



Assessing Knowledge and Use of English Vocabulary

Saengchan Hemchua*

Abstract

Assessing learners' knowledge and use of vocabulary is the key process in exploring vocabulary development. Knowing a word involves three main components: form, meaning, and use (Nation, 2001) and vocabulary acquisition is incremental (Schmitt, 2000). Therefore, the acquisition of word knowledge can be very difficult for L2 learners. Admittedly, not all facets of a learner' lexicon can be measured simultaneously. The decision on what aspects of vocabulary knowledge and use will be tested depends on the researcher's or teacher's purpose. This article examines various ways of measuring vocabulary knowledge and use. It then discusses and compares the methods of a receptive vocabulary knowledge test (Vocabulary Level Test) with a measure of vocabulary use (VocabProfile and Lexical Variation). These assessment tools have applications for L2 vocabulary teaching and learning.

Keywords: *Assessment, Vocabulary Test, Vocabulary Knowledge and Vocabulary Use*

* Deputy Dean for Academic Affairs & International Relations, International College for Sustainability Studies, Srinakharinwirot University
114 Sukhumvit 23, Bangkok 10110, THAILAND.
E-mail: saeng@swu.ac.th, saengchan_hemchua@hotmail.com

การวัดความรู้และการใช้คำศัพท์ภาษาอังกฤษ

แสงจันทร์ เหมเชื้อ*

บทคัดย่อ

การวัดความรู้และการใช้คำศัพท์ภาษาอังกฤษเป็นวิธีการสำคัญในการศึกษาพัฒนาการด้านคำศัพท์ของผู้เรียน เนชั่น (Nation, 2001) กล่าวว่า การรู้คำศัพท์ ประกอบด้วย ความรู้ด้านโครงสร้าง ความหมาย และการใช้คำ ชมิต (Schmitt, 2000) อ้างว่า การเรียนรู้คำศัพท์สะสมและเพิ่มขึ้นเป็นลำดับ ดังนั้น การรู้คำศัพท์ทุกแง่มุมจึงเป็นเรื่องยากสำหรับผู้เรียนภาษาที่สองและต้องรับว่าไม่สามารถวัดความรู้ทุกแง่มุมของคำศัพท์ได้ในเวลาเดียวกัน ทั้งนี้ขึ้นกับจุดประสงค์ของนักวิจัยหรือผู้สอนที่ต้องการวัดด้านใด บทความนี้ศึกษาเครื่องมือการวัดความรู้และการใช้คำศัพท์ โดยอภิปรายและเปรียบเทียบแบบทดสอบวัดความรู้คำศัพท์ภาษาอังกฤษ (Vocabulary Levels Test) และการวัดปริมาณการใช้คำศัพท์ (โปรแกรม VocabProfile และความหลากหลายของการใช้คำศัพท์) เครื่องมือทดสอบและประเมินนี้สามารถนำไปประยุกต์ใช้เพื่อการเรียนการสอนคำศัพท์ภาษาที่สองต่อไป

คำสำคัญ: การประเมินผล แบบทดสอบคำศัพท์ ความรู้และการใช้คำศัพท์

* รองคณบดีฝ่ายวิชาการและวิเทศสัมพันธ์ วิทยาลัยนานาชาติเพื่อศึกษาความยั่งยืน มหาวิทยาลัยศรีนครินทรวิโรฒ
เลขที่ 114 สุขุมวิท 23 กรุงเทพฯ 10110
E-mail: saeng@swu.ac.th, saengchan_hemchua@hotmail.com

Background

Assessing vocabulary knowledge is beneficial for the measurement of students' capacity and for finding ways to assist them. Learning vocabulary and the size of vocabulary have been extensively studied. For instance, approximately 54,000 word families (a base word plus its derivatives and inflections) have been identified in Webster's Third International Dictionary (1961). Researchers have calculated that native university graduates know approximately 17,000 word families. On the other hand, adult foreign learners of English typically know fewer than 5,000 word families (Nation & Waring, 1997). Laufer (1988) claims that 95% of text coverage is sufficient for reading comprehension. That is, university students require 2,000 high-frequency words plus 570 word families from the Academic Word List (AWL) (Coxhead, 2000). These 2,570 words cover 95% of general texts (Nation & Hwang, 1995). Hirsh and Nation (1992) claim that approximately 3,000 to 5,000 word families are sufficient for comprehension. Meanwhile, 2,000 to 3,000 word families are needed for spoken and written language use. Therefore, asking learners to read a text or produce writing when they have an inadequate vocabulary may result in comprehension difficulty and poor lexical choice.

Measuring learners' knowledge and use of vocabulary is usually restricted to a particular aspect of vocabulary and requires accurate interpretation. Previous research findings report how L2 vocabulary develops in different aspects. For example, an increase in receptive vocabulary size as measured by vocabulary tests is not guaranteed to increase low-frequency words used in writing because receptive vocabulary size and productive vocabulary use develop in different ways (Laufer, 1995, 1998). The results of lexical analysis do not simply reflect learners' vocabulary knowledge, but demonstrate how vocabulary is applied. Vocabulary use has been traditionally thought to involve making a connection between form and meaning; growth was considered to take place along a continuum from receptive to productive. It is now widely accepted that vocabulary use and development are not as simple as previously thought. As pointed out by Nation (1990, 2001), vocabulary knowledge involves three main components, form, meaning, and use, including various types of knowledge (e.g. spoken form, written form, grammatical patterns, collocations, associations, concepts, and constraints on use).

Receptive use (listening and reading) demands the ability to recognize and distinguish a word's sound and shape, have expectations regarding its grammatical pattern, what words it collocates with, know its register, recall its meaning, and be able to associate it with related words. In comparison, additional knowledge is required for productive use (speaking and writing). Productive knowledge involves the ability to pronounce a word, spell it, and use it correctly with grammatical patterns and other words with which it collocates. It also entails using high or low-frequency words in appropriate contexts and being able to use a word to express its meaning effectively.

However, L2 productive vocabulary use entails not only word knowledge, but other factors such as socio-cultural differences, learning situations, and communication strategies. In other words, vocabulary use is multi-dimensional and depends on several underlying factors. For example, words may be difficult to learn, difficult to spell, present phonological difficulties, or have confusing syntactic properties (e.g. *wish* exhibits more syntactic difficulties than *want*). An additional factor concerns *false friends*, such as *actuellement*; this word looks very similar to the English word *actual* but does not have the same meaning in French and English. Different cultural environments also present a challenge to learners (McCarthy, 1990). Consequently, learners may avoid using words that are too difficult for them and employ words with which they are confident (Hemchua & Schmitt, 2006). Learners may apply communication strategies when they are unable to retrieve a target word even if they know that word. Thus, they may give the definition of that word, or use synonyms. Students additionally apply a literal translation from their first language.

The assumption that “one’s lexicon receptive vocabulary is much larger than productive vocabulary and that reception precedes production” is supported by Melka (1997: 84). Being able to recognize words does not necessarily require knowledge of the complete form. “It takes less information about a word to interpret than it does to generate it...” (Norman, 1976: 135 cited in Melka, 1997: 87). Melka points out the factors that affect estimates of receptive and productive vocabulary as follows:

1. Type of words. If the choice of tested items is based on a dictionary, the size of the dictionary and the range of words (high-low frequency) will affect the results.
2. The grading of a test. The criteria (lenient-strict) to accept the correctly-tested items can influence the estimations.
3. The role of context. The context of the test setting can provide testees with a clue to recognize or generate a stimulus.
4. The notion of avoidance. In production, subjects sometimes avoid words that they are not comfortable with (e.g. taboo words), even if they know them, and also some words that they are not quite certain how to use.
5. Possible words versus actual words. The subjects may not use the words they know (possible words); they therefore remain in receptive knowledge.
6. The role of L1 and cognates. L1 cognates seem to facilitate reception but this effect may be decreased in higher stages of production since they are not completely equivalent.

Apart from the fuzzy boundary between receptive and productive vocabulary knowledge, the estimate of vocabulary size is complex. Waring (1999) finds that the size of receptive and productive vocabulary depends heavily on the language task involved. The degree of difficulty of tasks markedly affects whether language use is receptive or productive. Waring identifies higher scores on an easy version of a productive test than a difficult version of the receptive test with identical test items. This raises doubts about previous findings, showing that learners' receptive vocabulary size is greater than their productive vocabulary. No matter what factors apply in interpreting test results, it is necessary and beneficial for teachers to have background knowledge of assessment tools so that they can use them effectively.

Word Selection for Making a Test

Making a quantitative test to measure vocabulary size requires two crucial steps: selecting words and testing words (Nation, 1990). To select words for a test, a researcher needs to be confident that the words are representative samples, consider the number of words to test, and decide what type of test items to design.

To test words, a researcher must decide whether to test word recognition (knowing the meaning of words) or word recall (the production of words). Next, an appropriate type of text must be chosen, such as producing words in the mother tongue, a checklist test with some non-words, multiple choice for recognition tests, producing target words from the mother tongue, or testing synonyms or definitions in the target words.

To select and estimate vocabulary size, two primary methods have long been used: dictionary-based and frequency list occurrence. Using dictionaries to estimate vocabulary size seems to be a straightforward method but it is relatively complicated. To choose words from a dictionary to test, test makers have to randomly draw the words that they want to test. Estimating vocabulary size by this method is unreliable because it depends on the size of the selected dictionary. Nation (1993: 30) claims that “none of the 20 or more published studies of vocabulary size since 1907 has made full use of the necessary procedures. They have thus produced misleading estimates.” He suggests 8 steps that must be followed when sampling from a dictionary to create a test of vocabulary size:

1. Choose a dictionary that is big enough to cover the known vocabulary of the people being investigated.
2. Use a reliable way of discovering the total number of entries in the dictionary.
3. Use explicit criteria for deciding and stating (a) what items will not be included in the count and (b) what will be regarded as members of a word family.
4. Use a sampling procedure that is not biased towards items which occupy more space and have more entries.
5. Choose a sample that is large enough to allow an estimate of vocabulary size that can be given with a reasonable degree of confidence.
6. The sampling should be checked for the reliability of the application of the criteria for exclusion and inclusion of items.
7. The sample should be checked against a frequency list to make sure there is no bias in the sampling towards high-frequency items.
8. In the written report of the study, describe clearly and explicitly how each of the previous seven procedures was followed in sufficient detail to allow replication of any or all of the procedures.

Using a *frequency count* to estimate learners' vocabulary size is most widespread at present. Equal numbers of words are randomly drawn from different frequency levels such as the 1000 and 2000 word levels of West's General Service List (1953) or from specific word lists such as Coxhead's Academic Word List (AWL) (2000). Learners with high non-basic word (low-frequency and academic words) gain are considered to be advanced. However, problems exist since results rely heavily on the size of the frequency count and the materials selected to compile the lists.

Lexical Aspects for Testing

When vocabulary development is the main aim, Henriksen's (1999) three dimensions of lexical development is a reliable framework. She examines previous studies to see the focal processes in vocabulary learning and use and describes three dimensions of vocabulary development:

1. *Partial-precise knowledge*. This refers to vocabulary knowledge, ranging from the ability to recognize formal features (may not be able to recall meanings) to the ability to operate them (identify its associations and produce other word forms).

2. *Depth of knowledge*. This refers to lexical competence regarding various aspects of word knowledge (different shades of meaning, word classes, paradigmatic and syntagmatic knowledge).

3. *Receptive-productive knowledge*. This refers to the ability to use words receptively (comprehension) and productively (production).

Vocabulary development moves from vagueness to precision along the first dimension and associates with other aspects of word knowledge along the second dimension. "In this way, development along dimension 2 is seen as an important factor for lexical development along dimension 1" (p. 321). There seem to be unclear criteria regarding the difference between the receptive and productive competence of a particular word since no absolute division exists between reception and production. However, the depth of knowledge of a word is crucial for full understanding and for more thoroughly-learned words to extend into production. Dimensions 1 and 2 are in fact a knowledge continuum that involves

acquiring word meaning and elaborating knowledge of sense relations. Dimension 3 represents the levels of lexical use in reception and production.

Vocabulary Tests

Existing vocabulary tests generally attempt to evaluate various components of learners' vocabulary knowledge in different dimensions and various test formats, such as multiple choice, matching, checklists, and filling in blanks. To date, two prevalent dimensions for testing learners' lexical knowledge are breadth of knowledge and quality of knowledge (Read, 2000; Schmitt, 2000). The former estimates learners' vocabulary size, whereas the latter measures how well those words are known.

Breadth – Vocabulary Size Tests

Vocabulary size tests estimate how many words are known and are limited to partial knowledge of words; therefore, mainly the form and meaning of words are measured. Learners may know one meaning sense but not others. Similarly, they may understand the use of a word in speech but not in writing.

Vocabulary size can be measured receptively and productively in different test formats. Examples of receptive vocabulary size tests are, for example, Yes/No vocabulary tests, where learners are required to indicate the words that they know the meaning of (Beeckmans, Eyckmans, Janssens, Dufranne, and Van de Velde, 2001). Checklist tests are controlled by mixing non-words into the tests (Anderson & Freebody, 1983; Meara, 1994). Another format of the receptive vocabulary size test is Nation's Vocabulary Levels Test (1990) and the revised version by Schmitt et al. (2001). These tests are based on various word frequency levels that stem from the frequency with which L1 speakers use the language. They are used to estimate learners' word recognition by matching a word with its meaning. They are practical, provide usable results, and measure the breadth aspect of receptive vocabulary.

Vocabulary Levels Test

The Vocabulary Levels Tests (Versions 1 and 2) were revised and developed by Schmitt, Schmitt, and Clapham (Schmitt et al., 2001). They comprise five frequency

levels: 2000, 3000, 5000, 10,000 word frequency levels plus academic vocabulary. The frequency counts are based on Thorndike and Lorge (1944), Kucera and Francis (1967), and the General Service List (West, 1953). The academic vocabulary section was sampled from the Academic Word List (Coxhead, 1998 & 2000).

Each level contains ten clusters (altogether thirty items) (see Appendix) consisting of five noun clusters, three verb clusters, and two adjective clusters. The format of a noun cluster is as follows:

Choose the right word to go with each meaning. Write the number of that word next to its meaning.

1. business
2. clock 6 part of a house
3. horse 3 animal with four legs
4. pencil 4 something used for writing
5. shoes
6. wall

The Vocabulary Levels Tests (Versions 1 and 2) have been studied for their validity, reliability, equivalence, and practicality (Schmitt et al., 2001). The tests were administered to 801 international learners of English in England, New Zealand, Slovakia, Brazil, and Spain. The results from item analysis, profile analysis, factor analysis, and interviews provide evidence for their validity, reliability, and practicality. The results were one dimensional, scalable, and revealed learners' lexical knowledge. The two versions were equivalent for group scores but the authors suggested using the same version for longitudinal studies because individual learners always vary between versions. The Vocabulary Levels Test is a breadth of vocabulary knowledge test that estimates vocabulary size. It permits the assessment of vocabulary comprehension of English without requiring vocabulary production and it is used as a receptive test to explore learners' receptive vocabulary size.

Productive vocabulary size tests require learners to produce a target word (sometimes 1-2 letters are given to avoid near synonyms) in a given sentence (Laufer and Nation, 1999). This kind of test is confined to a single word form and

meaning of productive recall (e.g. I'm glad we had this opp____ to talk.). However, receptive and productive size tests are sometimes formed in a single test such as Laufer's (2002) *Computer Adaptive Test of Size and Strength*. This test is designed to assess two dimensions of word knowledge: size and strength. Size involves various word frequency levels. Strength concerns a combination of word knowledge according to the hierarchy of difficulty: receptive recognition, receptive recall, productive recognition, and productive recall.

Meara and Fitzpatrick (2000) also introduced a test of productive vocabulary called *Lex 30* for non-native speakers of English. It estimates learners' productive vocabulary by using a word association task. The test consists of thirty stimulus words (from Nation's first 1000 wordlist). Learners are required to write at least three response words for each stimulus word within thirty seconds. All response words are then analyzed by a computer program (similar to Nation's VocabProfile). Each response that is beyond the 1000 frequent content words scores one point. Though this test aims at estimating productive vocabulary, it also involves speed of access and requires more knowledge of word associations. Thus, it can also be considered a depth test to some extent.

Depth – Quality of Knowledge Tests

Apart from a single word's meaning and form, tests of the other aspects of paradigmatic and syntagmatic relations of words (collocations, grammatical functions, and register) show how well words are known. Most depth of knowledge tests aim to capture a learner's precise knowledge of a word, whether it is receptive or productive, but multiple levels of word knowledge are required. Examples of depth of knowledge tests are Read's (1993) word associates test, Greidanus et al.'s (2002) word associations test (French), and Hulstijn's (2002) polysemous words test as shown in the following examples:

Instructions: circle related words for this word: Edit

arithmetic film pole publishing

revise risk surface text (Read's Associates Test, 1993: 359)

Instructions: select 2 to 6 meanings for this word: Dash

- (a) (noun) A mixture of stylishness, enthusiasm and courage.
- (b) (noun) A straight, horizontal line used in writing.
- (c) (noun) A short fast race.
- (d) (verb) To be in a hurry and to have to leave immediately.
- (e) (noun) A small quality of something which you add when you are preparing food or mixing a drink.
- (f) (verb) To throw or push something violently, often so hard that it breaks.

(Hulstijn, 2002)

Compared to breadth or size tests, depth tests usually require more time to complete and fewer words are assessed. Even if a depth test cannot give an overall picture of a learner's lexicon, it is generally useful for some specific purposes.

Breadth of Knowledge Tests vs. Depth of Knowledge Tests

Obviously, breadth and depth of vocabulary knowledge and use are different dimensions and are measured by different types of tests. However, word knowledge is multi-dimensional and comprises various components. Thus, these dimensions may be related. Vermeer's (2001) empirical studies prove this claim. She assumes that "a deeper knowledge of words is the consequence of knowing more words,..." (p. 222). She studied fifty Dutch monolingual and bilingual kindergartens by using a receptive vocabulary task and a description task to measure breadth of knowledge. An association task (providing the features and characteristics of ten words) was used to measure depth of knowledge. She revealed connections between the breadth and depth measures and concluded that "if one knows more words, one can describe a stimulus word in greater depth" (p. 225). Furthermore, breadth measures are better able to distinguish between L1 and L2 learners, while depth measures depend on individual characteristics such as talkativeness. Her study of 1600 Dutch children, aged four and seven, including both monolingual and bilingual learners, shows that breadth and depth knowledge are affected by the same factors; frequency of language input also influences vocabulary acquisition. Therefore, size measures can be interpreted as giving some indication of depth of knowledge (Vermeer, 2001).

Tests of Vocabulary Use

The vocabulary size tests and depth of knowledge tests mentioned above are mostly discrete, selective, and context-independent. They are not suitable for measuring how many or how well words are used in learners' compositions and speaking. To do this, other comprehensive measures should be applied. For example, Laufer and Nation's (1995) VocabProfile measures how many words are used in a text at various word levels. A higher number of academic words indicates a more advanced text.

VocabProfile

VocabProfile (a new version is called *Range & Word*) is a computer program devised by Heatley, Kyongho, and Nation at Victoria University in New Zealand. VocabProfile provides three base lists which can be compared to any texts to determine whether or not the words in the texts are in the lists. The first (BASEWRD 1.DAT) and the second (BASEWRD 2. DAT) base lists are drawn from the first and second 1000 most frequently-used words from Michael West's General Service List (1953). The third base list (BASEWRD 3. DAT) is from the University Word List compiled by Xue and Nation (1984). The program can compare the similarity of words in two different texts, and the lists include both American and British spellings. Words with apostrophes are counted separately.

The results from VocabProfile show not only the amount of coverage of three base lists in a text, but also provide the frequency of word types in word families. This information is useful for grading reading texts aimed at foreign learners and also for exploring the number of words used in foreign language learners' free writing. The table below shows the quantity of coverage of each of the three base lists of a given text in terms of word tokens, word types, and word families.

WORD LIST	TOKENS/ %	TYPES/ %	FAMILIES
One	54/ 72.0	34/ 69.4	33
Two	2/ 2.7	2/ 4.1	2
Three	14/ 18.7	9/ 18.4	9
Not in the lists	5/ 6.7	4/ 8.2	?????
Total	75	49	44

(From the instructions of VocabProfile)

The example shows that 54 words in a sample written text are in base list one (the first 1,000 word level) and these 54 make up 72.0% of the total 75 words in this text. “Not in the lists” represents words outside the three levels.

VocabProfile is a practical tool to measure the vocabulary size of learners' compositions. It is based on words from word frequency lists and accounts for frequency of occurrence, coverage, and quantity of words in a particular corpus. In order to explore how a learner's vocabulary size progresses, the distinction between high-frequency words and low-frequency words should be a logical criterion for measurement since this is the way in which native speakers develop and use vocabulary. It reveals what types of vocabulary are used and what stage a learner is at in his or her vocabulary development because frequency levels discriminate between proficiency levels. Compositions with a higher number of infrequently-used words are considered more advanced because it is assumed that higher-frequency words are more likely to be known. Likewise, compositions with a higher number of academic vocabulary items are viewed as more skillful and sophisticated. The results from the profile are objective and stable over all compositions regardless of the group because the criteria are fixed within a certain band of word levels.

However, VocabProfile exhibits some shortcomings that need to be addressed. For one, the frequency lists depend on the corpus of texts in which counts are made. For instance, West's General Service List is outdated, based on a written corpus with no collocation information and has insufficient word coverage (Engels, 1968; Carter and McCarthy, 1988). Some researchers argue that VocabProfile is sensitive to text length and is less reliable with texts with fewer than 200 words (Meara& Bell, 2001). Apart from minor limitations, VocabProfile

is considered appropriate for analyzing written vocabulary size. Its criteria are similar to the Vocabulary Levels Test so the results from the two tools can be correlated.

An assessing tool similar to VocabProfile is Meara and Bell's (2001) P_Lex. This program works well with short texts of at least 120 words to assess the lexical complexity of low-level learners. P_Lex divides a text into ten word-segments and categorizes words into easy and difficult (easy words signify Nation's first 1000 word list and difficult words are defined as all other words).

Three widely-used measures for analyzing vocabulary usage include computing the ratio of specific types of words to the total words in the text. These measures are named *lexical sophistication*, *lexical density*, and *lexical variation*. Sophisticated words refer to advanced words relative to the level of students. Lexical words refer to nouns, verbs, adjectives, and adverbs. Lexical types refer to different word forms in a text. These measures have limitations and require careful application. For example, lexical sophistication depends on the level of students; lexical variation depends on text length because words tend to be repeated in a longer text. The ratio may not express a straightforward result since various factors are involved in language use. Currently, mathematical models that control for text length are shown to be a valid measure of lexical variation or lexical diversity. Malvern and Richard's (1997) model ($TTR = \frac{2}{DN} [(1 + DN) - 1]$), for example, is a computerized lexical density measure that claims to be independent of sample size.

Lexical Variation (LV)

A measure of lexical variation is the type/token ratio of a text. This ratio indicates the number of different word forms in a text divided by the total number of words in that text. For example, consider the following sentences:

I agree with many people who believe that smoking in public places should be banned. Nowadays, many people, especially teenagers, smoke.

The example contains 21 tokens but only 19 types since the words *many* and *people* both appear twice. Therefore, the type/token ratio of these two sentences is $19/21 = 0.90$. However, scholars generally agree that this ratio has its

limitations because it depends directly on text length (e.g. Chafe & Daniclewicz, 1987; Biber, 1988; Laufer& Nation, 1995; and Wesche&Paribakht, 1996). With very short texts, the type/token ratio has been shown to be unstable. Additionally, the ratio decreases when the total number of words in a text increases because a large number of different words used in the first part of a text tend to be repeated later. Hence, the type/token ratio appears to be higher in shorter texts. To avoid these limitations, the lengths of texts must be equalized before computing.

Another shortcoming of lexical variation is that it cannot show how advanced or sophisticated the vocabulary is that is employed in compositions. Lexical variation does not discriminate between low- and high-frequency words or simple and sophisticated words. Compositions with equal lexical variation may have different numbers of word levels. Lexical variation merely shows how many different words are used in a text of a particular length.

Regardless of its limitations, LV or type/token ratio (TTR) has been widely used in the research of lexis (Richards and Malvern, 1997). The fact is, people must choose words and phrases to convey their thoughts. Knowledge of the lexical options used to produce language to refer to objects, states, and events requires cognitive effort (Chafe & Danielewicz, 1987). For L2 learners, “LV shows how inclined the learner is to repeat the same words in his writing. The higher the LV, the more varied the active word repertoire” (Laufer, 1991: 442). This means that LV is a rough measure of vocabulary flexibility. Accordingly, it shows how well L2 learners vary their vocabulary in compositions.

Pedagogic Implications

The assessment tools mentioned above have considerable applications for L2 vocabulary teaching and learning. The tools can be applied to measure learners’ basic vocabulary knowledge in terms of the first version on the Vocabulary Levels Test and VocabProfile. The former is used to explore the breadth of subjects’ receptive vocabulary while the latter is used to investigate the percentage of word types at four frequency word levels in their written productions. Both tools are based on the measurement of precise vocabulary frequency levels rather than a global size consideration. It is a great advantage to use these measurements to

compare the two types of vocabulary breadth; namely, vocabulary comprehension and written vocabulary use.

It is assumed that lexical progress will be apparent as learners reach higher academic levels. To test this assumption, for example, productive vocabulary size (from free compositions analyzed by VocabProfile) and receptive vocabulary size (from the Vocabulary Levels Test) at all four levels of university students are compared to investigate vocabulary size development. To examine its development, mean percentages of word types at a particular word level (first 1000, second 1000, University World List [UWL] and words not in the lists) appearing in the compositions are compared across year levels. The evidence implies that productive L2 vocabulary use is remarkably complicated and slow moving. It is also very difficult to predict its progress in writing. Carson et al. (1990: 247) claim that “L2 literacy development is a complex phenomenon for already literate L2 learners.” Additionally, “production involves retrieval and use” (Carter and McCarthy, 1988: 80), which is exhausting. The results from various studies remain inconclusive concerning the development of production skills. Factors such as reading and writing experience, L2 exposure, and teaching styles may result in the variability of outcomes.

In order to examine the relationship between receptive and productive vocabulary size, the scores on Academic Word List (AWL) from the Vocabulary Levels Test and the percentage of word types on the UWL in the composition can be computed to find a correlation at each year level. Only academic vocabulary is applied since it is the only common frequency band for the measurement instruments and no equivalent frequency band exists between the two tasks. The AWL and UWL are similar but not exactly the same. This may lower the correlation to some degree.

Despite this, productive vocabulary use does not account merely for its meaning, by which receptive knowledge is measured. Rather, it includes factors such as spellings, grammatical patterns, associations, and collocations (Nation, 1990). In L2 writing, all kinds of word knowledge are utilized. Consequently, the number of receptive and productive vocabulary items may be affected by different methods of retrieving knowledge and procedures.

In order to investigate the development of lexical variation, means of the type/token ratio can be compared across year levels. However, the variation of words does not account for their frequency or difficulty. For example, Year 1 may vary more in vocabulary on the 2000 most frequently-used words, whereas Year 4 may vary more in vocabulary on the UWL or on the less frequently-used words. Recent studies also claim that the type/token ratio is not an effective measure of lexical richness (degree of vocabulary use). A more effective operative measure should be based on the extent of word difficulty (Vermeer, 2000). As mentioned, the importance of lexis in composition evaluation has empirical support (Linnarud, 1986; Astika, 1993). However, it is still uncertain whether or not the degree of LV will play a role in writing quality. Meara and Bell (2001: 6) call this kind of measure “Intrinsic Measures of Lexical Variety” since the variety is measured only among the words present in the text. Nevertheless, all English compositions with high LV written by Swedish university students in Linnarud’s (1975) study obtained low grades.

Apart from examining the vocabulary size of learners across different groups (cross-sectional study) within a yearly interval, teachers may also examine the vocabulary size of learners over an academic year among selected individuals (a longitudinal study). Productive vocabulary size can be measured by the percentage of word types appearing in timed compositions, while receptive vocabulary size can be measured by the scores on the Vocabulary Levels Test to determine progress. Teachers may also want to quantify the development of lexical variation as measured by the type/token ratio in timed compositions over the course of an academic year within a group of selected individuals.

Results will give a clear picture of learners’ vocabulary size and this will benefit teachers by establishing teaching guidelines. Although production is difficult, Long (1990) believes that adult learners can continue to gain new vocabulary throughout their lives. Productive vocabulary growth can be stimulated if people are provided opportunities to practice and use their vocabulary regularly. Promoting productive vocabulary is essential to language teaching because university EFL learners are expected to graduate with sufficient vocabulary.

Moreover, increases in knowledge of non-basic words assist learners in effectively communicating ideas in academic writing.

Conclusion

Breadth and depth are the conventional aspects for assessing knowledge and use of vocabulary. Vocabulary breadth tests provide an estimate of how many words are known, whereas depth tests show how well words are known. The VocabProfile can be used to analyze data for vocabulary size in written and spoken texts since it is a practical tool which classifies words in three levels. Results can then be compared against their receptive vocabulary size, which is assessed by the Vocabulary Levels Test. Nevertheless, lexical use is complicated and word knowledge comprises many component parts.

Vocabulary usage involves avoidance, lack of confidence, and use of various communication strategies (Hemchua & Schmitt, 2006). Consequently, a combination of vocabulary tests should be applied for full assessment. For example, the production of EFL learners changes over time. This can be measured by the proportion of low-frequency words used in writing and speaking. How students vary the vocabulary in their writing and speaking can be measured by the type/token ratio. In order to easily comprehend a reading text, 80% of the vocabulary in that text needs to be known. Teachers can measure students' receptive vocabulary size with the Vocabulary Levels Test. Vocabulary size demonstrates the amount of vocabulary gained receptively and productively. Nevertheless, these tools cannot demonstrate how well learners apply words or how correctly the words are used. Teachers must analyze texts that demonstrate learners' ability to employ and acquire language.

References

- Anderson, R.C. and Freebody, P. (1983). Reading Comprehension and the Assessment and Acquisition of Word Knowledge. *Advances in Reading/Language Research*. 2: 231-256.
- Astika, G.G. (1993). Analytical Assessment of Foreign Students' Writing. *RELC Journal* 24: 61-72.
- Beeckmans, R., Eyckmans, J., Janssens, V., Dufranne, M., and Van de Velde, H. (2001). Examining the Yes/No Vocabulary Test: Some Methodological Issues in Theory and Practice. *Language Testing*. 18(3): 235-174.
- Biber, D. (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.
- Carson, J.E., Carrell, P.L., Silberstein, S., Kroll, B. and Kuehn, P.A. (1990). Reading-Writing Relationship in First and Second Language. *TESOL Quarterly*. 24(2): 245-266.
- Carter, R. and McCarthy, M. (1988). *Vocabulary and Language Teaching*. London: Longman.
- Chafe, W. and Danielewicz, J. (1987). Properties of Spoken and Written Language. In Horowitz, R. and Samuels, S.J. (eds). *Comprehending Oral and Written Language*. London: Academic Press.
- Coxhead, A. (1998). An Academic Word List. English Language Institute Occasional Publication 18. Wellington: School of Linguistics and Applied Language Studies, Victoria university of Wellington.
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly* 34(2): 213-237.
- Engels, L.K. (1968). The Fallacy of Word Counts. *International Review of Applied Linguistics*. 6: 213-231.
- Greidanus, T., Bogaards, P., Van der Linden, E., Nienhuis, L and de Wolf, T. (2002). The Construction and Validation of a Deep Word Knowledge Test for Advanced Learners of French. Paper presented at Second Language Vocabulary Acquisition Colloquium, Universiteit Leiden, The Netherlands. March 2002.
- Henriksen, B. (1999). Three Dimensions on Vocabulary Development. *Studies in Second Language Acquisition*. 21(2): 303-317.

- Hemchua, S. and Schmitt, N. (2006). An Analysis of Lexical Errors in the English Compositions of Thai Learners. *Prospect: An Australian Journal of TESOL*. 21(3): 3-25.
- Hirsh, D. and Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language*. 8(2): 689-696.
- Hulstijn, J.H. (2002). Testing Knowledge of Polysemy Words: A Pilot Study. Paper presented at Second Language Vocabulary Acquisition Colloquium, Universiteit Leiden, The Netherlands. March 2002.
- Kucera, H. and Francis, W.N. (1967). *A Computational Analysis of Present Day American English*. Providence, RI: Brown University Press.
- Laufer, B. (1988). What percentage of text-lexis is essential for comprehension? In Lauren, C. and Nordmann, M. (eds). *Special Language: From Humans Thinking to Thinking Machines*. Clevedon: Multilingual Matters.
- Laufer, B. (1991). The Development of L2 Lexis in the Expression of the Advanced Learner. *The Modern Language Journal*. 75(4): 140-148.
- Laufer, B. (1995). Beyond 2000: A Measure of Productive Lexicon in a Second Language. In Eubank, L., Selinker, L. and Sharwood, M. (eds). *The Current State of Interlanguage*. Philadelphia: John Benjamins.
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? *Applied Linguistics*. 19(2): 255-271.
- Laufer, B. (2002). CATSS (Computer Adaptive Test of Size and Strength). Paper presented at Second Language Vocabulary Acquisition Colloquium, Universiteit Leiden, The Netherlands. March 2002.
- Laufer, B. and Nation, P. (1995). Vocabulary Size and Use: Lexical Richness in L2 Written Production. *Applied Linguistics*. 16(3): 307-322.
- Laufer, B. and Nation, P. (1999). A Vocabulary Size Test of Controlled Productive Ability. *Language Testing*. 16: 33-51.
- Linnarud, M. (1975). Lexis in Free Production, Swedish-English Contrastive Studies. Report 6. English Department, Lund university.
- Linnarud, M. (1986). *Lexis in Composition: A Performance Analysis of Swedish Learners' Written English*. Lund, Sweden: University of Lund.

- Long, M. (1990). The Least a Second Language Acquisition Theory Needs to Explain. *TESOL Quarterly*. 24: 649-666.
- Malvern, D. and Richards, B. (1997). A New Measure of Lexical Density. In Ryan, A. and Wray, A. (eds). *Evolving Models of Language*. Clevedon: Multilingual Matters: 58-71.
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- Meara, P. (1994). *LLEX: The Threshold Level Vocabulary Tests*. University of Wales, Swansea: Centre for Applied Language Studies.
- Meara, P. and Bell, H. (2001). P_Lex: A Simple and Effective Way of Describing the Lexical Characteristics of Short L2 Texts. *Prospect*. 16(3): 5-19.
- Meara, P. and Fitzpatrick, T. (2000). Lex 30: An Improved Method of Assessing Productive Vocabulary in an L2. *System*. 28: 19-30.
- Melka, F. (1997). Receptive vs. Productive Aspects of Vocabulary. In Schmitt, N. and McCarthy, M. (eds): 84-102.
- Nation, I.S.P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House.
- Nation, I.S.P. (1993). Using Dictionaries to Estimate Vocabulary Size: Essential but Rarely Followed Procedures. *Language Testing*. 10: 37-40.
- Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nation, P. and Hwang, K. (1995). Where would general service vocabulary stop and special purpose vocabulary begin? *System*. 23(1): 35-41.
- Nation, P. and Waring, R. (1997). Vocabulary Size, Text Coverage and Word Lists. In Schmitt, N. and McCarthy, M. (eds): 6-19.
- Read, J. (1993). The Development of a New Measure of L2 Vocabulary Knowledge. *Language Testing*. 10(3): 355-371.
- Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Richards, B. and Malvern, D. (1997). *Type-Token and Type-Type Measures of Vocabulary Diversity and Lexical Style: An Annotated Bibliography*. Reading: University of Reading.
- Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.

- Schmitt, N., Schmitt, D. and Clapham, C. (2001). Developing and Exploring the Behavior of 2 New Versions of the Vocabulary Levels Test. *Language Testing*. 18(1): 55-88.
- Thorndike, E.L. and Lorge, I. (1944). *The Teacher's Word Book of 30,000 Words*. Teachers college, Columbia University.
- Vermeer, A. (2000). Coming to Grips with Lexical Richness in Spontaneous Speech Data. *Language Testing*. 17(1): 65-83.
- Vermeer, A. (2001). Breadth and Depth of Vocabulary in Relation to L1/L2 Acquisition and Frequency of Input. *Applied Psycholinguistics*. 22: 217-234.
- Waring, R. (1999). Tasks for Assessing Receptive and Productive Second Language Vocabulary. Unpublished Ph.D. thesis. University of Wales.
- Wesche, M. and Paribakht, T.S. (1996). Assessing Second Language Vocabulary Knowledge: Depth versus Breadth. *The Canadian Modern Language Review*. 53(1): 13-40.
- West, M. (1953). *A General Service List of English Words*. London: Longman.
- Xue, G. and Nation, I.S.P. (1984). A University Word List. *Language Learning and Communication*. 3(2): 215-229.

Appendix

A Vocabulary Levels Test (Version 1)

The Vocabulary Levels Test (Version 1) was revised and developed by Schmitt, Schmitt, and Clapham (2001). It is an assessment tool used to measure the subjects' receptive vocabulary size.

Name _____

Nationality _____

Time required to finish the test: _____ minutes

This is a vocabulary test. You must choose the right word to go with each meaning. Write the number of that word next to its meaning. Here is an example.

1. business
2. clock _____part of a house
3. horse _____animal with four legs
4. pencil _____something used for writing
5. shoe
6. wall

You answer it in the following way.

1. business 6 part of a house
2. clock 3 animal with four legs
3. horse 4 something used for writing
4. pencil
5. shoe
6. wall

Some words are placed in the test to make it more difficult. You do not have to find the meaning of these words. In the example above, these words are *business*, *clock*, and *shoe*.

If you have no idea about the meaning of a word, do not guess. But if you think that you might know the meaning, then you should try to find the answer.

2,000 word level

1. birth
2. dust _____game
3. operation _____winning
4. row _____being born
5. sport
6. victory

1. choice
2. crop _____heat
3. flesh _____meat
4. salary _____money paid regularly for doing a job
5. secret
6. temperature

1. cap
2. education _____teaching and learning
3. journey _____numbers to measure with
4. parent _____going to a far place
5. scale
6. trick

1. attack
2. charm _____gold and silver
3. lack _____pleasing quality
4. pen _____not having something
5. shadow
6. treasure

1. cream
2. factory _____part of milk
3. nail _____a lot of money
4. pupil _____person who is studying
5. sacrifice
6. wealth

1. adopt
2. climb _____go up
3. examine _____look at closely
4. pour _____be on every side
5. satisfy
6. surround

1. bake
2. connect _____join together
3. inquire _____walk without purpose
4. limit _____keep within a certain size
5. recognize
6. wander

1. burst
2. concern _____break open
3. deliver _____make better
4. fold _____take something to someone
5. improve
6. urge

1. original
2. private _____first
3. royal _____not public
4. slow _____all added together
5. sorry
6. total

1. brave
2. electric _____commonly done
3. firm _____wanting food
4. hungry _____having no fear
5. local
6. usual

3,000 word level

1. belt
2. climate _____idea
3. executive _____inner surface of your hand
4. notion _____strip of leather worn around the waist
5. palm
6. victim

1. acid
2. bishop _____cold feeling
3. chill _____farm animal
4. ox _____organization of framework
5. ridge
6. structure

1. bench
2. charity _____long seat
3. jar _____help to the poor
4. mate _____part of a country
5. mirror
6. province

1. boot
2. device _____army officer
3. lieutenant _____a kind of stone
4. marble _____tube through which blood flows
5. phrase
6. vein

1. apartment
2. candle _____a place to live
3. draft _____chance of something happening
4. horror _____first rough form of something written
5. prospect
6. timber

1. betray
 2. dispose _____frighten
 3. embrace _____say publicly
 4. injure _____hurt seriously
 5. proclaim
 6. scare
-
1. encounter
 2. illustrate _____meet
 3. inspire _____beg for help
 4. plead _____close completely
 5. seal
 6. shift
-
1. assist
 2. bother _____help
 3. condemn _____cut neatly
 4. erect _____spin around quickly
 5. trim
 6. whirl
-
1. annual
 2. concealed _____wild
 3. definite _____clear and certain
 4. mental _____happening once a year
 5. previous
 6. savage
-
1. dim
 2. junior _____strange
 3. magnificent _____wonderful
 4. maternal _____not clearly lit
 5. odd
 6. weary

5,000 word level

1. balloon
2. federation _____ bucket
3. novelty _____ unusual interesting thing
4. pail _____ rubber bag that is filled with air
5. veteran
6. ward

1. alcohol
2. apron _____ stage of development
3. hip _____ state of untidiness or dirtiness
4. lure _____ cloth worn in front to protect your clothes
5. mess
6. phase

1. apparatus
2. compliment _____ expression of admiration
3. ledge _____ set of instruments or machinery
4. revenue _____ money received by the Government
5. scrap
6. tile

1. bulb
2. document _____ female horse
3. legion _____ large group of soldiers or people
4. mare _____ a paper that provides information
5. pulse
6. tub

1. concrete
2. era _____ circular shape
3. fiber _____ top of a mountain
4. loop _____ a long period of time
5. plank
6. summit

1. blend _____mix together
2. devise _____plan or invent
3. hug _____hold tightly in your arms
4. lease
5. plague
6. reject

1. abolish _____bring to an end by law
2. drip _____guess about the future
3. insert _____calm or comfort someone
4. predict
5. soothe
6. thrive

1. bleed _____come before
2. collapse _____fall down suddenly
3. precede _____move with quick steps and jumps
4. reject
5. skip
6. tease

1. casual _____sweet smelling
2. desolate _____only one of its kind
3. fragrant _____good for your health
4. radical
5. unique
6. wholesome

1. gloomy _____empty
2. gross _____dark or sad
3. infinite _____without end
4. limp
5. slim
6. vacant

10,000 word level

1. antics
2. batch _____foolish behavior
3. connoisseur _____a group of things
4. foreboding _____person with a good knowledge of art or music
5. haunch
6. scaffold

1. auspices
2. dregs _____confused mixture
3. hostage _____natural liquid present in the mouth
4. jumble _____worst and most useless parts of anything
5. saliva
6. truce

1. casualty
2. flurry _____someone killed or injured
3. froth _____being away from other people
4. revelry _____noisy and happy celebration
5. rut
6. seclusion

1. apparition
2. botany _____ghost
3. expulsion _____study of plants
4. insolence _____small pool of water
5. leash
6. puddle

1. arsenal
2. barracks _____happiness
3. deacon _____difficult situation
4. felicity _____minister in a church
5. predicament
6. spore

1. acquiesce
2. bask _____to accept without protest
3. crease _____sit or lie enjoying warmth
4. demolish _____make a fold on cloth or paper
5. overhaul
6. rape

1. blaspheme
2. endorse _____slip or slide
3. nurture _____give care and food to
4. skid _____speak badly about God
5. squint
6. straggle

1. clinch
2. jot _____move very fast
3. mutilate _____injure or damage
4. smolder _____burn slowly without flame
5. topple
6. whiz

1. auxiliary
2. candid _____bad-tempered
3. luscious _____full of self-importance
4. morose _____helping, adding support
5. pallid
6. pompous

1. dubious
2. impudent _____rude
3. languid _____very ancient
4. motley _____of many different kinds
5. opaque
6. primeval

Academic Vocabulary

1. benefit
 2. labor _____work
 3. percent _____part of 100
 4. principle _____general idea used to guide one's actions
 5. source
 6. survey
-
1. element
 2. fund _____money for a special purpose
 3. layer _____skilled way of doing something
 4. philosophy _____study of the meaning of life
 5. proportion
 6. technique
-
1. consent
 2. enforcement _____total
 3. investigation _____agreement or permission
 4. parameter _____trying to find information about something
 5. sum
 6. trend
-
1. decade
 2. fee _____10 years
 3. file _____subject of a discussion
 4. incidence _____money paid for services
 5. perspective
 6. topic
-
1. colleague
 2. erosion _____action against the law
 3. format _____wearing away gradually
 4. inclination _____shape or size of something
 5. panel
 6. violation

1. achieve
 2. conceive _____change
 3. grant _____connect together
 4. link _____finish successfully
 5. modify
 6. offset
-
1. convert
 2. design _____keep out
 3. exclude _____stay alive
 4. facilitate _____change from one thing
 5. indicate
 6. survive
-
1. anticipate
 2. compile _____control something skillfully
 3. convince _____expect something will happen
 4. denote _____produce books and newspapers
 5. manipulate
 6. punish
-
1. equivalent
 2. financial _____most important
 3. forthcoming _____concerning sight
 4. primary _____concerning money
 5. random
 6. visual
-
1. alternative
 2. ambiguous _____last or most important
 3. empirical _____something different that can be chosen
 4. ethnic _____concerning people from a certain nation
 5. mutual
 6. ultimate

