was 5,000 (Level 5000), almost the same number needed to know the words of grade 12 textbook.

Senior high school textbooks showed similar profile, where textbook for the first year students was at level 4,000, then raised one level at the second year, and dropped one level at the third year. The difference was that senior high school textbook raised two levels at the second year and dropped one level level after that. On average the vocabulary level of junior high school textbooks was at 4,000 and that of senior high school was 5,000. If we take a look at Tabel 7, there were three points worth highlighting in vocabulary levels of the textbooks: (1) there was an increasing tendency from junior to senior (average: K-4.33 to K-5), (2) the books showed similar patterns, where there was an increase in year two but slope in year three, and (3) the vocabulary level of year two was always the highest for each level of education. For senior high school level grade two students needed the largest number of vocabulary, where even with 6,000 words students would only know 94.78% of the entire words. With such vocabulary levels the textbooks have reading them would mean working on quite a number of new words.

Table 6. Vocabulary Levels of JHS and SHS NE texts

Texts		**Vocab Levels**	**Coverage (%)**
JHS NE	2011	K-4	94.92%
	2012	K-4	96.10%
	2013	K-4	95.74%
	2014	K-4	95.25%
	2015	K-4	95.33%
Average		**K-4**	**95.46%**
SHS NE	2011	K-4	96.07%
	2012	K-3	95.41%
	2013	K-4	96.24%
	2014	K-4	95.12%
	2015	K-5	95.12%
Average		**K-4**	**95.60%**

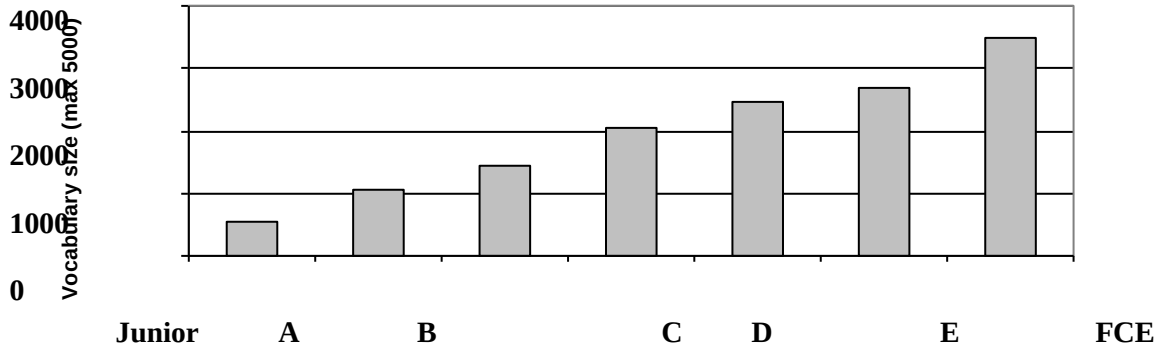
Similar picture is seen in the national examination texts, where the high vocabulary levels and unsystematic pattern of vocabulary level among the years of administration and levels of education were apparent. In year 2011 the vocabulary level for JHS NE texts reached K-4 with 95.46% coverage, which meant Indonesian junior high students at that year needed to know at least 4,000 words to attain 95.46% comprehension, and in the following years the level was the same but with varying degrees of coverage. In 2012, the texts for senior high school examination were at K-3, which was lower than the texts for junior high school examination in the same year. On average the vocabulary level of junior high school examination texts throughout five years of administration was more or less the same as that of the high school.

If comparison was to be made between the vocabulary input of TBs and NE texts some interesting points were worth highlighting. First, on average the vocabulary level of TBs was higher than that of NE texts, where the average level for JHS TBs was at K-4.33 while for the JHS NE texts the level was at K-4.20; and for SHS level of education, the SHS TBs was at K-5 while SHS NE texts was at 4. Secondly, despite the fluctuation among grades the vocabulary levels of TBs showed an increasing tendency from the lower JHS to the higher SHS. This could particularly be seen from the average K-4.33 for JHS TBs and the average K-5 for the JHS TBs. Thirdly, when TBs showed a slight increase in the vocabulary levels from the lower JHS TBs to the higher SHS TBs, NE texts showed a slight decrease, where the JHS NE texts reached K-4.20 but SHS NE texts only K-4. This was because during the five years of administration there was one year when the texts level dropped to K-3 (year 2012), while the other years remained stable similar to those of JHS.

The high levels and fluctuating patterns among the grades in the vocabulary input of Indonesia's textbooks might hint that the books were developed without proper consideration of language components. This suggests the fact signalled by Richards (1976: 76) when stating that issues on the vocabulary teaching and learning have never attracted the same extent of interest within the realm of language instruction as have other issues like grammatical competence, speaking, reading, or writing. In lexical domain, aspects such as the selection of words (how many and which words), their organization, presentation, practice, and review of the words were not given sufficient attention.

As demonstrated by Milton's (2006) research result on Greek learners of English as a foreign language across all the classes, about 500 words are learned every year in around 100 meeting hours (see Figure 1). The data on learners' vocabulary uptake was collected using a test developed using the most frequent 5000 words in English. As seen in Figure 1 below the Junior class acquired around 500 words within one year study, and so did the other classes, and at the end of the sixth year study (Class E), the learners had known about 2750 words.

Figure 1. Vocabulary growth among Greek learners of EFL (Milton 2006)



When it comes to how many words are to be presented to the learners, Nation (2001) asserts that the first 2000 most frequent words are imperative to language learning. This view is supported by Milton (2009) who contends that the systematic teaching of the words is worth the instructional time since this number accounts for around 80% of a text coverage, while the next 2000 frequent words account for a mere 8% coverage. To this point, McCarten (2007) concludes that firstly, the 2,000 up to 5,000 most frequent vocabulary items are a moderate amount to be presented to English learners in classrooms; secondly, considering the huge number of words that the learners need

to know to become self-sufficient readers it is equally essential to help them cope with *how* to learn the words as well as with *what* to learn. She also recommends that teachers identify which words among the 2,000 to 5,000 are to be selected and give them priority in teaching. In other words, after the selection of the target words on a certain criteria like culture and needs, the teacher needs to distribute the target words among the classes on priority basis as well.

Given the information above, it was quite obvious that Indonesian textbooks did not pay adequate attention to the vocabulary selection criteria, organization or arrangement of the vocabulary, presentation techniques, practice to use the vocabulary items, recycling of the items throughout the classes, and word learning strategies. The textbooks concentrated on building the language skills, listening, speaking, reading and writing, on which the books were organized. Each chapter comprised of listening, speaking, reading, and writing sections and vocabulary was introduced only in reading section in the form of an exercise.
The exercise required the students to find out synonims of ten difficult words. There were no instructions of how to find the synonims or of what to do with the words other than finding the synonims.

Lexical variety of JHS and SHS TBs and NE texts
One of the measures of text difficulty level other than vocabulary level is lexical variety. This measure refers to the proportion of different words (type) to the entire running words in a text (token). The smaller the lexical variety index a text holds the easier the text is assumed, because there are more repeated words in the text. Table 7 below shows that the average lexical variety of all textbooks across grades and across education levels is at the same index 0.23. One interesting case occured in textbook for Grade IX, where its index reached 0.26, or the highest of all classes. This again supports the fact that the books were not developed with ample attention to vocabulary input.

Table 7. Lexical variety of JHS and SHS textbooks

Texts		**Lexical Variety**
JHS TBs	Grade VII Grade VIII Grade IX	0.23 0.21 0.26
Average		**0.23**
SHS TBs	Grade X Grade XI Grade XII	0.22 0.23 0.25
Average		**0.23**

A sound picture in this study appears in the lexical variety of NE texts, where the average index for SHS was higher than that of JHS. If we see the pattern one by one education level, we can see a U-shaped curve in the JHS data and a rather sloping line in the SHS.

Table 8. Lexical variety of JHS and SHS NE texts

Texts		**Lexical Variety**
JHS NE	2011 2012 2013 2014 2015	0.32 0.28 0.24 0.26 0.28
Average		**0.27**

SHS NE	2011 2012 2013 2014 2015	0.43 0.41 0.43 0.26 0.39
Average		**0.38**

If we see Table 8 above we may conclude that SHS NE texts are slightly more difficult because the average index of lexical variety is higher, even though the vocabulary items belong to the same level. This is in part because in SHS NE texts there were less repeated words.

Interdisciplinary Academic words in JHS and SHS textbooks

Because Coxhead's (2000) academic words in the AWL are assumed to reach nearly 10% coverage of the whole words in general academic texts, they definitely "gives more return to learners who wish to pursue further study than the next 1,000 words after the 2,000 basic English words" (Aziez 2011: 24).

Table 9. Academic words and words beyond the 2000-word level in JHS and SHS textbooks

Texts		**Academic words**	**Words beyond 2000-word band**
JHS TBs	Grade 7 Grade 8 Grade 9	2.35% 1.39% 1.53%	2.35% 1.39% 1.53%
Average		**1.75%**	**1.75%**
SHS TBs	Grade 10 Grade 11 Grade 12	2.11% 2.90% 5.67%	2.11% 2.90% 5.67%
Average		**3.56%**	**11.87%**

Table 9 above displays the coverage of interdisciplinary academic words in each English textbook for both levels of education. The table proves that SHS TBs bore more academic words than the JHS TBs. A normal pattern was shown in SHS TBs in which the percentage tended to raise from the lower to the higher grade. The table shows that Grade 12 textbook bore the most academic words across levels of education (5.67%) and Grade 8 TBs the least (1.39%). A divergent index occured in Grade 7 (2.35%), where it was not only the highest in the education level, but also higher than that of Grade 9 (2.11%). In conclusion, on average the vocabulary items in terms of academic words in JHS and SHS TBs demonstrate a normal profile. Besides, the vocabulary input in terms of academic words coverage of the TBs shows that there was a correlating pattern between the academic words and words beyond the 2000-word level: the indices seem to correlate quite well. Furthermore, though not exactly the same, the pattern also applied to the JHS and SHS NE texts.

Table 10. Academic words and words beyond the 2000-word level in JHS and SHS NE texts

Texts		**Academic words**	**Words beyond the 2000-word band**
JHS NE	2011 2012 2013 2014 2015	3.77% 2.55% 3.73% 3.73% 2.53%	13.15% 9.00% 11.46% 11.78% 10.62%
Average		**3.26%**	**11.20%**
SHS NE	2011 2012 2013	4.70% 6.13% 6.26%	13.56% 13.94% 15.93%

	2014 2015	5.42% 5.72%	15.33% 16.73%
Average		**5.65%**	**15.10%**

As Table 10 shows, the fall and raise of academic word indices are followed by the fall and raise of the words beyond 2000-word band indices. For instance, for JHS NE texts in 2011 the academic words reached 3.77% coverage and in 2012 the index dropped to 2.55% coverage. The same case happened to words beyond 2000-word band with the drop from 13.15% to 9.00%. the same pattern appeared in other years of administration.

Table 10 above also demonstrated that NE texts contained more academic words than TBs, where throughout the five years of administration on average JHS NE texts contained 3.26% compared to JHS TBs, which contained only 1.75%; and SHS NE texts contained 5.65% compared to SHS TBs which only contained 3.56%.

Words beyond the 2000-word band

For the words beyond the 2000-word level (using BNC HFWL) a question which can lead to our discussion here is "How many more words would Indonesian students need to know in order to satisfactorily comprehend the texts they encounter in their classes and examination rooms?".

As stated in the instrumentation part of this paper the word list used in this study was BNC HFWL, which listed exactly 2000 word families. Assumed that the 2000 word families are the target number of words for junior and senior high school students, how many more new words the students would meet in their textbooks and how many of them would enable them to read the texts successfully. Table 10 above shows that the junior high school students across grades in Indonesia on average were not familiar with around 7.89% of the words in the textbooks, while the senior high school students were not familiar with on average 11.87%. That implies for the JHS students to know 95% of the running words the texts they would have to learn about 2000 more words.

Table 10 above also suggests that when JHS students still needed to know 2000 more words to attain around 95% comprehension SHS students would have to know many more vocabulary items to reach the same level of comprehension. Using Milton's approach to the text coverage beyond the 2000-word band, SHS students had to know more about 3000 words to attain 95% or more comprehension.

The national examination texts show a significantly rising profile, in which NE texts for both JHS and SHS contained more words beyond the 2000 most frequent word families than the textbooks. The words beyond 2000-word band in JHS textbooks covered 7.89%, while in JHS TBs amounted up to 11.20%. In SHS TBs the coverage reached 11.87%, whereas in SHS NE texts the coverage amounted up to 15.10%.

Pedagogical implication

The results of this study suggest that the textbook authors and examination constructors take the following into consideration. First, determining the amount of the target vocabulary input is imperative if the textbooks, and consequently the examination texts, are to be within the learners' assumed linguistic development. The decision on the ultimate target for the whole six year period can then be broken down into yearly target then followed even into semesterly. Second, the decision of the number of words to be taught can then be followed by grading them by level of education. Grading means also distributing words by frequency, the vocabulary levels. Third, selecting the words to be taught using criteria as the ones proposed by some educators like Stahl & Nagy (2006) is crucial: how important are the words for the students to learn in order to be able to communicate at their level (*importance*), whether the words are useful in other contexts (*transferability*), and do the words contain root, base word, and afix that can be used to learn related words (*usefulness for generative study*). Fourth, when using authentic material for the textbooks, adoption and adaptation principles must be put into practice. The adoption of authentic material without adaptation will only result in daunting texts as many of the them were written by and for native speakers.

Finally, the presentation of the target vocabulary must take the *what* and *how* into consideration. Concerning the *what*, Qi Pan and Runjiang Xu (2011) stated that there were three parts of vocabulary that needed to be taught: pronunciation

and spelling, grammar, and word formation. Students need to know what a word looks like and what it sounds like; they should also know the change of form of the word in certain grammatical context; and they should also be aware that many words can be broken broken down into its components. As to the *how*, Qi Pan and Runjiang Xu suggested the use of some strategies: teaching in context, employing semantic field, expanding vocabulary through word formation, creating mental linkages through mental association, and teaching cultural connotation and cultural differences. Through implementation of these principles in organizing the vocabulary presentation it is hoped that students would get better vocabulary uptake.

Conclusion

There are several important findings obtained from this study: (1) on the average, JHS and SHS textbooks belonged to level K-4 and K-5 respectively; (2) in both education levels year two constantly held the highest vocabulary levels; (3) the lexical variety of both JHS and SHS textbooks was the same; (4) SHS textbooks contained more academic words than the JHS textbooks; and (5) SHS textbooks had more words beyond the 2000-word band than JHS TBs; (5) through the five years of administration JHS and SHS NE texts belonged to the same level of K-4, though JHS NE indicated lower coverage, which meant a little bit more difficult; (6) lexical variety of NE texts was higher than that of TBs; (7) lexical variety of SHS NE texts was higher than that of JHS NE texts; (8) SHS NE texts contained more academic words than JHS NE texts; (9) NE texts contained more academic words than TBs; and (10) SHS NE texts held more words beyond the 2000-word band than JHS NE texts;

To this point, the findings suggest that first the vocabulary levels of both JHS and SHS textbooks were too high. The findings above also indicated that Indonesian textbooks were not developed through a careful consideration in the area of vocabuary organization. This of course is different from Taiwanese English textbooks, where the Taiwanese Ministry of Education set the vocabulary goal of merely 2000 words for primary and secondary school students (Hsu, 2009), which means that the textbooks are organized around, and projected towards the acquisition of, the first 2000 high frequency words.

The findings also reflect some contrasts with the American Council on the Teaching of Foreign Languages (ACTFL) Performance Descriptors for Language Learners (2015). According to the descriptors novice learners, who are at similar level to the students of primary and secondary schools, are expected to be able to produce only a number of high frequency words and formulaic expressions and able to use a limited variety of vocabulary on familiar topics, which all are at the K-2 or below. Furthermore, if we follow the research finding on Greek learners' vocabulary knowledge (Milton, 2006), where the average vocabulary uptake of Greek students was about 500 words for every study year, the vocabulary levels of Indonesian textbooks are also too high. Actually, if Indonesian textbook authors set 500 new words in every textbook for each education level the vocabulary level would be much more manageable. Provided the number, at the end of the secondary school period students would have learned around 3000 words.

However, with the findings of the study, it is quite clear that Indonesian English textbooks may fail to carry out the missions both as guides to the teaching and learning activities and as sources of words. The high vocabulary levels the textbooks set for the texts, practices, and instructions, may lead to the low readability index the textbooks have. And, as claimed by Thornbury (2002) the realization of textbooks as sources of words can be found in the content of the books by means of segregated vocabulary activities, integrated text-based activities, grammar explanation, and task instruction. These vocabulary presentation strategies are apparently not present in the textbooks studied.

References

American Council on the Teaching of Foreign Languages. (2015). *ACTFL Performance Descriptors for Language Learners*. USA, Alexandria VA: ACTFL

Aziez, Furqanul. (2011). Examining the Vocabulary Levels of Indonesia's English National Examination Texts. *Asian EFL Journal*. Professional Teaching Articles – CEBU Issue. *51*

Chujo, K. (2004). Measuring Vocabulary Levels of English Textbooks and Tests Using a BNC Lemmatised High Frequency Word List. *Language and Computers 51*(1) 231-249, October 2004.

Cob, T. (2008). Web Vocabprofiler (Version 2.6) [Computer software].

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly. 34* (2).

Davies, Mark and Timothi L Face. (2005). *Vocabulary Coverage in Spanish Textbooks: How Representative is It?*. Minnesota: Cascadilla Proceedings Project.

Doddy, Achmad, Ahmad Sugeng, Effendi. (2008). *Developing English Competencies for Senior High School*. Jakarta: National Education Department.

Donzelli, G. (2007). "Foreign Language Learners: Words they Hear and Words they Learn: A Case Study". *Estudios de Lingüística Aplicada (ELIA)*.

Flognfeldt, M. E., & Lund, R. (2016). *English for Teachers and Learners*. Oslo: Cappelen Damm akademisk.

Hayashi, H. (2002). *Eigo no Goi Shidou* (Teaching English Vocabulary). Hiroshima: Keisuisha. (In Japanese).

Hsu, W. (2009). College English Textbooks for General Purposes: A Corpus-based Analysis of Lexical Coverage. *Electronic Journal of Foreign Language Teaching 6*(1) 42-62

Jiménez, Catalán and R. Mancebo, R. (2008). "Vocabulary input in EFL textbooks". *Revista Española de Lingüística Aplicada, 21*. 147-165.

Jimenez, Dolores. (2014). A Critical analysis of the vocabulary in L2 Spanish textbooks. *Porta Linguarum 21* 163-181

Laufer, B. (1989) 'What Percentage of Text Lexis Is Essential for Comprehension?' in C. Lauren and M. Nordman (eds.), *Special Language: from Humans Thinking to Thinking Machines*. Clevedon: Multilingual Matters. 316–323.

McCarten, Jeanne. (2007). *Teaching Vocabulary: Lesson from the corpus lesson for the classroom*. New York: CUP.

Milton J. (2006). X_Lex: the Swansea Vocabulary Levels Test. In C. Coombe, P. Davidson and D. Lloyd (eds), *Proceedings of the 7th and 8th Current Trends in English Language testing (CTELT) Conference (4)* UAE: TESOL Arabia, 29-39.

Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters.

Milton, James and Thomaï Alexiou. (2010). Developing a vocabulary size test in Greek as a foreign language. *Advances in Research on Language Acquisition and Teaching: Selected Papers 2010 GALA.*

Nation, P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Orío, Susana Fernández. (2014). "Vocabulary knowledge dimensions in EFL textbooks". *Encuentro (23)*. 30-37.

Qi Pan and Runjian Xu. (2011). Vocabulary Teaching in English Language Teaching. *Theory and Practice in Language Studies. 1*(11), 1586-1589.

Read, J. (2000). *Assessing Vocabulary.* Cambridge: Cambridge University Press.

Schmitt N. & M. McCarthy. (1997). *Vocabulary, Description, Acquisition and Pedagogy.* Cambridge: Cambridge University Press.

Stahl S. A., & Nagy, W. (2006). *Teaching Word Meanings.* Mahwah, NJ: Lawrence Erlbaum Associates.

Tono, Y. (ed.) .(1997). *Eigo Goi Shuutoku-ron* (Theories of Teaching and Learning English Vocabulary). Tokyo: Kagensha. (In Japanese).

Wardiman, Artono, Masduki B Jahur, and M Sukirman Djusma. (2008). *English in Focus for Junior High School series (for grade VII-IX)*. Jakarta: National Education Department.

About the Authors

Furqanul Aziez teaches at Muhammadiyah University of Purwokerto, Indonesia. His main research interests include vocabulary, TEFL, and language testing. He can be reached at furqanulaziez@ump.ac.id or f.aziez2010@gmail.com Afiliation address: University of Muhammadiyah Purwokerto Jl. Dukuhwaluh, Kembaran, Purwokerto 53182, Indonesia

Feisal Aziez teaches at Muhammadiyah University of Purwokerto, Indonesia and is now a Ph.D student of Multilingualism Doctoral School, University of Pannonia, Hungary. His research interests are TEFL and language acquisition. He can be reached at feisalaziez2010@gmail.com Afiliation address: University of Muhammadiyah Purwokerto Jl. Dukuhwaluh, Kembaran, Purwokerto 53182, Indonesia