A STUDY ON ENGLISH VOCABULARY SIZE OF STUDENTS MAJORING IN INTERNATIONAL NATURAL SCIENCE EDUCATION, YOGYAKARTA STATE UNIVERSITY

by

Syafrianto SM., SS., M.Hum.
STMIK EL RAHMA Yogyakarta

Maria Wulandari, S.Pd., M.Hum.
IAIN Surakarta

ABSTRACT

The current study is an attempt to discover the vocabulary size of students majoring in International Natural Science Education and the developmental pattern of their vocabulary size. The proposed research questions are: (1) What are the vocabulary sizes of students of international natural science education?, and (2) Is their vocabulary size different across levels (freshmen, sophomore and junior)?

To achieve the research aim, a cross-sectional research design was used. The total number of 79 students from three semesters studying in International Natural Science Education participated in the study. The instrument used was a Vocabulary Size Test developed by Nation and Beglar (2007). A descriptive statistics and an inferential statistics One-Way ANOVA continued by a post-hoc test (Scheffe) were run in order to answer the proposed research questions.

The results of computation revealed that the vocabulary size was 5,271 word families for the freshmen level, 6,142 word families for the sophomore level and 6,524 word families for the junior level. Based on the result of inferential statistics One-Way ANOVA, the students vocabulary size across levels are different. The F-ratio indicates that the difference between these groups was 7,369 in the level of significance 0,001. Furthermore, based on the Scheffe test, the significant difference lies between semester 2 and semester 4. Meanwhile, the difference of the vocabulary size between semester 4 and semester 6 was found to be less significant. In general, it can be concluded that there is a difference on the development of students vocabulary size reflected by the result of statistical analysis. The developmental pattern of students’ vocabulary size seems to be positive since the higher the level, the larger the students’ vocabulary size.

Keywords: vocabulary size, international natural science education, developmental pattern

A. INTRODUCTION

Vocabulary is the group of words that a person or a group of people knows how to use (Ansarin, Zohrabi and Zeynali, 2012: 1841). Since vocabulary is a general term for all languages, in this study, the researchers will only refer to the vocabulary in English language. Vocabulary is a core component of language proficiency and provides much basis for how well learners speak, listen, read and write (Richards and Renandya, 2002: 255). In the same light, Nation (1993: 131) states that vocabulary is an essential prerequisite to the development of the skill in language use. Hence, knowledge of vocabulary is inevitably important. There are two types of vocabulary knowledge. They are vocabulary breadth and vocabulary depth (Read, 2000). One of the most important aspect that has been drawn the researchers’ attention is vocabulary breadth or vocabulary size. Some studies on vocabulary size and vocabulary acquisition have been conducted in the English language program. For example a study by Ibrahim, et al. (2013: 44) who investigated EFL learners vocabulary size in the pre-university intensive English programme in International Islamic University Malaysia found that there was a strong relationship between language proficiency and vocabulary size. Another study about vocabulary size was conducted by Ansarin, Zohrabi and Zeynali (2012: 1841). They found that advance level learners have higher vocabulary size in relation to other proficiency levels.
However, only few studies were available concerning the study of vocabulary size in the field of non-English language program. Non-English language program can be defined as an international program or class studying specific discipline such as mathematics, biology, natural science, and chemistry using English as the medium of communication during the teaching and learning process (bilingual). The international class has been grown in the recent time in some universities as a response to the demand of international schools teachers. If it is compared with the regular class, students of international class get more opportunity to learning English and acquiring vocabulary. Having a large vocabulary size is important for students of international because it will help them in reading books, journal and other sources in English more easily. Anderson and Freebody (1981 in Eyckmans, 2004: 14) stated that vocabulary size was found to be a good predictor of reading comprehension. In relation to the amount of the vocabulary students can master, it might be related to the availability of sufficient exposure in the environment. Besides that, it is also influenced by ways of learning the vocabulary itself. In fact, vocabulary can be learnt implicitly and explicitly. Which technique is effective is based on the students individual differences and other possible factors.

For international class students, the higher their level (semester), the more books or academic texts they should read in their fields. It means that along with the increase of their levels they are expected to have larger vocabulary in order to read more comprehensively. Therefore, the researchers was interested in conducting the study about the vocabulary size of students of international class across levels. The sample used is the international class of natural and science education in Yogyakarta State University. This study was intended to measure the students’ vocabulary size and its developmental pattern. The benefit of measuring vocabulary size and finding out the developmental pattern is for learners to be more aware with their knowledge of English vocabulary and to improve it in order to become more efficient and effective in their fields. Besides that, for the institution, it is useful as a part of the evaluation component in the international class offered by the institution.

The research questions proposed in this study are two-fold: (1) What is the vocabulary sizes of students of international natural science education in each level?, and (2) Is their vocabulary size different across levels? Concerning the second research question, the hypothesis formulated in this study is: “the students vocabulary size is different across levels.”

B. LITERATURE REVIEW

Vocabulary Acquisition and Development

Vocabulary acquisition of L2 is a developmental process. Bogaards and Laufer (2004: 103) stated that the process of getting to know the meaning of new words can be divided into two stages. The first one is comprehension stage. This is the initial understanding of a word’s meaning or the initial mapping of new word forms to existing meanings or concepts in the learner’s mind. This includes a critical part of word retention and production because a word without meaning is not likely to be retained for long-term use. The other is the gradual developmental process or the development stage. In the development stage, the learners face the continuous task of checking the original semantic content of a word against meaning of the words as shown in different contexts for match and mismatch. Nowadays, across countries, there is an increasing awareness of the importance of vocabulary development in the learning of second language, as well as in improving communication skills (Milton, 2009 in Biswas, 2013: 6). Some researches have been conducted in relation to the vocabulary development. One of the notable findings from previous research is about the importance of instructional intervention in vocabulary learning which was found to be useful for assisting learners to acquire vocabulary knowledge. It was emphasized that natural exposure alone may not provide enough impetus for semantic restructuring and development (Bogaards and Laufer (2004: 121).

Vocabulary Knowledge

Nation (1990: 31-33) divides the vocabulary knowledge into two types. Receptive vocabulary knowledge and productive vocabulary knowledge. Receptive knowledge involves the ability to recognize the word when it is heard or when it is seen. Meanwhile, productive knowledge includes receptive knowledge and extends it. Meaning to say, it also involves knowing how to pronounce the word, how to write and spell it, how to use it in correct grammatical patterns along with the words it usually collocates with. Another distinction has also often been made in relation to vocabulary
knowledge: vocabulary depth and vocabulary breadth or size. Vocabulary depth is the quality of the vocabulary knowledge, namely how well a particular word is known. Meanwhile, vocabulary size is related to the quantity. It means the number of words of which the learner knows at least some significant aspects of the meaning (Eyckmans, 2004: 13). In this study, the researchers focuses on vocabulary size.

**Vocabulary Size Test**

To discover the vocabulary size, administering a test is necessary. The test is usually called vocabulary size test. Based on some literature, vocabulary size test has various purposes. It can be used to see if the learners know sufficient words to perform certain tasks, chart the growth of the learners’ vocabulary, evaluate the extent to which a certain program meets its objectives, choose the right instructional path for learners, and design relevant curricula and course materials and so on (Beglar, 2010). In this study, the purpose is focused on illustrating the growth of the learners’ vocabulary size. It is to look at the development of the international natural science education students in terms of their English vocabulary size. Referring to the types of vocabulary knowledge, there are also two types of measurement of vocabulary size, i.e. receptive and productive vocabulary size. The researchers decided to measure the receptive vocabulary size only since it was considered more feasible to the context. The consideration is that because of the students’ short-term goal in mastering vocabulary knowledge, that is for reading academic texts. It has been argued and empirically stated by many researchers that the mastery of vocabulary size has strong relation with the reading success in L2 (Stæhr, 2008: 140). Therefore, receptive vocabulary size is selected to be measured by the researchers. It is to know whether their vocabulary is developed significantly considering the increase demand of the students’ need for reading academic texts in English.

In measuring vocabulary size, usually the learners are asked to match each word with a synonym, a dictionary-type definition or an equivalent word in their own language (Read, 2000: 16). Some instruments to measure vocabulary size have been developed by the experts such as the Vocabulary Levels Test (Nation, 1983; Beglar and Hunt, 1999), the Eurocentres Vocabulary Size Test (Meara and Jones, 1990) and the Vocabulary Size Test (Nation and Beglar, 2007). However, for the purpose of this research, the researchers used Vocabulary Size Test (VST) by Nation and Beglar (2007). The researchers selected this instrument to gather the data because this is the most recent instrument. Besides that, the VST has been validated and well-designed. Also, it is considered practical in administering and scoring because it is in the form of multiple choice questions. Therefore, in relation to the limitation in this research (time limitation to develop new instrument and to administer the test to the students), the VST was believed to be appropriate considering the criteria of validity, reliability and practicality like what have been stated by Nation (2001).

**C. METHODS**

**Research Design**

This study used survey research design. Survey research is the procedure in quantitative research in which investigators administer a survey to a sample or to the entire population of people to describe the characteristics of the population (Creswell, 2012: 376). Best (1981: 20) stated that survey research is not concerned with the characteristics of individuals as individuals. It concerns with the generalized statistics when data are abstracted from a number of individual cases. The type of survey research design adapted for this study is cross-sectional. Cross-sectional design is a type of survey research which is used to collect data at one point in time. It compares two or more educational groups (Creswell, 2012: 377-378). There are three groups compared in terms of their vocabulary size: freshmen, sophomore and junior.

**Instrument**

In order to assess the participants’ vocabulary size, Vocabulary Size Test developed by Nation and Beglar was used. The vocabulary size test is a measure of written receptive word knowledge. It is based on word family frequency estimates derived from British National Corpus (BNC) (Nation, 2006 in Karami, 2012: 54). The vocabulary size test includes 14,000 word families. That is, it is generally assumed that the first 1,000 level contains the items most likely to be known and the 14,000 includes those less likely. A word family is defined as a headword plus the inflected and closely related
derived forms. A requirement for the word family is in fact that the headword should be a free form (Karami, 2012: 54). In the test paper, each 1000 word families is represented by 10 items. Therefore, there are 140 multiple-choice items in the test. However, considering the time limitation provided by the institution in which the vocabulary size test was conducted, the researchers only adopted 100 multiple-choice items for the test. Concerning the validity of the test, there has been a strong evidence from Beglar (2010). He argued that the items in Vocabulary Size Test have been carefully written and the distractors function well. The Vocabulary Size Test items used in this research can be seen in Appendix.

Setting and Participant
The data was collected in Faculty of Mathematics and Natural Science, Yogyakarta State University especially in International Science Education Class. In order to gather data, 79 students across three levels, i.e. freshmen, sophomore and junior (semester 2, semester 4 and semester 6) were involved. The students from senior level were not possible to be involved in the study because of they did not have any regular class anymore so that it would be difficult to administer the test to them. Their ages ranged approximately between 18-21 years old. All of them spoke Indonesian as their first language. They were studying for their bachelor degree in international natural science education. International natural science education is a study program which offers students to develop knowledge and skill concerning their discipline as well as to prepare them in facing challenges of teaching in some international schools. At the first time before selected to be students of international natural science education, they were tested in order to measure their English proficiency. In addition, when they started their first up to fourth semester, they had taken an English for Specific Purposes as their general subjects.

Research Procedure
The researchers followed a systematic procedure in conducting the research. The procedure are described as follows. First, the researchers determine the topic of the research, read some books and journals, linked it with the local context and found the gap. From the gap, the researchers then tried to find a new research topic and then formulated the research questions. Hypothesis was formulated, the methodology was selected according to the aim of the research, then the data collection process was done. The data collection process is started by the activity of selecting the instrument for the vocabulary size test and asking the permission to the institution in which the test would be administered.

The data were collected by the researchers after getting the permission from the Dean of Faculty of Mathematics and Natural Sciences in Yogyakarta State University. As stated previously, the participants were from 3 levels (semesters) of students of International Natural Science Education. A lecturer was chosen by the dean to assist the researchers to conduct the study. The instruments were administered within the time scheduled by the researchers, the lecturer and the students from three levels so that it did not disrupt their regular class schedule. The test was administered separately among the students of each level. The students of semester 4 did the test on May 20, 2014 at 10 a.m. and the students of semester 6 did the test on the same day at 11 a.m. Meanwhile, the students of semester 2 did the test on the other day, that is on May 26, 2014 at 7 a.m. This is due to the tight schedule of the students since it was close to the end of the effective class hours in the academic calendar. In the test, the students were given time limitation of 45 minutes. They were not allowed to look at the dictionary and they should answer the question in the provided paper.

After the data collected, the researchers then computed the data to get the total score of students vocabulary size. The procedure for scoring the test was based on Paul Nation (2012) which states that a learner’s total score needs to be multiplied by 100 to find the learner’s total vocabulary size. After that, the researchers conduct statistical analysis using SPSS software in order to test the hypothesis of difference or to answer the research questions. The result of statistical analysis, then was interpreted and presented in the research report. Finally, the conclusion was derived from the findings and discussion.

Data analysis
The data gathered (students’ test results) were calculated manually to obtained the vocabulary size test score (the total of the correct answer). Then, the scores of every student were entered numerically to
the SPSS 17.0 software. The 79 name of the participants and their semester were entered as nominal scale while their scores were entered in SPSS software as interval scale. To answer the proposed research questions, there are two main statistical analysis conducted. To answer the first question (what are the vocabulary size of students across levels/semesters?), the researchers conducted descriptive statistics analysis in order to describe the data obtained from the vocabulary size test. Meanwhile, to answer the second research question (is the students’ vocabulary size different across levels?), the researchers conducted statistical analysis using One-Way Anova because the researchers would like to discover whether the variance of the means score is different among the three groups which reflect the existence of development of students’ vocabulary size. In addition, Scheffee test was run to determine in which level the vocabulary size is significantly different. Scheffee test was chosen because the number of sample size is different across levels.

**D. FINDINGS & DISCUSSION**

**The English Vocabulary Size of Students Majoring in International Science Education**

In order to answer research question one, what is the vocabulary sizes of students of international natural science education in each level?, the result of computations consisting of mean, standard deviation, minimum score, maximum score, range and variance are presented as follows.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Range</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshmen</td>
<td>24</td>
<td>35</td>
<td>34</td>
<td>69</td>
<td>52.71</td>
<td>9.787</td>
<td>95.781</td>
</tr>
<tr>
<td>Sophomore</td>
<td>26</td>
<td>37</td>
<td>42</td>
<td>79</td>
<td>61.42</td>
<td>10.856</td>
<td>117.854</td>
</tr>
<tr>
<td>Junior</td>
<td>29</td>
<td>48</td>
<td>37</td>
<td>85</td>
<td>65.24</td>
<td>14.389</td>
<td>207.047</td>
</tr>
</tbody>
</table>

Table 1 shows that the mean score of each level is varied. The mean score of the freshmen level is 52.71 with the standars deviation of 9.787. The mean score of the sophomore level is 61.42 with the standard deviation of 10.856. The mean score of the junior level is 65.24 with the standard deviation of 14.389. Besides, its range for each level are respectively 35, 37, and 48. Table 2 also manifests the minimum and the maximum score of the students’ vocabulary size. The minimum score are respectively 34, 42 and 37. Meanwhile, the maximum score are respectively 69, 79 and 85.

Nevertheless, descriptive statistics alone is not sufficient to determine the students’actual vocabulary size in terms of word families. The scoring rule of the Vocabulary Size Test should be considered as well. According to Nation (2012), a learner’s total score on the Vocabulary Size Test needs to be multiplied by 100 to find the learner’s total vocabulary size. Nation illustrates, a score of 35 means that the learner’s vocabulary size is 3,500 word families. Hence, to find out the average of students’ vocabulary size, the mean score above needs to be multiplied by 100. As a result, the average of students’ vocabulary size in the freshmen level is 5,271 word families, the sophomore level is 6,142 word families and in the junior level is 6,524 word families. The figure can be seen below.
From the figure above, it can be seen slightly that there was a gradual development of the students’ vocabulary size from one level to another. In other words, the increase of the vocabulary size from one level to another is the sign that there were vocabulary growth. If the increase value is calculated, the growth of the vocabulary size from one level to another can be presented as follows.

Table 2. The Growth of the Students’ Vocabulary Size

<table>
<thead>
<tr>
<th>Vocabulary Size</th>
<th>The Growth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freshmen</td>
<td>5,271</td>
</tr>
<tr>
<td>Sophomore</td>
<td>6,142</td>
</tr>
<tr>
<td>Junior</td>
<td>6,524</td>
</tr>
<tr>
<td>Total Growth</td>
<td></td>
</tr>
</tbody>
</table>

The vocabulary size growth from the freshmen level to sophomore level was 871 word families. Meanwhile, the vocabulary size growth from the sophomore to the junior level was 382 word families. The overall growth was 1,253 word families. From Table 2, it can be concluded that the relatively significant growth lies from the freshmen level to the sophomore level. On the other hand, the growth from the sophomore to junior level was found to be smaller than that of freshmen to sophomore level.

The Difference of English Vocabulary Size of Students Majoring in International Science Education

To answer research question two, is the students’ vocabulary size different across levels (freshmen, sophomore and junior)?, analysis of variance was employed. In order to conduct analysis of variance, the data should fulfil some pre-requisites: normality and homogeneity. Hence, the data were tested first. For normality, the data were tested using Kolmogorov-Smirnov Test. Meanwhile, for homogeneity, the data were tested using Levene Test. The result of normality and homogeneity test shows that the data is normally distributed and homogeneous as well (See Appendix). The distribution is normal with the significance value of 0.792 (more than 0.05) and the data is homogeneous with the significance value of 0.029 (more than 0.01). After testing the normality and homogeneity, the statistical analysis ANOVA were run using SPSS 17.0 software. Table 3 shows the result of analysis of variance (ANOVA).
In One-Way Anova, F value and Significance (Sig.) value is important. F value in the table above is usually called F-observed. If the F-observed is larger than F-table, it means that there is a significant difference between groups compared in this test. On the other hand, if F-observed is smaller than F-table, then there is no difference between groups being compared. From the table above, the F value is 7.369. Meanwhile, the F table after being checked in ANOVA table is 3.117. The F observed (7.369) is larger than the F-table (3.117). Hence, it can be concluded that there is a significant difference between groups being compared. This is also supported by the significance value of 0.001. It is smaller than 0.05 indicating that the difference is significant. However, it is still unclear where the differences lie. Therefore, the statistical analysis then was continued to a pos-hoc analysis employing Scheffee test multiple comparisons considering that the sample sizes are not equal. Here is the result of Scheffee test.

Table 4 shows the multiple comparisons between groups. The result of comparisons indicates that the difference between the groups were not the same. The significant difference lies between the freshmen and the sophomore level and between the freshmen and the junior level. However, there is less significant difference between the sophomore and the junior level. It can also be seen in the Figure 2 below. Figure 2 below illustrates the means plot of the students’ vocabulary sizes among the three levels.
To conclude, the findings show that there was a difference of the students’ vocabulary size which reflects the existence of the vocabulary development among the students across levels. The trend of the development is positive since from the statistics it is known that the higher the students level, the larger their vocabulary size. However, the significant difference which also reflect the significant development lies between the freshmen level and the sophomore level. Meanwhile, the line from the sophomore level up to the junior level seems to show less significant development because the difference is not really significant based on the Scheffe test.

As it has been stated previously, the finding shows that there is a sort of development on the students’ vocabulary size in international natural science education. The hypothesis formulated at the beginning has been accepted. The null hypothesis which states that that there is no difference of students’ vocabulary size across level is rejected. The students’ vocabulary size increases along with their length of study as shown by the result of data analysis. From this finding, it can be inferred that length of study may have an influence to the students’ vocabulary size. Students of semester 6 might perform better in the vocabulary size test since they have longer time involving in learning English vocabulary through the international class environment.

In addition to the finding discussed above, there is another interesting finding to be discussed in this study, that is the development of students vocabulary size from sophomore to junior level which was found to be less significant than the development of students vocabulary size from the freshmen to sophomore level. This is probably related to the exposure and the focus of their learning. Meaning to say, the students of international natural science education were exposed to English for Specific Purposes subject during semester 1 up to 4 taught by the native speakers. Therefore, rapid development occurs among the students of freshmen level up to the sophomore level. This is different from the development of vocabulary size from sophomore level up to the junior level. As shown in Figure 2, the line represents less significant development of students vocabulary size from sophomore level up to the junior level. It is probably due to the less exposure and input to learning English with formal instruction like English for Specific Purposes as they got in semester 1 up to 4.

Those findings might confirm some theories and previous studies that have been done by other researchers. Ellis (2008: 101) has summarized some studies about how learner’s lexicon develop quantitatively. He comes to a conclusion that over time, given input, learners will learn more words. So, some possible factors in the increase or development of vocabulary acquisition is the number of years the learners learn and the input. The more time the students have to learn vocabulary, the larger their vocabulary size, provided that there are sufficient input of target language in the environment.
In the study of developmental pattern, there might be various factors influencing the result. This study is only focused to discover the developmental pattern. Findings regarding the factors affecting the vocabulary size development above is still necessary to be confirmed by conducting further research. Meaning to say, it is still unclear enough whether the development pattern as reflected in the data analysis is a consequences of instruction (the existence or the inexistence of ESP subject in international natural science education), or whether it reflects the nature of lexical learning, or probably other factors.

In relation to the implications, this study about developmental pattern of vocabulary size is considered useful since it can be an input or consideration for the students to increase their vocabulary knowledge especially their vocabulary size because they are usually exposed to various English reading texts in their fields. International class students are usually required to read books and journals in English. Knowledge of vocabulary is important to understand reading text more easily and comprehensively. Nation (1993: 115) states that good vocabulary knowledge enables good comprehension. Some studies have also shown that a strong relationship exist between knowledge of word meaning and ability to comprehend passages containing those words (Anderson and Nagy, 1992 in McLaughlin, et al. 2000: 131). In another study, Qian (2002 in Rouhi and Negari, 2013: 41) also obtained the result confirming the important role of breadth of vocabulary knowledge in reading comprehension. However, it is not to say that vocabulary size alone is sufficient in reading. Rather, the point is that the adequate knowledge of words is a prerequisite for effective language use (Read, 2000: 83).

E. CONCLUSION

The aim of the study described above is to find out the students’ vocabulary size of international natural science education and to discover if the students vocabulary size different across levels. The result of the study shows that the students the vocabulary size was 5,271 word families for the freshmen level, 6,142 word families for the sophomore level and 6,524 word families for the junior level. Based on the result of inferential statistics One-Way ANOVA, the students vocabulary size across levels are different. The F-ratio indicates that the difference between these groups was 7,369 in the level of significance 0.001. Furthermore, based on the Scheffe test, the significant difference lies between semester 2 and semester 4. Meanwhile, the difference of the vocabulary size between semester 4 and semester 6 was found to be less significant. The possible explanation of this latter finding is that the inexistence of the formal instruction, i.e. English for Specific Purposes class, which is assumed to be the cause of the decrease of the students vocabulary size. In general, it can be concluded that there is a development of students vocabulary size reflected by the result of statistical analysis. The developmental pattern of students’ vocabulary size seems to be positive since the higher the level, the larger the students’ vocabulary size.

In this study, however, it is still unclear enough whether the development pattern as reflected in the data analysis is a consequences of instruction (the existence or the inexistence of ESP subject in international natural science education), or whether it reflects the nature of lexical learning, or probably other factors. Further research is still necessary to discover and to explain about that.
REFERENCES


