A Lexical Study of EFL Reading Materials in Developing English Competencies 3

Ro’ifah¹, Rosiana Silfani¹

¹Universitas Muhammadiyah Surabaya, Indonesia
*Correspondence: roifah@um-surabaya.ac.id

ABSTRACT

A textbook is one of the essential learning sources for EFL/ESL learners, especially in vocabulary acquisition. It must be evaluated and selected when performing eight categories (aims and approaches, design and organization, language content, skill, topic, methodology, teacher’s book, and practical considerations). The language content category evaluates whether the EFL/ESL materials used in the learning process for vocabulary teaching are adequate in terms of vocabulary quality and range, the emphasis placed on vocabulary development, and strategies for individual learning. It is possible to say that learning a language is fundamental and so deeply dependent on vocabulary. This study is aimed to identify the number of lexical units found in the reading materials of an EFL textbook, Developing English Competencies 3, and to measure the sufficiency of lexical units found in the reading materials of the textbook in vocabulary quantity and range by comparing to the general frequency list, Michael West’s GSL (1953). The book became the object of this descriptive quantitative study with the specific linguistic field, corpus. A computational tool, RANGE, and FREQUENCY, designed by Nation (2001) and then developed by Heatley, Nation, and Coxhead (2002) was applied to count the lexical units easily. It was found that the book was clearly imbalanced concerning the amount of 5,064 tokens and 1,434 types with the distribution of 57.1% in range one, 14.9% in range two, and 5.2% in range three throughout three ranges described by Nation. In summary, the book didn’t propose the foreign learners’ expectations and necessary opportunities for automatizing vocabulary. Besides, it is also limited to offering all possible comprehensive situations, which the students may find in real life.

1. Introduction

It is very vital for English teachers to select teaching and learning materials, especially reading materials presented in a textbook that will be used in the learning process. It is caused by appropriate material to the student’s level that can develop the student’s language proficiency. They can understand most of the message delivered by the vocabulary and the structural elements of the texts. Reading is beneficial since it helps to improve the mind. The mind is a muscle that can be exercised. It needs to be worked out. Understanding the written word is one of the ways the mind develops. Teaching young children to read aids their language development and aids their learning of how to listen. A universal objective of English as a Foreign Language / EFL teaching in secondary school is to improve reading skills, especially in English (Rahman & Weda, 2019; Hasnia et al., 2022). Reading is admitted as one of the most important goals for most foreign language learners to get information, pleasure, and study purposes. The ability to read can open the mind up about worlds for both children and adults. However, the ability of Indonesian students to comprehend the reading material is very poor (Junaidi et al., 2020; Sukmawaty et al., 2022; Mokoginta et al., 2021).

A textbook is usually evaluated and selected as a good book when it performs eight categories such as aims and approaches, design and organization, language content, skill, topic, methodology, teacher’s book, and practical considerations (Lodhi et al., 2019; Farisatma et al., 2017). The third category, language content, evaluates whether the material for vocabulary teaching is adequate in terms of quality and range of vocabulary, the emphasis placed on vocabulary development, and strategies for individual learning (Richards, 2017). Materials of EFL/ESL textbooks are supposed to be sufficient quantity and range of vocabulary, vocabulary improvement, and strategies for self-learning. In vocabulary acquisition, the amount of lexical information should be entered and consolidated in long-term memory efficiently.
Vocabulary is inevitably an integral part of the language system and vocabulary knowledge has a significant role in language comprehension and production (Seddigh, 2012; Indarwati et al., 2022). It is a fundamental language component and affords more bases for how well learners speak, listen, read, and write. Second-language readers of English need to have a vocabulary of at least 3000-word families in order to have some knowledge of 95 percent of the running words in a text (Read, 2000). In order to comprehend a text, the readers have to be familiar with a 3000-word class at any level or 95 percent of the words in the text. Vocabulary and better meaning of words are one of the key components in understanding discourse. It is possible to say that learning a language is deeply dependent on vocabulary. Consequently, a good textbook is essential to hold up the enrichment and development of vocabulary which becomes an objective of greater importance in all areas of EFL textbooks.

The most obvious feature of a language is consisting of words. Words are not the only units of meaning that bring with them some consequences to language teaching and learning. Larger lexical units must be identified, defined, and conveniently presented first as an object for teaching and learning. Words possess two characteristics: (i) they have meanings that are unpredictable and so must be listed in dictionaries – as lexical items, and (ii) they are the building blocks for words and phrases – words [3]. Both characteristics do not always be together. The previous one, any lexicon includes a range from simple isolated words to formulaic cliché phrases, to tightly bound idioms. Lexical entries comprise the set of information necessary to identify understand and use the associated unit. The latter one is formulated quite vaguely; however, making the formulation more precise belongs not to a book on word formulation but to a book on syntax [3]. A word has a more complex phenomenon because it can have the same, similar, or different meanings if it is used in different situations or for different effects. Besides, it is possible to have a variety of overlapping meanings.

Although many words have predictable meanings, there is nevertheless a tendency to lose motivation over time. Hence, a word that does not start out as a lexical item may in due course become one. Conversely, many of the lexical items in the form of phrases or sentences (idioms or proverbs) have meanings, which can be seen as metaphorical extensions of a literal meaning; so to that extent, their interpretation remains motivated. In fact, there is no perfect equal between words and lexical items in the dictionary list. The practice of most dictionaries reflects cooperation. Some are more generous than others in listing idioms and some are more generous than others in listing words with entirely predictable meaning.

A sentence can be made of many words even repeated ones. In unwritten language, the ‘word’ may be a very elusive thing. It is caused by language is a system that consists of the development, acquisition, maintenance, and use of complex systems of communication, particularly the human ability to do so; and also specific examples of such a system or perspective on education’s importance for development (Hamsia, 2021). There are several ways of deciding what words will be considered as type and token. The term ‘word’ would be ambiguous between a ‘token’ and ‘type’ interpretation. Both spoken sentences and also written sentences may be said to be composed of word-tokens, but they are clearly not word-tokens listed in dictionaries. At one level, words as listed in dictionary entries are types and tokens. The type-token distinction is pertinent to the concept of ‘word’.

The token is defined as an instance and individual occurrence of a type (Mc Carthy et al., 2002). What is sometimes called ‘running words’ is a simple way to count each word form in a spoken or written text. If the same word form occurs more than once, each occurrence is counted as a token. Another way, type or ‘different word’, if we see the same word again, we do not count it again (Mc Carthy et al., 2002). Therefore, a repeated word is calculated as a single type only.

For example, how many words are in the following sentence, Kam bodja goes to Edinburgh next week, and she intends to go to Washington next month. In this sense, the third word (to) is as same as the eleventh, and the fifth word (next) is as same as the thirteenth, so there are fourteen tokens but only twelve types in the sentence. On the other side, the third and eleventh words of the sentence are distinct tokens of a single type. Equally, the third and eleventh word as a type is represented twice in this sentence, and similarly the fifth and thirteenth words.

Over the past two decades, many researchers have exposed a great deal about the kind of vocabulary instruction that is most effective for helping students to improve their knowledge of words. There is no single instructional method, which is sufficient for optimal vocabulary learning; therefore, effective instruction must use a variety of methods to help students acquire new words and increase the knowledge depth of words over time. The most effective instruction can be done through opportunities for both incidental and intentional vocabulary learning. A distinction is commonly drawn between incidental and intentional learning.
Classroom practice and most teaching materials merge both options, allowing for explicit and incidental vocabulary acquisition. The opinion extremely often focuses more on teaching explicitly the most usual words, while less frequent items are left to incidental learning. Incidental learning, however, is heavily dependent on usage. The common problem in most classrooms is that very poor opportunities for communicative interaction. The lack of real communicative contexts and contact with native speakers enlarges further problems to the efficiency of this type of learning.

On the other hand, teaching materials are not adequately substituting for real and intensive language use since they are limited by nature and cannot offer the linguistic richness needed by the students for incidental learning. One criterion might be that vocabulary is well-contextualized (Ro’fah & Hamsia, 2022). Sufficient vocabulary enables learners to be skillful in communicating and accessing information about various fields of science. Therefore, vocabulary enrichment for learners and development must be a greater objective of importance in all areas of EFL. It makes sense to shoulder that EFL/ESL teachers should do everything they can, to ensure their students enlarge the size of their vocabulary. A teacher must teach vocabulary using strategies that force the students to learn words or to encourage vocabulary self-learning by their students (self-study).

However, some Indonesian textbooks have a long tradition that focuses on grammar. LOOK AHEAD 2 has monotonous and inadequate vocabulary (Prayitno, 2011). The importance of vocabulary has not always been adequately emphasized. How often the words rise in the texts is not considered. In addition, another student book is found that the vocabulary of reading texts for the seventh grade SMP/MTs English textbook published by CV. Nadia Sarana Utama Balikpapan only met the criteria of 500 words and the variations to use the words are limited so that certain words only appear (Setiawan, 2006; Jannah et al., 2022). Many textbook authors, both language teaching theoreticians and practitioners seem to have less attention on whether the textbook is suitable or not to help and support easy and effective vocabulary acquisition. They do not consider on some elements that maintain vocabulary acquisition faster.

An analogous reality was also discovered. The certain book in his study reveals a clear unbalance in all the aspects of vocabulary selection if this is compared to the expected frequency list of general English (Criado, 2009). It does not seem to take into account some of the most basic issues affecting vocabulary acquisition, both from the point of view of which words should be learned first and the conditions, which govern vocabulary acquisition. The authors do not seem to consider the corpora, which are the best source to define words more efficiently in linguistic communication.

Therefore, the researcher is eager to know how many lexical units of reading materials are presented in a student e-book. It is caused by the textbook should be a medium and a resource designed to meet educational standarts efficiently (Khasanah et al., 2023). Related to vocabulary acquisition, the electronic books will be analyzed on the following parameters of distribution along general English and the relevance of the learners’ needs in their level, informative level.

2. Methodology

Research design can be well thought-out as a conceptual blueprint within which research is conducted. The descriptive quantitative method was used in this study due to the fact that the data was in the form of numbers. This research used a computational tool to collect the data due to quantitative research is characterized by the use of numbers to represent its data (Read, 2000). It is aimed to find out how many lexical units in the electronic book hold up the vocabulary acquisition especially whether the number of lexical items mentioned in textual materials positively approaches their adequacy to the general frequency list of English i.e. 1,000, 2,000, and 3,000.

The object of this study was Developing English Competencies 3 for Grade XII of Natural and social science programs of senior high school (SMA/MA), in 2008, the e-book was written by Ahmad Doddy, Achmad Sugeng, and Effendi where the copyright has been sold to Department Pendidikan Nasional and published by Pusat Perbukuan, Jakarta. The book was assessed by Badan Standar Nasional Pendidikan (national standard board of education) and decided to be one of the school textbooks because it undertakes the possibility necessary to use for the learning process. It applies the principle not only in the School-Based Curriculum 2006 but also in Peraturan Menteri Pendidikan Nasional No. 34 in 2008. The methodology of the e-book is a communicative approach designed to facilitate the students of senior high school to communicate in English based on the context of the language use. Besides, many senior high schools use the book, especially for natural and social science programs, which enables students can access various science and knowledge to prepare the students to go to university.
Data is definite information used as the basis substantial of the analysis to get the answer to the research questions. In this research, at first, the data was corpus which is in the form of words, lexical items of the reading materials that appeared in the student book. Then, the data were counted in the form of numbers. These data were obtained from the e-book, Developing English Competencies 3 is organized into 5 chapters with 26 pages on average. Every chapter contains more than one text related to the theme. The texts were composed of 39 essays consisting of 4 kinds of genres; narrative, explanation, discussion, and review text, and 11 short functional texts consisting of ‘advertisement’, ‘announcement’, ‘notice’, and ‘poster’. It is worth noticing that the glossary includes about 141 words, which are defined as ‘the most frequent words’, with no further specifications, but it doesn’t include the data.

Despite the corpus used as the source of the study is reading materials in the form of short functional text and essay texts, the short functional text used in this study consists of two only; advertisement and announcement. In detail, two narrative texts from Chapter 1 and four items from Chapter 4, three explanation texts from Chapter 2, two discussion texts from Chapter 3, and two review texts from Chapter 5 are chosen as the samples of essays to study. The glossary clearly contrasts against the 1,434 types with files ranging in length from 107 to 3,994 words (tokens) and the total length of the entire corpus of 5,064 words.

There were two instruments used in this study. They were observation and computational tools. In this case, observation by reading and identifying reading materials was used to select texts in every chapter. Then, to make it easier, the researcher used a computational tool, RANGE and FREQUENCY designed by Nation (2001) and developed by Heatley, Nation, and Coxhead (2002) (Nation, 2001) that is available at http://www.vuw.ac.nz/lals/staff/paulnation/nation.aspx. The computational tool was needed to collect the data easily (Ro’ifah, 2022). By this tool, the lexical units of reading materials found in the textbook were counted and compared to the vocabulary of the three base-word vocabulary ranges defined by Nation in 2001 (1,000, 2,000, and 3,000 most frequent words of English) which the 2,000-word level usually has been set as the most suitable limit for high-frequency words (Nation, 2001).

To obtain the needed data for this research, at first, the data was collected from reading materials by observing and selecting between short functional text and essay. After that, the collected texts were inputted into the computational tool. Then, it was obtained the data, lexical items were classified into three categories; the first 1,000, the second 1,000, and the third 1,000 most frequent words of general English (Nation, 2001). The words which did not include within these first three categories appeared as ‘off ranges’. Afterward, the words were classified as tokens (every word form in the text repeated or not), types (different words in the text such as ‘friend’ and ‘friends’ are two types), and word families (the headword, its inflected forms and its closely related derived forms). Additionally, the identification of tokens vs. types allowed for a contrast between the raw vocabulary inputs and the new words really introduced in the text.

In qualitative research, data were thematically analyzed (Febriana et al., 2022). In this research, the data, which was in the form of a corpus was analyzed using the classic list of GSL. General service words are looked into account suitable as the basis for learning English as a foreign language (Nation, 2001). The list encompasses 2000 words or 80% of the running words in the text are high-frequency words and are considered to be supposed to give first in teaching English as a foreign language. The words found in the textbook were compared to the vocabulary with the three base-word vocabulary ranges defined by Nation in 2001 (1,000, 2,000, and 3,000 most frequent words of English) (Nation, 2001). It would be an indicator of vocabulary knowledge matching the ‘expected’ vocabulary according to the English general list, particularly in relation to the first 3,000 most frequent words.

3. Result and Discussion

A. Result

Developing English Competencies 3, the e-book contains 5,064 tokens or running words. The number of types or distinct words is 1,434 words that are classified into 817-word families as in the following Table 1.

<table>
<thead>
<tr>
<th>WORD range</th>
<th>TOKENS%</th>
<th>TYPES%</th>
<th>FAMILIES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1000</td>
<td>3994 / 78.9</td>
<td>819 / 57.1</td>
<td>566</td>
</tr>
<tr>
<td>2000</td>
<td>343 / 6.8</td>
<td>213 / 14.9</td>
<td>186</td>
</tr>
<tr>
<td>3000</td>
<td>107 / 2.1</td>
<td>74 / 5.2</td>
<td>65</td>
</tr>
</tbody>
</table>
The e-book doesn't familiarize the equal tokens and types. Regarding the word ranges defined by Nation (2001), 3994 tokens, or 78% are categorized as range one. A member of range two is about 343 tokens or 6.8%. The tokens that belong to range three consist of 107 (2.1%). The rest, 620 (12.2%) of the tokens are owned by off-ranges. In total, the distinct words, 1,434 types consist of range 1, 2, 3, and off-ranges.

Range 1 accounts for 57.1% (819 types). Range 2 involves 213 types (14.9%). Then, range 3 contains 74 types (5.2%) only. While off-ranges entail about 328 types (22.9%). About 3,994 tokens (78.9% of the total) that cover only 819 types (57.1%) in 566-word families belong to range 1. This amount is significantly higher and reasonable. The imbalance between tokens and types is a result of the height of lexical item frequency in range 1, which does not contradict the normal distribution of words in texts. In other words, 181 types of range 1 do not occur in the textbook. It is hard to assume that students have already been fully strong on those 181 high-frequency items.

Tokens of range 2 are 343 words or 6.8% of the total with 213 types or 14.9% of the total in 186-word families. This quantity is too low if it is compared to the total tokens in the textbook and also more specifically to the amount of range 1 (as the percentage clearly shows). The amount of types, however, has not been significant in equal, 213 types or 14.9%; likewise in the case for word families. The relative lack of balance in the number of tokens and types regarding the total of lexical units in the textbook designates serious negative consequences.

Range 3 presents a sharp contrast because of containing 6.5 words only with 2.1% of the total or 107 tokens and 5.2% or 74 types. This summation is too small when it is assessed to a total number of tokens in the reading texts. It is far away from expectancy if the goal of the textbook is to reach the informational level since this level for ‘independent users’ needs fluent comprehensive English use in daily life, very much in line with the third 1,000 words included in range 3 and the 2,000 words from the previous ranges, 1 and 2.

‘Off ranges’ that is a more advanced level comprising 620 tokens or 12.2% with 328 types or 22.9% of the total types. This expanse has actually not reached the goal of the textbook. However, the figure is higher than the total types in both ranges 2 and 3. Contrasting between the amount of ‘off ranges’ and the word in ranges 2 and 3 also shows the same case. The whole types of range 3 are 287 types or 20.1%. Thus, the extent of ‘off ranges’ is considered higher than the total types’.

The e-book, Developing English Competencies 3 presents the imbalanced distribution of types. Range 1 proposes high enough types. But, the stooping for new vocabulary breaks off in range 2. In addition, the curve comes down even more in range 3. Nevertheless, it turns up abruptly in ‘off ranges’. However, the distribution of the highest figure is supposed to fit in ranges 2 and 3, while the ‘off ranges’ interval ought to grasp a lower percentage, and range 1 must be granted a moderate representation for consolidation objectives.

B. Discussion

The textbook can be appraised and picked out as a good book when it performs eight categories (Richards, 2017). One of them precisely in the second point of the third category, language content evaluates whether the reading material for vocabulary teaching is ample in terms of quantity and range of vocabulary that emphasizes placed on vocabulary development, and strategies for individual learning. The reading materials of ELT books are recommended not only to proffer the sufficient words supposed to acquire but also to boost the number of tokens and types, vocabulary improvement, and strategies for self-learning. However, the electronic book, Developing English Competencies 3, seems far away from the expectation in terms of the amount of lexical unit. Besides, the distribution also reveals a strongly marked imbalance in the new words brought together in ranges 1, 2, and 3 versus ‘off ranges’.

The token or ‘running words’ is an instance or individual occurrence of a type. It is a simple way to count tokens since it counts each word. In fact, the reading texts of Developing English Competencies 3 cover 5,064 tokens. In addition, the type or different word, a repeated word is counted as only a single type. The amount of distinct words or types of 5,064 tokens is 1,434 words classified into 817-word families as mentioned in Table 1 above.

The reading and understanding of authentic written texts necessitate a vocabulary of three to five thousand word families, word family: a base word completed with its derivatives [1]. To achieve good learners’ understanding of the text, it is required at least the first 3,000 most frequent words or the three base-word vocabulary ranges (1,000, 2,000, and

<table>
<thead>
<tr>
<th>Off ranges</th>
<th>620 / 12.2</th>
<th>328 / 22.9</th>
<th>??</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>5064</td>
<td>1434</td>
<td>817</td>
</tr>
</tbody>
</table>
3,000 most frequent words of English). Nevertheless, another different case is found in this study such as acquainting with too small numbers of both types and word families. The contrast is demonstrated between the amount of ‘off ranges’ and the word ranges 2 and 3. The higher figure should belong to ranges 2 and 3, while the ‘off ranges’ interval is supposed to take a lower percentage.

4. Conclusion

In this case study, it is found that the e-book is clearly unstable both quantity and distribution of vocabulary offered throughout the three ranges defined by Nation in 2001. The range of 2,000 to 3,000 word families is a reasonable reference for measuring reading capabilities for learners. But, the lexical items of the reading material explored in the electronic book do not seem to reflect some of the most basic issues affecting vocabulary acquisition. The e-book does not offer teachers and students the expected and necessary opportunities for automating vocabulary since the types and the word families are less than the minimal limitation. The words offered and presented in the e-book as a goal to reach or vocabulary acquisition are not in line with the word list of general English.

In addition, the comparison of the results does not allow for an optimistic assumption. Imbalanced points need to worry regarding the number of lexical units presented in Developing English Competencies 3. In the e-book, the ascending curve for new vocabulary breaks off in range 2, falls down even more in range 3, and ascends abruptly in ‘off range’. Then, the amounts of word range 2 and 3 (287 types or 20.1% of the total) that the learners need to increase their vocabulary are lower than of off ranges that is not the goal of the e-book (328 types or 22.9% of the total). This means that the learners learn new words of poor potential for comprehension in the informational level. It can be said that it is equitable to acquire vocabulary for range 1 only but not for ranges 2 and 3.

The textbook, Developing English Competencies 3 is limited to propound all possible comprehensive situations in real life. It is advocated for the teachers to pick the materials, which can augment the learners’ vocabulary. Furthermore, a lot better for the book authors to be more selective using the reading materials in the student book by considering the words including GSL (General Service List) of English Words. More study, however, is anticipated to carry out the whole content of the book evaluation.

References


