An Investigation of EFL Students’ Vocabulary Size and Level at an Indonesian Private Tertiary Institution

Fenty L. Siregar¹, Henni²
Universitas Kristen Maranatha, Indonesia¹,²
Email: fenty.siregar@outlook.com¹

| Submitted: 18/12/2022 | Revised: 28/01/2023 | Accepted: 27/07/2023 |

E-ISSN : 2579-4574  
P-ISSN : 2549-7359

Abstract. Having good vocabulary knowledge is essential for every student who learns a new language. However, there are only limited previous research projects which have studied the development of English language students’ vocabulary knowledge. Thus, this current study investigated English students’ vocabulary progress of participants who studied at a private tertiary institution in West Java, Indonesia. The participants came from three batches: 2018, 2019, and 2020 to find out the difference of development of vocabulary among them after studying English for two semesters. In its investigation, this study used two different instruments: the Vocabulary Size Test and the Vocabulary Level Test which respectively were created by Nation & Beglar (2007) and Webb et al. (2017). The findings show that after two semesters, the most improved vocabulary size occurred to participants from batch 2019. On average, the vocabulary size achieved after two semesters is enough to read more advanced texts. However, the average score of their vocabulary level test shows that the high-frequency words were not yet mastered, with a decrease on the average score of 2020 students. The finding might imply that there should be more intervention in the teaching and learning of high-frequency words.

Keywords: EFL Students, Vocabulary Size, Vocabulary Level

https://ojs.unm.ac.id/eralingua

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
INTRODUCTION

Mastery of vocabulary plays a vital role in L2 learners’ writing and reading competencies. Thus, to help learners improve those abilities, it is essential that teachers know the learners’ vocabulary size and level. Studies have shown that when teachers can identify the L2 learners’ vocabulary size and level, they can gain insight on their students’ reading ability (e.g. Laufer, 1997; Qian, 2002), their writing quality (Llach & Gallego, 2009; Schoonen et al., 2011; Yang et al., 2019), and speaking ability (e.g. Derakhshan & Enayat, 2020; Enayat & Derakhshan, 2021; Uchihara & Clenton, 2020). Derakhshan & Enayat (2020) found high-frequency vocabulary could be used as a tool to determine the L2 overall speaking performance. Furthermore, the study of Enayat & Derakhshan, (2021) showed that receptive vocabulary size could inform teachers the dimension of lexical mastery of L2 speakers. Thus, knowing how much vocabulary learners have mastered will give teachers insights on what advice and help they can provide for their students to develop their vocabulary size.

Nation (2012) suggested that to help students increase their vocabulary level, teachers need to find out to which level their students belong in the three vocabulary-level groups: high-frequency (1000-2000 word-levels), mid-frequency (3000-9000 word-levels) and low-frequency words (from 10,000 word-level and beyond). Nation further explained that different word levels required different learning processes. If students were not yet familiar with the high-frequency words, it was important to engage students in reading graded readers and being involved in deliberate teaching and learning processes (Nation, 2012). When they had passed the first level, the students could read mid-frequency readers and engage in intentional learning in order to recognize the mid-frequency 3000-9000 word families (Nation, 2012). Whereas the acquaintance with low frequency word list or level 10,000 and above could be done through broad reading and specialized analysis of words related to a certain subject area (Nation, 2012).

The study of Goulden et al. (1999) found that the word-families mastered by adults whose L1 is English are approximately 20,000 words. Coxhead et al. (2015) investigated secondary school students who were L1 English users in New Zealand and their study revealed that the participants mostly had mastered around 9000 word-families. Nation (1990) stated that the average vocabulary increase of English L1 users was 1000 - 2000 new words per year. Regarding L2 English learners, Nation (2012) noted that the non-European L2 English successful undergraduates at an English-speaking university mastered 5000-6000 word-families; while the PhD students mastered 9000 word-families. Mclean et al. (2014) studied the size of vocabulary of 3449 Japanese English learners in different universities and their findings showed that the size of participants’ vocabulary was 3715.20 word-families on average. Another study by Gibriel (2017) found that the second and the fourth semester Egyptian EFL students respectively knew around 6751 and 7566 word-families. The result of a study by Yang et al. (2019) showed that on average, Chinese graduate students knew 7274.75 word-families. Ozturk (2016) examined the growth of English receptive vocabulary from 174 learners studying at an English Program at a university in Turkey by using VST. The participants’ vocabulary growth was
measured at different stages of the study and the study found that their average vocabulary size was between 5000 and 6000. Although their vocabulary size increased by 500 words a year, the size of the vocabulary decreased in their final year of study. This happened because of the reduced use of English by the students (Ozturk, 2016).

Research on VLT of English L2 learners found that most of the participants in the studies have not yet mastered the first-two of the 1000-word families. An example is the study of Webb and Chang (2012) that examined the level of vocabulary of 166 Taiwanese English learners. Their findings showed that more than 50% of the learners still failed to master high-frequency words in the first 1000 word-families and just around 16% of them were proficient in the 2000 word-families, even though the learners had learned English for nine years.

In the context of Indonesia, several studies on vocabulary size were present and overall, the findings of these studies indicated that the more recent the study was, the higher the students’ average vocabulary size would be (see Kusumarasdyati & Ramadhani, 2018; Nurweni & Read, 1999; Romadloni, 2019; Umam, 2016). The study of Nurweni and Read, (1999) which researched 324 EFL students’ vocabulary receptive knowledge found that their vocabulary size was only 1226 English words. Nurweni and Read (1999) concluded that the students should have acquired 4000 words because that amount was expected from the students upon entry to the university. The study done by Umam (2016) examined the vocabulary size of 111 fifth semester Indonesian EFL students. The study came out with the result that the average vocabulary size was 5,873 word-families, with 8,800 words as the highest and 2,800 as the lowest. Another study by Kusumarasdyati and Ramadhani (2018) took data from 216 students at State University of Surabaya. The data were taken from four batches of students majoring at the English Department, and they compared the outcome of each batch. The result of the study showed that the average vocabulary size of the first-year participants was 5425 words, the second was 5641.8 words, the third was 5987.8 words, and the fourth was 6141.3 words. They also found that the vocabulary size of the participants was expanded by 238.8 words per year. Romadloni (2019) replicated Kusumarasdyati and Ramadhani’s study, analyzing the vocabulary size of 242 undergraduate students from the same university, taking data from the classes of 2015 to 2018. He found that on average, the vocabulary size for each batch was 6519.78 words, 7028.13 words, 7040.91 words and 8202.33 words. He also found 2.3 times higher result (approximately 560.85 words every year) when the average vocabulary sizes were compared.

A similar study by Sudarman and Chinokul (2018) examined the vocabulary level of first year students at Kutai Kartanegara University. They found that participants of the study were not yet shown mastery of the 2000 and 3000 word-levels (Sudarman & Chinokul, 2018). Kirana & Basthom (2020) administered Vocabulary Level Test to 319 participants majoring in the English Department, Institut Agama Islam Negeri Ponorogo. The participants were students from five semesters, from semester one to five. They discovered that on average, the students were only familiar with over 1,366-word families, which did not meet the
expected lexical threshold. The researchers proposed the promotion of vocabulary development through relevant courses in every semester in order to ensure that the students could receive enough exposure that would help them acquire 3000-word families.

Two recent studies conducted by Siregar (2020a, 2020b) investigated the mastery of word knowledge among undergraduate English language learners in Indonesian universities. Using Vocabulary Size Test (Nation and Beglar, 2007) and Vocabulary Level Test (Webb, Sasao and Balance, 2017) as the instruments of measurement, the research findings indicated that most of the students had limited grasp of high- and mid-frequency vocabulary. In particular, Siregar (2020a) concluded that the vocabulary size of as many as 92.5% of the students ranged from 6000 to 15400 word families, with an average size of 8732.5 word families. Although this large vocabulary size would allow the students to engage with reading materials containing 8000-9000 word families, the result furthermore revealed that few students had a proper understanding of mid- and high-frequency words. Since only ten out of forty participants (25%) had mastered words at the mid- and high-frequency levels, a more extensive approach to reading comprehension is deemed necessary vocabulary learning. The other study of Siregar (2020b) involving a higher number of participants enrolled in EAP subjects at two private universities concluded similarly, suggesting that the entire participants of the study had low knowledge of words ranging in high- and mid-frequency. It was observed that an overwhelming majority of students were incapable of understanding 1000-5000 word levels despite having a big vocabulary size. This finding pointed out the obstacles that the students might face as they failed to master the most common words from the first to the third 1000 word families. Not only was vocabulary knowledge inadequacy challenging for the students, language instructors had also expressed concerns regarding the issue since the current consensus has established that a lexical coverage of over 98% is crucial for comprehending a variety of texts. Another recent study by Hartono and Prima (2021) investigated the vocabulary level of 168 first-year students from a private university in Jakarta. The participants were not majoring in English, yet English was used as the medium of instruction in the courses. Findings from the research showed that only 20% of the participants can pass all levels. Even though the study gave insight on how acquisition of vocabulary impacted students’ achievements on IELTS, the fact that the students were not English major students nor those studying English for general or specific purposes, might not give a conclusive description of the students’ low mastery of vocabulary and their English proficiency. Furthermore, as the study only involve participants who were in the first semester, it did not provide information on how much the participants’ English proficiency has improved when they have experienced interventions in their English learning.

Even though there have been previous studies on vocabulary size and level in Indonesia, they are still limited. Thus, there is still room to conduct more research to investigate students’ vocabulary size and level. More similar studies are needed to know whether the increase of the average vocabulary size will still become a trend or not. Following a number of studies that have devised some tests to
determine English learners’ vocabulary knowledge (e.g., Mclean & Kramer, 2015; Nation, 1993; Nation & Beglar, 2007; Schmitt et al., 2001), in this paper, two tests are going to be the focus. The first is Vocabulary Size Test (VST) which was created by Nation and Beglar (2007). The test aims at measuring the learners’ vocabulary knowledge on words required for reading. The scoring of the vocabulary will be based on the new vocabulary level test (VLT) formulated by Webb et al. (2017), with the goal of measuring students’ receptive knowledge of higher frequency words (1000-5000).

Most previous studies on vocabulary size and level took freshmen as participants upon their entry to the university. There has been no study yet that investigated three different batches of students for two times in one year of their studies or in two successive semesters. Thus, in order to gain deeper insight on students’ vocabulary size and level, the study included students from three different batches. Moreover, the data were taken from two semesters, the odd and even semesters in the academic year 2020/2021. Thus, it enabled this study to obtain information on how much the students have progressed in the end of the academic year. It is also worth noting that other studies have involved a comparison of data among students of different batches, however, their focus was more on students’ vocabulary size and not vocabulary level. Therefore, it is expected that this current study can provide further information on how far English vocabulary level can improve in one year of study regarding both vocabulary size and level.

Using the aforementioned measurements, the present study has the goal to know the vocabulary size and the vocabulary level of undergraduate students who were majoring in English in a private tertiary institution in West Java, Indonesia, to find answers to the following research questions:
1. How large are the EFL undergraduate students’ vocabulary sizes?
2. To what extent have the EFL undergraduate students mastered 1000-5000 word-levels?

RESEARCH METHOD

The study adopted a quantitative research approach. Thus, the study focused its analysis on statistical data that was gathered for the research descriptions. Since the data was calculated by a computer, the researchers can save time in data gathering and calculation. Consequently, researchers could invest time to describe the result of the study. As mentioned in the previous part, this current study involved undergraduate EFL students from a private tertiary institution. They were students who were from the following batches: 2018, 2019, and 2020.

In this study students completed two vocabulary tests in each semester. First, they completed the vocabulary size test (VST) of Nation and Beglar (2007) 14000 or 20000 versions. Then, the students did the vocabulary level test (VLT) of Webb et al. (2017), after which the result was analyzed quantitatively. The former was used because it was the most useful vocabulary size test and the latter was chosen since it was the latest and most updated vocabulary test. The former was also used in the previous recent studies of Siregar (2020, 2021) and the latest was employed in the investigation of Siregar (2020, 2021) and Hartono and Prima (2021). Employing the same tests which are still reliable to be used is essential when
An Investigation of EFL Students' Vocabulary Size – Fenty & Henni (305-315) 310

comparing the result of this study with the result of the previous studies. The study followed the recommendation of Nation and Beglar (2007) when counting the VST result: the students’ correct answers were multiplied by 100. For instance, a learner who correctly answered 60 questions, would get their vocabulary size stated as 6000 words. For the second test or VLT, the analysis followed the recommendation of Webb et al. (2017). Therefore, the cutting point when calculating whether or not the students have passed 1000 to 3000 word-levels was set at 97%. The 97% mastery equals 27 correct answers out of 30 questions. The vocabulary mastery of 4000 and 5000 word-levels was set at 80% which means that students have to answer 24 questions correctly.

RESULT AND DISCUSSION

Result

Table 1 presents the results of a total of 73 English major students from three batches who participated in the Vocabulary Size Test. The mean scores acquired showed that the students from all batches increased their vocabulary size after one year of study. However, there was a decrease in the highest score achieved by students in batches 2018 and 2020; also in the lowest score of the student in batch 2020. The gap in the standard deviation from before and after one year of study was quite big, especially for batch 2020. This indicates that there is a wide range of proficiency among the participants.

Table 1. Vocabulary Size 2020 and 2021

<table>
<thead>
<tr>
<th>Year</th>
<th>2020</th>
<th>2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>N total</td>
<td>25</td>
<td>21</td>
</tr>
<tr>
<td>Highest</td>
<td>16400</td>
<td>10000</td>
</tr>
<tr>
<td>Lowest</td>
<td>7400</td>
<td>8000</td>
</tr>
<tr>
<td>Mean</td>
<td>11745</td>
<td>8800</td>
</tr>
<tr>
<td>SD</td>
<td>2899.78</td>
<td>883.18</td>
</tr>
</tbody>
</table>

Table 2 shows the findings of students’ vocabulary level. Broadly, it can be observed from the table that after one year of study there were more students who passed 1000-5000 word-levels. As previously mentioned, the calculating point for mastering 1000 to 3000 word-levels was set at 97% while the mastery of 4000 and 5000 word-levels was set at 80%. Thus, the data shows that after one year of study, the average score of participants from batch 2020 has passed the cutting point for 1000-word level with an increase of 12%. This is the highest improvement among the three batches after one year of the study. Despite the students’ vocabulary development, it is noticeable that after three years of study, the batch average score for 3000 word level is still disappointing. The mean score shows that averagely students have not passed that level for all batches.

Moreover, there are still 48% of the students who have not mastered the first 5000 word-level. Some students also have not mastered the 1000-2000 words. Only all students of Batch 2019 passed 1000 word-level after one year of study.
Table 2. Vocabulary Level 2020 and 2021

<table>
<thead>
<tr>
<th>Year</th>
<th>Batch 2018</th>
<th>2020</th>
<th>2020</th>
<th>2018</th>
<th>2019</th>
<th>2020</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>25</td>
<td>21</td>
<td>27</td>
<td>22</td>
<td>21</td>
<td>27</td>
</tr>
<tr>
<td>1000 L</td>
<td>Highest</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Lowest</td>
<td>93</td>
<td>97</td>
<td>80</td>
<td>93</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>98.80</td>
<td>99.71</td>
<td>95.19</td>
<td>98.80</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.40</td>
<td>0.91</td>
<td>5.72</td>
<td>2.33</td>
<td>0</td>
</tr>
<tr>
<td>2000 L</td>
<td>Highest</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Lowest</td>
<td>87</td>
<td>80</td>
<td>67</td>
<td>70</td>
<td>60</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>96.20</td>
<td>96.43</td>
<td>93.07</td>
<td>91.75</td>
<td>92.96</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>4.55</td>
<td>6.31</td>
<td>8.16</td>
<td>9.98</td>
<td>9.07</td>
</tr>
<tr>
<td>3000 L</td>
<td>Highest</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Lowest</td>
<td>70</td>
<td>57</td>
<td>17</td>
<td>67</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>89.76</td>
<td>88.53</td>
<td>77.65</td>
<td>92</td>
<td>90.79</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>8.90</td>
<td>11.05</td>
<td>20.34</td>
<td>9.23</td>
<td>13.94</td>
</tr>
<tr>
<td>4000 L</td>
<td>Highest</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Lowest</td>
<td>63</td>
<td>57</td>
<td>13</td>
<td>67</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>87.80</td>
<td>87.71</td>
<td>76.79</td>
<td>92</td>
<td>90.79</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>10.63</td>
<td>12.73</td>
<td>23.80</td>
<td>9.23</td>
<td>13.94</td>
</tr>
<tr>
<td>5000 L</td>
<td>Highest</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Lowest</td>
<td>63</td>
<td>40</td>
<td>23</td>
<td>47</td>
<td>43</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>86.40</td>
<td>84.24</td>
<td>68.77</td>
<td>91.20</td>
<td>90.48</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>10.64</td>
<td>14.45</td>
<td>23.30</td>
<td>12.76</td>
<td>12.03</td>
</tr>
<tr>
<td>NS (pass all Cutting points)</td>
<td>10 (40%)</td>
<td>8</td>
<td>5</td>
<td>13</td>
<td>9</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>(38%)</td>
<td>(19%)</td>
<td>(52%)</td>
<td>(43%)</td>
<td>(30%)</td>
<td></td>
</tr>
</tbody>
</table>

Discussion

The mean scores from the Vocabulary Size Test (Table 1) show that in the first semester, the average vocabulary size is about 10,000. This confirms what previous studies has found that on average, students' vocabulary size increases on more recent studies (Siregar, 2020a; Kusumarasdyati & Ramadhani, 2018; Nurweni & Read, 1999; Romadloni, 2019; Umam, 2016). Table 1 also shows that the vocabulary size of participants from batch 2019 is the lowest. The fact that students from batch 2020 have quite a big score in their vocabulary size test might give an initial impression that compared to their senior, students from batch 2020 have acquired better proficiency when they enter their study. However, in the second semester, the mean scores show an increase in the three batches, with the most significant increase in the result of participants from batch 2019. This suggests that it is essential for the teachers to review and improve the English courses as well as the teaching and learning process for other batches 2018 dan 2020 - so that their progress will be as high as that from batch 2019.

With regard to Vocabulary Level Test, the result shows that all students from batch 2019 passed the 1000 word level cluster in their second term; however
when it comes to 2000 word level, some students were still struggling. Similar findings were also found in previous studies done by Sudarman and Chinokul (2018) in which some of the participants failed to acquire the 1000-2000 word levels. In the current analysis, the students who had issues in mastering the 1000-2000 word families were from batch 2019 and 2020, who had only studied at the English Department for one or two years. Although the length of study could affect the students’ vocabulary level, it is expected that the students, especially those majoring in the English Department, are able to master the 1000-2000 words since such ability is crucial in understanding English, both spoken and written (Webb, Sasao & Ballance, 2017). On the other hand, a steady improvement can be seen in the number of students who passed cutting points. Within a year, the vocabulary level of students increases by 9% on average. It can be argued that continuous exposure and study demand helped enhance the students’ lexical competence. The exposure also should incorporate the 3000-word level which students on average have not yet mastered. We agree with Kirana and Basthomi (2020) who suggest promoting vocabulary development through relevant courses in every semester so that the students could receive enough exposure that would help them acquire 3000-word families. This highlights the crucial role of teachers in making sure the exposure present as Dang (2020) argues that learners tend to learn vocabulary that their teachers expose to them. In other words, teachers’ well-planned vocabulary exposure can significantly enlarge students’ knowledge of high frequency words. On top of that, as Nation (2012) argues, engaging students in reading graded readers and in deliberate teaching and learning processes is the vital key for high frequency word mastery. Thus, it is important that the students’ extensive reading activity which involves graded readers should be encouraged even after they pass their extensive reading courses.

It is also interesting to note that although on average students’ vocabulary size increase within their one year of study, the students with the highest scores from batch 2018 and 2020 show a decline in their achievements. The student from batch 2020 with the lowest achievement in 2020 also shows a decline in 2021. Moreover, when the mean is compared, the increase in the number of vocabulary is lowest for the batch 2020. This might indicate that in 2020 the students might not be exposed to English as much as the previous batches. The students in batch 2020 are those who start their tertiary study in the pandemic era. Thus, it can be assumed that online learning might somehow influence the students’ achievement in developing their vocabulary. The study of Amiruddin and Jannah (2021) showed a similar trend. Their study involved a sample of English major students in Madura University. They found that students who were not engaged in online learning during the covid-19 pandemic showed better proficiency than those who had experienced online learning.

However, the students’ increase in their vocabulary level shows that even in the 1000-word level not 100% of the students have mastered the level after one year of study. Even though students from batch 2020 show the highest increase in the mean, the achievement of the student with the lowest score is the least among the others. On the other hand, the finding in 2000-word level shows that students
from batch 2020 had increased their level on average. This is important to note as the students from the other two batches show an opposite trend. Therefore, ensuring that the students are consistently exposed to high frequency words during their study is important, no matter what semester they are in.

CONCLUSION

In summary, the results of this research project found that generally students have a better mean score in both vocabulary size and level tests and more students passed the cutting points of 1000-5000 words after studying for two semesters. In spite of the improvement, each batch's top score declined, and batch 2020's standard deviation was particularly large before and after a year of study, showing a wide range of proficiency among the participants. Also, some students still failed to understand the 1000–2000-word level. Since vocabulary mastery is crucial for children's language proficiency, these students need a lot of support in order to increase their vocabulary. Specifically, for those who have not yet mastered the most frequent words in English, it is advisable to encourage them to read graded reader books extensively. The study only conducted the vocabulary size and level tests twice in 2020 and in 2021 and did not include too many students. Therefore, further study can replicate this study with a bigger number of students to get a clear picture of students' vocabulary development which this study cannot provide.

ACKNOWLEDGMENTS

The researchers would like to thank Universitas Kristen Maranatha (Maranatha Christian University), especially the Institution of Research and Community Service (LPPM) at the University, for supporting and funding this research project.

REFERENCES


