

## CHAPTER 2

### LITERATURE REVIEW

This chapter explains the theories of English vocabulary, categorization of vocabulary, receptive and productive vocabulary, vocabulary mastery, vocabulary level, vocabulary level test, and conceptual framework

#### 2.1 Vocabulary

Knowing how many words in English is not easy. In every year, new words come up and use widely. The last phenomenon word is selfie (self photo). According to The Global Language Monitor, there are 1,025,109.8 on January 1, 2014. *Webster's Third New International Dictionary*, Unabridged, together with its addenda section, includes some 470,000 entries. In addition The Second Edition of the 20-volume *Oxford English Dictionary* contains full entries for 171,476 words in current use, and 47,156 obsolete words. To this may be added around 9,500 derivative words included as subentries. Over half of these words are nouns, about a quarter adjectives, and about a seventh verbs; the rest is made up of exclamations, conjunctions, prepositions, suffixes, etc. And these figures don't take account of entries with senses for different word classes (such as noun and adjective).

Despite such difficulties, researchers have tried to estimate how many words native speakers know in order to assess the number of words learners need to learn. Estimates for native speakers vary between 12,000 and 20,000 depending on their level of education. One estimate is that a native speaker university

graduate knows about 20,000 word families (Goulden, Nation, and Read, 1990). Nation and Waring (1999) explained that the second language learners need to know the 3,000 or so high frequency words of language.

Vocabulary is important for reading achievement in the university level due to the increasing demands of the texts students are required to read in their specific area. As being stated by Wiwat from Bangkok University, for many activities and productives skills, like speaking and writing, if leaners don't have an adequate number of words, they can't convey their message (Wiwat,2011).

Vocabulary is the core component of language proficiency and provides much of the basis for how well leaners speak, listen, read, and write. This idea supported by Nation (2001). Laufer (1989) suggested that an amount of vocabulary needed in order to comprehend a text, saying that "learners would need at least 95% coverage of the running words in the input in order to gain reasonable comprehension and to have reasonable sucesss at guessing from context. Therefore, students will be able to use vocabulary effciently if they know enough words.

### **2.1.1 Categorization of Vocabulary**

Vocabulary is usually divided into two, three, four, and even more levels, it can be divided by how often a word (words) can occur (frequency) and how widely it occurs (range). Nation (2012) divided frequency words into three levels, there are High-Frequency words, Mid-Frequency words and Low-frequency words. Nation had developed vocabulary test from 1,000 words level, 2,000

words level, 3,000 words level, 5,000 words level (the university word level), and 10,000 (the expert of non-native speaker) and last academic words.

A study by Laufer (1998) distinguished the knowledge on vocabulary into three types, they are: (1) receptive knowledge, that is the knowledge to understand a word that is used most frequently; (2) controlled productive knowledge, that is the knowledge to recall words using hints, and (3) free productive knowledge, that is the knowledge to produce vocabulary through various activities such as writing assignments.

Some researchers have found out that the success of explicit teaching of vocabulary was greatly influenced by breaking words into their component parts. Emadin & Moghadam (2007) and Tabtimasai, (2003) concluded, the important technique for building vocabulary is breaking words into their component, and using vocabulary learning exercises. The other technique, as being mentioned by Yuen (2009), teaching students how to use context clues while reading improves their understanding of new vocabulary words. Summat (2009) said that using task-based activities can improve student's vocabulary learning and their confidence. These results were found to be successful in enhancing students' vocabulary knowledge.

The research conducted by Wiwat (2011) concluded, students in each class (nine faculties including Humanities, Business Administration, Accounting, Communication Arts, Fine and Applied Arts, Sciences, and Engineering) at the Bangkok University, have different levels of language proficiency. During their

first year at the Bangkok University, the freshmen are in a big trouble when they are assigned to read more in English. Wiwat (2011) said, a limited vocabulary makes them encounter the problem of text comprehension. If they have more vocabulary knowledge, they are likely to be more successful in their learning. Accordingly, providing students with more vocabulary is the most appropriate method to help them improve English vocabulary skill.

As being stated by Wiwat (2011), the researcher belief that for most students any one particular technique of learning vocabulary may not prove efficient enough. A combination of certain techniques might enables students to better acquire and retain words unknown to them.

Yang (2009) concluded in research that (a) college English teachers should keep the students better informed of the significance and specific functioning of context clues in contextual guessing, and (b) college English teachers should encourage the students to guess word meaning from context instead of inhibiting it when there are adequate context clues offered. But, Yang (2000) stated, since a word may have several meanings, the use of context clues will prove effective only when a general meaning of a word is adequate. When the surrounding words are not familiar, when the definite is required, or when the word is question as a key word, using context clues might not be good idea.

### **2.1.2 Receptive Vocabulary and Productive Vocabulary**

Nation (2000) stated that Receptive carries the idea that we receive language input from others through listening or reading and try to comprehend it.

Productive carries the idea that we produce language forms by speaking and writing to convey messages to others. Essentially, receptive vocabulary use involves perceiving the form of a word while listening or reading and retrieving its meaning. Productive vocabulary use involves wanting to express a meaning through speaking or writing and retrieving and producing the appropriate spoken or written word form.

## **2.2 Vocabulary Mastery and Vocabulary Learning**

Mastery of academic vocabulary, mainly that have many meanings in different context, will help the students to be able to study new matters and strategic skill (Marzano, 2005). Research by Stahl and Fairbanks (1986) has proven that academic vocabulary, especially the one that specifically related to certain discipline or instruction, can directly increase and improve the students' understanding in reading text book.

Yamamoto (2011: 227) said, scholars (e.g. Horst, Cobb, & Meara, 1998; Nation, 2001, 2009) expressed their view that the step of learning vocabulary by way of reading many books in limited time is not the most effective way to obtain many vocabulary. In this sense, Waring & Nation (2004) gave the example of ineffectiveness of using the strategy of widening reading to increase vocabulary. The example was when a student reads in one hour, he/she would only obtain between 3 to 6 new words. Therefore, according to Waring & Nation (2004), when a student reads for only 30 minutes per day for 13 weeks (equal with a semester at university), he/she is only able to read for 52.5 hours, however the

potency to obtain new words is only 157 to 315 words. Moreover, continued Waring & Nation (2004), after three months, almost 50% of the newly obtained vocabulary through the reading will disappear. It means that only few of the new vocabulary that will be obtained by students when he/she use the strategy of mastering vocabulary through reading.

Then, which one is the most effective way to learn vocabulary in big number? The scholars (e.g. Horst, Cobb, & Meara, 1998; Nation, 2001, 2009), according to Yamamoto (2011: 227) were more confident in believing that the most effective way is to learn vocabulary explicitly that implicitly through reading. In this sense Waring & Nation (2004) described that learning vocabulary explicitly is that the student must find the word minimum eight times in three-month period to obtain 50% opportunity of mastering the word. Moreover, said Waring & Takaki (2003), if the students find the new word only less than five times in three months, the opportunity to master the word is considered none. In a study, Clockwork Orange Study, Sarai, et. al (1978) concluded that the requirement of more frequencies, in which for a new word to be mastered appropriately by a student, the new word shall be found minimum 10 times. Regarding this frequency to find the word, Horst, et. al (1998) as being compared with by Yamamoto (2011) concluded that the word shall be found minimum eight times to be mastered appropriately by the student.

It means that there is correlation between the frequencies of finding new words with the effective way to study vocabulary. While other researcher proposed other formula (other than frequency of finding new word) to be able to

master vocabulary appropriately. Research by Laufer (2003) found that reading alone is less effective in mastering vocabulary than learning word in the view of doing assignment such as answering questions with certain vocabulary that has already been targeted by taskmaster.

Meanwhile, Nation & Ming-Tzu (1999), stated that learning vocabulary with the purpose of exploring a matter would be more effective since the new word assignment will tend to go in-depth as well.

### 2.3 Vocabulary Level

There are theoretically a number of classifications that divide the level of vocabulary capacity. There are some experts that are doing research for grader vocabulary. Nation and Ming (1999) explained level 1-5 are found on fiction and daily vocabulary and level 6-7 are found on non-fiction and academic vocabulary. The distribution is on table 2.1 below.

**Tabel 2.1. Level and Word Families**

<b>LEVEL</b>	<b>WORD FAMILIES</b>
Level 1	500 word families
Level 2	700 word families
Level 3	1.000 word families
Level 4	1.500 word families
Level 5	2.000 word families
Level 6	3.000 word families
Level 7	5.000 word families

Table 2.1. Adapted from Nation and Ming Tzu (1999, p. 377)

Furthermore Nation (2001, 2009) on The Oxford Bookworms series (2001, 2009) stated, which mainly consists of condensed versions of classic novels, it has six levels (Level 1 [400 headwords] to level 6 [2,500 headwords]). According to

Nation (2009), the total word families are roughly the same as the graded readers' level in the Oxford Bookworms series, as shown in the table 2.1 below.

**Table 2.2. Total New Words and Word Families at Each Oxford Bookworms**

Level	New words	Total word families at each level
1	400	400
2	300	700
3	300	1,000
4	400	1,400
5	400	1,800
6	700	2,500

Adapted from: Nation (2009) and Maruyama (2011).

Milton had done many research related on CEFR levels and XLex (orthographic vocabulary) 5,000 max word-level. CEFR levels start from A1 into C2. A1 has XLex <1500 which is tested by Starters, Movers and Flyers from Cambridge. A2 has XLex 1500-2500 which is tested by Kernel English Test from Cambridge. B1 has XLex 2750-3250 which is tested by Preliminary English Test from Cambridge. B2 has XLex 3250-3750 which is tested by First Certificate in English from Cambridge. C1 has XLex 3750-4500 which is tested by Cambridge Advanced English from Cambridge. C2 has XLex 4500-5000 which is tested by Cambridge Proficiency in English from Cambridge. The table is shown below.

**Table 2.3 CEFR Levels and Xlex (5000max)**

CEFR Levels	Cambridge exams	XLex (5000 max)
A1	Starters, Movers, and Flyers	<1500
A2	Kernel English Test	1500-2500
B1	Preliminary English Test	2750-3250
B2	First Certificate in English	3250-3750
C1	Cambridge Advanced English	3750-4500
C2	Cambridge Proficiency in English	4500-5000

Adapted from Approximate vocabulary size scores associated with CEFR Levels

Milton (p. 224).



**Table 2.4 CEFR Levels and Vocabulary Range**

	<b>VOCABULARY RANGE</b>
<b>C2</b>	Has a good command of a very broad lexical repertoire including idiomatic expressions and colloquialisms; shows awareness of connotative levels of meaning
<b>C1</b>	Has a good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions; little obvious searching for expressions or avoidance strategies. Good command of idiomatic expressions and colloquialisms.
<b>B2</b>	Has a good range of vocabulary for matters connected to his/her field and most general topics. Can vary formulation to avoid frequent repetition, but lexical gaps can still cause hesitation and circumlocution.
<b>B1</b>	Has a sufficient vocabulary to express him/herself with some circumlocutions on most topics pertinent to his/her everyday life such as family, hobbies and interests, work, travel, and current events.
<b>A2</b>	Has sufficient vocabulary to conduct routine, everyday transactions involving familiar situations and topics. Has a sufficient vocabulary for the expression of basic communicative needs. Has a sufficient vocabulary for coping with simple survival needs.
<b>A1</b>	Has a basic vocabulary repertoire of isolated words and phrases related to particular concrete situations.

Vocabulary Range Criteria from Council of Europe (2001, p. 112)

The vocabulary range criteria and CEFR levels by Council of Europe (2001) has the values lay in the ability of its users to apply these criteria consistently and accurately but in the absence of more detailed criteria this may be difficult to do in practice. If a student at university possesses many thousand words, including idiomatic and colloquial expressions, and is comparable to a native speaker in his or her foreign language vocabulary knowledge then this

would be a good evidence that he or she would be at C2 level, at least in terms of vocabulary range. A student at university with only a few hundred foreign language words would probably be at level A1 level in terms of vocabulary range and almost inevitably would be much more limited in their skill in using the foreign language.

Other classification of vocabulary level was stated by Laufer, and Waring & Nation. Laufer (1992) stated, in order to read for pleasure without consulting a dictionary, a number of conditions need to be met. First, there is a general consensus that 95% of lexical coverage (one known word every 20 words) or higher is essential for minimal comprehension of reading text. This means that a threshold of 3,000 word families (4,800 lexical items) has to be attained for minimal comprehension. Nation (2001) defined, a word family refers to a “headword, its inflected forms, and its closely related derived forms. For example, “accept”, “acceptability”, “unacceptable”, “accepted”, “accepting”, “accepts,” and “acceptance” are all counted as one word family.

Waring & Nation (2004) stated, to read a short unsimplified novel for pleasure, learners need to know more than 98% of the words in the text (i.e., less than one known word in every 50 running words). According to Hirsh and Nation (1992), having a recommended lexical coverage of 98% translates into knowing roughly about 5,000 word families (8,000 lexical items).

In the level of vocabulary mastery, there is among other “Tier Model” that classifies the level of vocabulary mastery into three categories. The three

categories are tier 1, comprises of Beck, McKeown, and Kucan's, 2000 general words that have basic nature. The Tier 1 is words that are used most often among ESL-type English speakers. Vocabulary in this category is not or only a few instruction. Examples of words in Tier 1 are table, swim, cars, dog, book. The words are related to discernible things (concrete things), functional words, and names of other objects; all of them are included in the category of Tier 1. Tier 2, is the vocabulary that concerned with terms that are used often with many frequencies. These words are important to understand text and are usually used in the curriculum. The examples of words in Tier 2 are analyze, compare, conclusion. The words are often being used on academic setting whether in instruction communication, in discussion forum, in examination, and delivery of assignment by lector. In addition, included in

Tier 2 are the words that have many meanings such as set, bat, base, and check; which all of them have several meanings, so when students use them, they shall link them with the context for the readers or audience to understand. Students can master vocabulary appropriately in general by frequently using the words that included in Tier 2 (Sibold, 2011). Tier 3, vocabulary for usage with more specific content since they are only used in certain fields. As being expressed by Vacca and Vacca (2008) words in Tier 3 are the words with more technical nature in certain area such as centimeter, kilogram, and deciliter in science class and math; or words like abolitionist, emancipation, and secession in history class.

Tier 2 and Tier 3 categories are the category of vocabulary for academic area. Certainly the academic vocabulary tends to be difficult for students to study compared with daily conventional language. This difficulty, according to Sibold (2011: 24) is due to the words are more specific and tend to be more abstract, that make students are more difficult to master. The knowledge that related to specific science disciplin, such as social, science, math, or literary, are directly linked to content of sciences that are more dominant in using Tier 2 to Tier 3 vocabulary (Sibold, 2011).

In addition, Nation (1983, 1990) and also Schmitt (1993) has classified the level of vocabulary capacity as 2,000 words-level, 3,000 words-level, academic words-level, 5,000 words level, and 10,000 words-level (further, see Vocabulary Level Test).

### **2.3.1 Vocabulary Level Test**

The Vocabulary Levels Test was designed to make estimation regarding the measurement of vocabulary mastery for second language (L2) learners of general or academic English.

There are a number of tests to find out the level of vocabulary mastery. Nation (2001, 2009) tested vocabulary mastery using the instrument of series of novel from Oxford Bookworms that have six level of vocabulary mastery. The Level test was categorized on Level 1 (400 headwords) to Level 6 (2.500 headwords), in which, said Nation, the level of vocabulary mastery is equal with vocabulary mastery in the series.

Meanwhile, Nation (1990, 2001), Schmitt & Clapham (2001) stated that there are five sections in the vocabulary mastery level test, they are: (1) the 2,000 word level; (2) the 3,000 word level, (3) the university level. Furthermore Golkar and Yamini (2007) explained this version has 10 clusters and each cluster has six words and three definition. The ratio is 3 (verb) : 2 (noun) : 1 (adjective) in each word level there are 15 nouns, 10 verbs, and 5 adjectives. The test is divided into 150 items, so there are 75 nouns, 50 verbs, and 25 adjectives included in the whole test.

### **2.3.2 Why Testing Vocabulary?**

Historically, The Vocabulary Levels test or frequently called as “Level Test” was originally designed by Paul Nation as a diagnostic vocabulary test for use by lecturers. The vocabulary test was started to emerge in 1983 which was then republished in a book in 1990. Since then, as being illustrated by Schmitt, et al (2001), the vocabulary level test is used internationally, in which there are four versions, version A (original version of Nation), and other three version (version B, C, and D) which were developed by other English experts.

Why is the vocabulary test needed? According to Schmitt, et al (2011) the vocabulary level test will be useful to estimate the students' capability in relation with academic vocabulary, since it is the vocabulary needed in their study in university. Information on the result of vocabulary test can be used by lecturer and administrative department in educational context, to determine whether the student included in the category of students that have already able to read original

English text or not. Cobb (1997), Schmitt and Meara (1997), Laufer and Paribakht (1998) said, the information can also be used to identify possible lexical deficiencies which might need addressing. Likewise, the result of vocabulary level test can be used in research studies where an estimate of lexical size at the relevant frequency levels is considered informative.

An important reason on the importance of vocabulary test is based on the result of study by Schonell, et al, (1956) that the measurement of vocabulary mastery is directly related to the ability of speaking English in various ways. Schonell, et al, 1956) in Schmitt, et. all (2001) cited the vocabulary capability level 2000 words enables the student to have capability of oral communication in basic English for daily communication. Student's mastery at 3,000 word level enables the student to have additional material to have a discussion, however, Schonell emphasized that this 3,000-word level is at the crossroad to be able to read authentic text in English. Since, according to Schonell, et al, (1956), the most results of study have concluded that at the 5,000-word level of vocabulary mastery that students are able to have enough vocabulary so that they can read authentic English text. What if the students with 5,000 word-level of vocabulary finds new words? Schonell, et al, 1956) stated that of course there are still many words the students do not know if he/she only masters 5,000 words, however the level enables the students infer the meaning of many of the novel words from context, and to understand most of the communicative content of the text.

According to Schonell, et al, (1956), in the meanwhile, the students who have reached 10,000-word level of vocabulary can be said to learn new

vocabulary. In this sense, Hazenberg and Hulstijn (1996) found that the students with 10,000-word mastery are the minimum that students need to face language challenges at the university.

### **2.3.2 Vocabulary Testing Techniques**

According to Schmitt, Schmitt & Clapham (2001) as being quoted by Laufer, Kaovski (2010), students' vocabulary mastery can be measured by Nation's (1983) version Vocabulary Levels Test (VLT). The Nation's Test consists of level 2,000, 3,000, 5,000, and 10,000 words, ones that frequently appear; and 3,000, 5,000, 10,000 and academic vocabulary.

Yamamoto (2011) said, the vocabulary profile calculated and divided the words of each composition into four categories according to their frequency: (a) the most frequent 1,000 English words, (b) the second-most frequent thousand English words, (c) academic English words, and (d) the remainder, or less frequent words; that is, words that are not in any of the above lists. Each word appeared in the test represents "a word family" (i.e., the word, its infections and common derivations). Each level contains 30 items, except for academic vocabulary section that contains 36 items

What is the technique of calculating the score of each test participant? Each correct answer on the item gets 1 (one) point, while each wrong answer or no answer get 0 (zero) point. If the student participate in the tests level 3,000 words and level 5,000 words at once, by learning the technique of calculation undertaken by Schmitt, Schmitt & Clapham (2001), level test 3,000 words has 0

point (if all wrong) to 30 points (if all correct). So do the level test 5,000 words will get 0 point (if all wrong) to 30 points (if all correct). According to Laufer and Kalovski (2010), since each frequency level is 30 items for maximum score in which it represents the students' mastery level of 5,000 word-level, it will be  $30 \times 5 = 150$  which is the maximum score for 5,000 word-level test results.

Each VLT (for example the 3000-words level, the 5000-words level) was comprised of many sections. Each section was comprised of six words and three definitions, and each section the test takers were asked to match the words on the left with the definition given on the right.

For example:

- 
- |             |                               |
|-------------|-------------------------------|
| 1. original |                               |
| 2. private  | ___1___ first                 |
| 3. royal    | ___2___ not public            |
| 4. slow     | -----3-----all added together |
| 5. sorry    |                               |
| 6. total    |                               |
- 

The both results (level 3,000 word, and level 5,000 word) are combined. For example, if the student succeeded to answer 26 questions for level test 3,000 word, and 22 correct answers for level 5,000; the result is  $26 + 22 = 48$  points.

How to estimate the number of word families known by a learner? According to the formula created by Laufer (1998), word families' estimations were calculated based upon the total raw scores of receptive and controlled



productive tests. First, even though students were tested on three levels (2,000-word level, 3,000-words level, and 5,000-words-level), the 1,000-words, and 4,000 words were accounted for statistically. For the 1,000-words level, students were assumed to have the same score as for the 2,000 level. For the 4,000 level, the average score for the 3,000 and 5,000 words was taken. Then, the total score was added and multiplied by 5,000 (as students were tested up to a size of 5,000 word families). For example, if a student scored 20 out of 30 at the 2,000-word level, 15 out of 30 at the 3,000-word level, and 5 out of 30 at the 5,000-word level; it is estimated that a student has 20 out of the 30 at the 1,000-words level and 10 out of 30 at the 4,000-words level. The total score adds up to 70.

**Table 2.5 The Example of Total Score (Statistically)**

Vocabulary Levels Test	Scoring	
	Scores (based on the testing)	Estimated scores
1,000-words level	XXX	Have the same score as for the 2,000-words level → 20
2,000-words level	20	
3,000-words level	15	
4,000-words level	XXX	$\frac{1}{2} (15 + 5) \rightarrow 10$
5,000-words level	5	
<b>Total</b>	<b>40</b>	<b>30</b>
<b>Grand Total</b>	<b>70 (40 + 30)</b>	

Adapted from Nation (2009) and Maruyama (2011)

This grand total is multiplied by 5,000 and the sum is divided by 150 which roughly equals to 2,333 words families. This showed that the students were in control 1,500 – 2,500 word families (CEFR level; Tabel 2.1)

**Table 2.6 Total Score, Calculation, and English Vocabulary Level (Equality to the Number of Word Families)**

<b>Total Score</b>	<b>Calculation</b>	<b>English vocabulary level</b>
<b>150</b>	<b><math>150 \times 5,000 : 150</math></b>	<b>5,000 word families</b>
130	$130 \times 5,000 : 150$	4,333.33 word families
<b>120</b>	<b><math>120 \times 5,000 : 150</math></b>	<b>4,000 word families</b>
<b>90</b>	<b><math>90 \times 5,000 : 150</math></b>	<b>3,000 word families</b>
88	$88 \times 5,000 : 150$	2,933 word families
<b>50</b>	<b><math>50 \times 5,000 : 150</math></b>	<b>1,666 word families</b>
39	$39 \times 5,000 : 150$	1,300.00 word families

Adapted from Nation (2009) and Maruyama (2011)

### **2.3 Conceptual Framework**

The researcher wants to find out the capability of vocabulary mastery on the students in Department of English, State University of Jakarta. There are instruments to find out the level of vocabulary mastery made by Nation (1983, 1990) which was then modified by several other English specialists such as Schmitt (2001). The instruments called Vocabulary Test Level (VTL), consists of five categories of word-level, it is 2,000 words-level, 3,000 words level, academic vocabulary level, and 10,000 word-level.

In this research, researcher only used research instrument to find out 2,000 words-level, 3,000 words-level and 5,000 words-level on senior students of Department of English. The VTL instrument Schmitt (2001) is used for this study. The instrument has been used by many experts (Schmitt (2001), Golkar and

Yamini (2007), Maruyama (2011), and others), which the validity, reliability and practicality are valid. Each instrument measured the test takers' capability separately. The result of research will be processed using software SPSS to find out the descriptive statistic. The vocabulary level from CEFR is the basis to describe the profile of their vocabulary level.