University Freshmen’s EFL Receptive and Productive Recall Vocabulary Knowledge and Use*

Donald Makarchuk
(Kyonggi University)


This article reports on a study that compared EFL freshman university students’ receptive and productive recall vocabulary knowledge and their ability to use that knowledge, as it remains unclear in the literature to what extent knowledge of vocabulary, especially productive knowledge, is indicative of learners’ ability to use vocabulary to communicate. The study first measured the vocabulary size of 169 students from 24 majors. Next, deeper meaning word knowledge and vocabulary use-related knowledge were assessed by administering depth tests to 51 of the students who completed the size tests. The study found that the participants were able to use only 60.3% of the words that were known receptively and productively on the size tests, and that receptive deeper meaning vocabulary knowledge was 43.0% greater than productive knowledge. Finally, the study results support the concept of a vocabulary knowledge continuum, but highlight the importance of including both receptive and productive knowledge, as they were found to develop in a dissimilar manner.

**Key words:** vocabulary acquisition, vocabulary knowledge and use, receptive and productive vocabulary knowledge, vocabulary knowledge continuum

1. **INTRODUCTION**

Because of the growing importance of English communicative competence for many EFL university students and vocabulary’s essential role in L2 communication (McCarthy, 1990; Read & Chapelle, 2001), the present study investigated freshman university students’ vocabulary knowledge and use in the South Korean context. Having an extensive vocabulary has long been regarded as beneficial to an individual’s success. However, it is

* This work was supported by a Kyonggi University Grant 2011.
not just the number of words one knows that is important, but also how well the words are known and to what extent they can be used for communicative purposes (Shea, 2010). The ability to communicate well in English is vital for both native speakers and learners of English as a second/foreign language. In both ESL and EFL contexts the use of English is essential not only for developing language skills, but also for learning subject matter in the increasing number of English-medium courses available to learners (Fouser, 2011; Lee, 2011). Insufficient English language proficiency places learners at a disadvantage in English-medium courses (Wiley & Hartung-Cole, 1998), and this can have negative academic and psychological consequences (Han, 2011; Oh, 2011). In addition, as some vocabulary research in EFL settings has found university students less than adequately prepared for tertiary-level education in English (Cobb & Horst, 2000; Mokhar, 2010), the question of whether EFL learners have the vocabulary skills needed to succeed at university is of great importance.

As Nation (2001) writes, merely acquiring vocabulary knowledge is generally not the learner’s goal; rather, they seek to speak, write, read, and listen more effectively with the aid of an expanding lexicon. The distinction between vocabulary knowledge and the use of vocabulary knowledge for the purpose of communication has been noted by a number of researchers (Coady, 1997; Corson, 1995; Laufer, Elder, Hill, & Congdon, 2004; Nation, 2007; Read, 2000; Schmitt, 2010), but the dividing line between the two is less than clear. Clearly, one must have vocabulary knowledge before one can use it, but having such knowledge is not a guarantee that it can be used, as other knowledge (e.g., grammar, constraints on use) is generally required before vocabulary can be used communicatively.

One approach to conceptualizing vocabulary knowledge has been to posit a learner’s vocabulary knowledge as existing on a continuum with easier to acquire aspects at one end and harder to acquire ones further along (Henriksen, 1999). The vocabulary continuum metaphor has also been used to describe the receptive-productive vocabulary knowledge dimension with Schmitt (2010) suggesting that the more contextualized aspects of word knowledge (e.g., register, collocation) would likely be acquired after less contextualized knowledge such as basic form-meaning knowledge. One concern of researchers has been to better understand the development of vocabulary knowledge as it proceeds along the hypothetical continuum. For example, according to Read (2000), Lauffer and Nation (1999) claim that having productive vocabulary knowledge of a word implies that the word is available for use. Read, however, avers that productive recall knowledge represents “production in only a very limited sense” (p. 157), and is unlikely to meet his test of vocabulary use, which is to use words in speech or writing. Vocabulary use claims and counter-claims highlight the importance of understanding the process of vocabulary acquisition as it progresses incrementally along the continuum from basic vocabulary knowledge to the everyday communicative use of words. However, according to Schmitt
most vocabulary research has focused on explicit/declarative knowledge such as form-meaning knowledge with much less attention having been given to implicit/procedural knowledge, which is the ability to use lexical items. In response to this need, the present study sought to better understand the nature of the vocabulary knowledge continuum by assessing learners' basic form-meaning knowledge, more precise meaning knowledge, and knowledge related to the use of words. To measure learner vocabulary knowledge, the study conducted a battery of receptive and productive recall vocabulary size and depth tests.

2. BACKGROUND

2.1. Vocabulary Knowledge: Size and Depth

It is generally accepted that the central component of vocabulary knowledge is the form-meaning link (Laufer et al., 2004; Nation, 2001), as to know a word is to know at least its meaning. However, word knowledge entails more than a basic link between a symbol and concept. For example, many words have more than one sense, and word knowledge is enriched by knowledge of associated words (Meara, 2009). Chapelle (1998) in an attempt to define the construct of vocabulary knowledge identified three approaches: trait, behaviorist, and interactionalist. According to Chapelle, a vocabulary trait approach seeks to define vocabulary knowledge in terms of the individual's vocabulary size, word characteristics (e.g., "phonemic, graphemic, morphemic, syntactic, semantic, and collocational features" (p. 36)), lexicon organization, and processes associated with vocabulary access. When assessing learners, trait theorists try to minimize context-related effects on performance. A behaviorist approach to vocabulary construct definition focuses not on the individual's internal vocabulary characteristics, but rather on external contexts that give rise to the use of vocabulary. For example, using a behaviorist approach a researcher might examine the effect that aspects of context such as setting and participants have on the use of inflectional morphemes. An interactionalist approach would view the use of vocabulary as being the result of the individual's internal vocabulary system under the influence of external contextual factors.

Henriksen (1999) emphasizes the trait dimension of vocabulary knowledge in a three-part definition of vocabulary knowledge. The first component is partial-precise knowledge, which refers to basic versus deeper knowledge of the meaning of words. The meaning of a word can be described in terms of what it denotes and how it relates to other words. The relationship of a word to others is sometimes referred to as sense relations, and three important such relations are synonymy, antonymy, and hyponymy (McCarthy, 1990). The more that an individual knows of these relations relative to a particular word, the greater is
his or her understanding of its meaning. Read (2000) found that there were three main types of associated words: paradigmatic, syntagmatic, and analytic. Paradigmatic relations refer to words in the same lexical set or lexical field as the test word. For example, the lexical set of animals might include “dog,” “cat,” and “bear,” and these words would be co-hyponyms. The word “animal” is a more general category of word semantically and is termed a superordinate. Read also includes synonyms in this class. Syntagmatic relations refer to words often used together or collocations, and analytic associates are words that include at least one component of the target word’s meaning (e.g., export-overseas).

Henriksen’s second element of vocabulary knowledge is depth knowledge, and this includes aspects of word knowledge other than the basic meaning such as the word characteristics in Chapelle’s definition above. The final part of Henriksen’s definition is receptive and productive word knowledge. Nation (2001) also emphasizes the importance of meaning and depth knowledge, and the receptive-productive dimension, but adds a use component.

Measurements of vocabulary knowledge that focus on the basic form-meaning link are often obtained with the use of size (or breadth) tests. Statements about an individual’s vocabulary size tell us for how many words the learner can associate a word form and the word’s primary meaning. Examples of tests that assess this type of knowledge are ones that match words and their definitions, and word translation tests. However, while form-meaning word knowledge is necessary, it is widely accepted that vocabulary knowledge is multidimensional, and not limited to what is measured by size tests. The testing of these additional dimensions of word knowledge, or depth testing, has tended to be less common than size testing (Richards & Malvern, 2007). Depth tests can tell researchers more about finer shades of meaning knowledge and a learner’s ability to use the target vocabulary. Testing for depth knowledge in addition to size knowledge is especially important because it has been suggested that increases in size knowledge are not automatically accompanied by increases in depth knowledge (Zheng, 2011).

While the construct of depth knowledge is not always evident (Read, 2004), Schmitt (1999) writes that two approaches have been used to assess this type of vocabulary knowledge, one that focuses on development of knowledge and one that assesses dimensions of depth knowledge. Developmental approaches seek to measure the change in word knowledge over time while dimensional approaches specify the component knowledge associated with depth knowledge and try to measure it. For example, Nation (2001) suggests that in addition to basic form-meaning word knowledge, knowing a word deeply includes knowledge of associated words and more precise knowledge about the limits of the concept a word refers to, as well as word use components such as grammatical functions, collocations, and constraints on use.
2.2. Use-Related Vocabulary Knowledge

Nation (2007) cites Read’s (2000) view that a comprehensive assessment of an individual’s vocabulary ability ought to include a measurement of word use. Supporting this position, Schur (2007) found that while a group of Chinese English learners had a very large word size, they experienced great difficulty in understanding and producing English for the purposes of communication. One contribution of Chapelle’s (1998) analysis is to highlight the importance of vocabulary use. According to one definition, to “use” vocabulary means to deploy words with the intention to communicate a meaningful message as opposed to displaying vocabulary knowledge in, for example, a classroom context (Richards, Platt, & Platt, 1992). This definition of use focuses on only speaking and writing, which is in accord with Read’s (2000) view in which reading and listening are called comprehension. Other writers (Bygate, 1987; Nation, 2001) take a broader view of use which includes both the productive and receptive macro-skills.

With regard to measuring vocabulary use, Nation (2001) recommends examining the learner’s ability to use target words grammatically correctly, word collocations, and constraints on word use such as those affected by cultural differences and register. He writes that collocation is local context information. This lower-context information, though important in using words, provides less evidence of the ability to use words than is suggested by some writers. Behaviorists, according to Chapelle (1998), focus their research on vocabulary used at the discourse level because it is at this level that a test taker’s ability to use vocabulary in particular contexts becomes more reliably apparent. However, Read (2000) notes that vocabulary use tests that entail comprehending or producing discourse tend not to be favored by researchers as they require a great deal of time and often include test conditions that limit test validity such as guessing the meaning of unknown words from the context of the test item (also Schmitt, 1999). This presents a conundrum as the discussion thus far has advocated the testing of vocabulary use in addition to underlying trait characteristics. One way to contextualize vocabulary to a higher degree than collocation tests yet avoid prohibitively time-consuming test items composed of extensive discourse that can provide unwanted answer hints to the test taker is to use single-sentence length items. Nation, for example, suggests a number of sentence-related activities for developing vocabulary use (e.g., matching sentence halves, ordering words in a sentence) that could be used as written test items to test whether learners can deploy grammatical knowledge related to words. However, while it may be practical to test vocabulary use at the sentence level, doing so limits the inferences that can be made about the test taker’s use of words to those levels. If one wished to make inferences regarding use and wider contextual concerns such as register, the test item would need to provide more contextual information for the test taker to respond to (Read & Chapelle, 2001).
The above discussion reinforces the need to define and measure vocabulary as a multidimensional construct. One way to envisage vocabulary knowledge is using a continuum that includes basic word-meaning knowledge near one extreme and other, later acquired, aspects of word knowledge such as deeper meaning knowledge, lower-context use knowledge, and higher-context use knowledge further along the continuum (Read, 2004; Schmitt, 2010). However, in order to locate these other aspects of word knowledge on the continuum they would have to be measured relative to each other. To do this, the study assesses basic form-meaning knowledge using size tests, deeper meaning knowledge with word association tests, lower-context use knowledge with tests of collocation, and higher-context use knowledge with sentence writing tests.

2.3. Receptive and Productive Recall Vocabulary Knowledge and Use

A number of researchers have explored the differences between receptive and productive vocabulary knowledge (Laufer et al., 2004; Makarchuk, 2010; Melka, 1997; Mondria & Wiersma, 2004; Nation, 2001; Webb, 2008). Receptive vocabulary is vocabulary used in the process of listening and reading, and in psycholinguistic terms involves a movement from the received word form to its meaning. This knowledge can be assessed by providing L2 word prompts and asking the test taker to supply the appropriate meaning or L1 translation. Productive word knowledge is that which is used in speaking and writing. To operationalize these skills the learner begins with a meaning to be expressed and proceeds to the correct word form. Productive vocabulary knowledge can be assessed by asking learners to provide evidence of L2 word form knowledge when given a meaning or an L1 word prompt. Assessment of the relative amounts of these types of vocabulary knowledge is important because it provides researchers with information about how vocabulary is learned and has implications for vocabulary use. Even if one accepts, after Read (2000) above, that size tests provide very limited evidence of a learner’s ability to use a known word, it seems only sensible to acknowledge that if a learner’s receptive vocabulary size were much greater than his or her productive size the skills of reading and listening would likely be superior to those of speaking and writing (Schmitt, 2010).

Research results have varied on comparisons of learners’ receptive and productive vocabulary sizes. According to Melka (1997), some L2 studies support the view of receptive vocabulary being much larger than productive while others report more similar sizes. Supporting the latter view, Webb (2008) found that among Japanese university students receptive vocabulary size was only slightly larger than productive size on form-meaning translation tests. One concern regarding some studies (Laufer, 1998; Laufer & Paribakht, 1998; Waring, 1997) that found receptive vocabulary to be higher than productive was the use of versions of the Vocabulary Levels Test. The receptive version of
this test is a recognition test which allows test takers to choose among a number of possible correct answers, whereas the productive test requires that test takers recall the correct answer; this difference can inflate receptive vocabulary scores (Stewart & White, 2011; Webb, 2008). This concern suggests a need for further research on learners’ receptive versus productive vocabulary size using a research design that makes more valid comparisons. In addition to the importance of exploring receptive and productive size knowledge, it is necessary to know more about how other aspects of vocabulary ability, such as collocation, progress along the receptive-productive continuum, as little research has been conducted in this area (Schmitt, 2010). With regard to measuring vocabulary use, Schmitt suggests that recall tests are essential as interpersonal communication requires that the form-meaning link be known at the recall level, not just the recognition level. To measure receptive vocabulary use, Nation (2001) recommends that learners be able to identify collocates and grammatically correct sentences related to test words, and to demonstrate productive use learners should be able to provide collocates and grammatically accurate sentences.

3. RESEARCH QUESTIONS

The study has three inter-related research foci that have been developed to address questions raised in the literature.

1. What are the participants’ receptive and productive vocabulary sizes at the 1,000 and 2,000 word frequency levels?
2. How much deeper meaning vocabulary knowledge accompanies basic form-meaning vocabulary knowledge when vocabulary knowledge is considered receptively, productively and on a combined basis?
3. To what extent is use-related vocabulary knowledge associated with form-meaning vocabulary knowledge when vocabulary knowledge is considered receptively, productively and on a combined basis?

4. METHOD

4.1. Participants and Context

The participants in this study consisted of 169 freshmen from intact classes in 24 majors studying at a South Korean university. Each participant completed the size tests and a
background questionnaire, while 51 of the 169 students were randomly selected to also undergo depth testing. There were 88 males and 81 females who took the size tests, and a similar ratio of males to females was selected for the depth tests with each of the majors being represented in accordance with their proportional contribution to the overall study sample. Ninety-six percent of the students were between 19 and 21 years of age, and at the time the research was conducted they had studied English in educational institutions for 8.8 years on average. According to the participants, their English education before entering university had focused largely on the sub-skills of vocabulary (24.5%) and grammar (16.9%), and the receptive macro-skills of reading (24.4%) and listening (16.0%). Only 7.6% and 6.3% of their educational time had been devoted to writing and speaking development, respectively, while pronunciation practice accounted for the final 4.3% of the time. Because of the participants' extensive background in EFL vocabulary acquisition, they were deemed an appropriate group for the purposes of the study.

4.2. Procedure

In the first week of the semester, the participants were given the background information questionnaire and the first of two vocabulary size tests. The second size test, and an information and consent form were administered in the second week of the semester. In the eighth week of classes, students began taking the depth tests. To measure deeper meaning vocabulary knowledge and vocabulary use-related knowledge, the study assessed components based on Nation's (2001) typology, all of which were tested receptively and productively. The deeper meaning tests measured synonym, antonym, and lexical field knowledge, and the use tests consisted of collocation tests and tests of grammatical use. All of the tests assessed recall knowledge of words known in translation on both the receptive and productive size tests. Thus, unique depth tests were created for each of the 51 participants based on their performances on the size tests. The tests included equal numbers of words from the 1,000 and 2,000 words bands, and a counter-balanced design was used to assuage threats to test validity. Finally, the test instructions were written in the participants' L1 to promote test purpose clarity, and the tests were trialed to check for typographical and other errors that might have led to test taker confusion. Each student's depth test battery was administered in a one-hour session, and the testing was conducted over a five-week period.

4.2.1. Vocabulary size tests

Two 80-item size tests were administered during regular class times to assess form-meaning vocabulary knowledge. The 160 test items consisted of 80 unique words, which
were tested both receptively and productively. A counter-balanced test design was used to compensate for possible score inflation in the form of a learning effect due to meeting a test word twice, once when tested receptively and again on the productive test. Thus, one half of the test words were measured first receptively then productively, while the other half were assessed productively first and receptively second. The same words were tested twice to avoid threats to validity caused by the possible differences in difficulty of words even when drawn from the same word frequency level. Of the 80 test words, 40 were chosen from the midrange of the first 1,000 word frequency level of the British National Corpus (BNC) and 40 from the midrange of the second 1,000 most frequently used words. A lemmatized frequency list (Kilgarriff, 1995) was used instead of a word family list, as word family lists include both inflections and derivatives as knowledge known along with a headword, an assumption that has been questioned (Schmitt, 2010). Only knowledge and use of words at the 1,000 and 2,000 word frequency levels were measured as the purpose of the study was not to determine general language proficiency, but to compare aspects of vocabulary knowledge. In addition, previous vocabulary size testing (Makarchuk, 2010) with more advanced learners in a similar population revealed that they knew only 46.8% of the words both receptively and productively at the 3,000 word level. As a result, it was considered unlikely that sufficient useable responses would be obtained using the 3,000 word level to justify including words at this level.

To measure word size knowledge, Laufer and Goldstein’s (2004) item design was adapted to the Korean language context. An example of a productive test item is as follows: \( \text{city} \) \( c \) ____. This is a test of productive word knowledge because the test taker is given the L1 word and asked to produce the L2 translation (i.e., city). According to Laufer and Goldstein, the first letter of the answer is provided to prevent test takers from answering with words that are synonyms of the word being tested. This is necessary because if a synonym were selected, the test taker could answer the item correctly without providing information regarding knowledge of the target word. To measure receptive knowledge, Laufer and Goldstein’s design template was also used; for example, \( \text{city} \) A. \( c \) ____ B. _____. This item tests receptive knowledge because the test taker is presented with the target language word, in this case an English word, and is asked to supply the L1 translation (i.e., \( \text{city} \)). As was the case with the productive test, the first letter of the answer is provided. This item type was modified to allow test takers to answer with a synonym of the L1 answer as the test is intended to measure only knowledge of the English word, not particular L1 words. With regard to construct validity, receptive knowledge is measured by test items that provide the L2 word form and elicit its meaning or L1 translation, and productive knowledge is assessed by items that prompt the test taker with an L1 word or its meaning and require the L2 translation. As these procedures match definitions of receptive and productive knowledge in the literature (Laufer et al., 2004;
Nation, 2001), construct validity is supported. It should be noted that both the receptive and productive test items described above test recall knowledge in that test takers are not provided with possible answers to choose from, as is the case with recognition tests. This feature of test design eliminates validity concerns that have been raised with regard to studies that have based their results on a mix of recall and recognition tests. Finally, the reliability of tests based on this test design was measured in a trial with 74 learners similar to those who participated in the study (Makarchuk, 2010). Using the split-half method, reliability estimates of 0.8995 for the receptive test and 0.8933 for the productive test were found. This was judged to be sufficient based on work by Schmitt, Schmitt, and Clapham (2001), which found a reliability level of 90% acceptable for vocabulary tests.

4.2.2. Word association tests

The tests of deeper meaning included 16 items that measured synonym, antonym, and lexical field knowledge related to words that were known receptively and productively on the size tests. The tests were counter-balanced to compensate for meeting a word twice, and eight words were tested at the 1,000 and 2,000 word frequency levels. For the receptive test, two possible English synonyms, antonyms, and paradigmatic words, for a total of 6 sub-items, were given for each test item (Appendix A). A combination of correct and incorrect sub-items was given with both types being drawn from the same frequency word level as the test word. The test takers had to decide whether or not each sub-item was a synonym, antonym or paradigmatic word relative to the given test item. This is a receptive recall test because it is believed that the test takers would have to recall the meaning or L1 translation of the given L2 words and then search for meaning associations with the known test prompt in their mental lexicons. For the productive test, the test takers were given the English test word and asked to recall and write six English words matching the sub-items types described above. The English word form was used as a prompt because the size tests had already demonstrated that the test takers had linked the L2 and L1 word forms. This is a productive recall test because the test takers were required to recall L2 words associated with the L1/L2 word form complex.

4.2.3. Use-related vocabulary knowledge tests

Two use-related knowledge tests were administered to better understand the test takers’ ability to use vocabulary. One test examined collocation knowledge and the other assessed the participants’ knowledge of grammar related to the test word. Both tests were given in receptive and productive versions with a counter-balanced design that tested 16 words, eight at the 1,000 word frequency level and eight at the 2,000 word level. Once again, the
words were chosen from among the ones each participant had answered correctly on both the receptive and productive size tests.

The receptive collocation test (Appendix A) supplied the test takers with three possible L2 collocations for each known test word, some of which were correct and others incorrect. Correct collocations were drawn from the *Oxford Collocations Dictionary* (McIntosh, Francis, & Poole, 2009) and non-collocations from the same word frequency level as the test word. The productive test gave the test takers 16 L2 test words and asked them to recall and write three L2 collocations for each.

The receptive test of the ability to use grammatical knowledge provided the test takers with 16 L2 sentences that included one test word each. Some of the sentences had grammatical errors and some were correct. The test takers were asked to underline the errors in the sentences. The productive test provided only the 16 test words and asked the participants to write a grammatically correct sentence for each one. Neither test type provided possible answers, but rather required the participants to recall knowledge from memory. The psycholinguistic justification for identifying these tests as receptive and productive recall tests is the same as that for the word association tests above.

5. RESULTS AND DISCUSSION

To investigate research question one, the study measured the participants’ receptive and productive recall form-meaning vocabulary knowledge at the 1,000 and 2,000 word frequency levels.

5.1. Form-Meaning Tests

First, total receptive and productive size was 70.2% at the 1,000 and 2,000 word levels combined with word knowledge falling from 82.2% to 58.1% as word frequency decreased (Table 1). Total receptive word size was 76.2% and total productive word size was 64.1%, a 12.1% greater receptive word size. To check for the statistical significance of this difference, a Wilcoxon Signed Ranks Test was used, as the data were collected using a non-random sampling procedure (intact classes). The difference between total receptive and productive vocabulary knowledge was statistically significant \(Z = -17.040, p = 0.000\). At the 1,000 word level receptive knowledge was 10.3% greater than productive knowledge, and at the 2,000 word level the difference in favor of receptive knowledge was 13.9%. The differences at both the 1,000 and 2,000 word levels were statistically significant (1,000: \(Z = -12.385, p = 0.000\); 2,000: \(Z = -11.908, p = 0.000\)). These findings support previous research (e.g., Melka, 1997) that has claimed that second language
learners tend to have more receptive vocabulary knowledge than productive knowledge, but not researchers who claim the difference to be large. Rather, the total knowledge difference of 12.1% is supportive of researchers (e.g., Webb, 2008) who found EFL learners’ receptive word knowledge to be slightly to somewhat larger than their productive vocabulary knowledge.

### TABLE 1

<table>
<thead>
<tr>
<th>Word Frequency Level</th>
<th>1,000</th>
<th>2,000</th>
<th>Total Mean (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Receptive (R)</strong></td>
<td>17.5/87.3/2.02</td>
<td>13.0/65.0/3.51</td>
<td>76.2</td>
</tr>
<tr>
<td><strong>Productive (P)</strong></td>
<td>15.4/77.0/2.66</td>
<td>10.2/51.1/3.19</td>
<td>64.1</td>
</tr>
<tr>
<td>% Difference (R - P)</td>
<td>10.3</td>
<td>13.9</td>
<td>12.1</td>
</tr>
<tr>
<td>% Mean (R + P)</td>
<td>82.2</td>
<td>58.1</td>
<td>70.2</td>
</tr>
</tbody>
</table>

The results suggest that the participants would face challenges functioning in English as they knew only 70.2% of the words receptively and productively at the 1,000 and 2,000 word levels combined, and this would make the guessing of unknown words problematic (Nation, 2001). Moreover, they would be especially challenged by words at the 2,000 word frequency level as only 58.1% of them were known receptively and productively on average. The degree to which the learners would struggle with English communication would also be determined by how well they knew the vocabulary beyond basic form-meaning knowledge.

### 5.2. Deeper Meaning Tests

To answer research question two, the depth tests of deeper meaning knowledge measured receptive and productive knowledge of synonyms, antonyms, and paradigmatic relations (superordinates and co-hyponyms) for words known both receptively and productively at the 1,000 and 2,000 word frequency levels on the size tests. For the productive test, answers were marked according to English native speaker benchmarks because of the variability of the test words with regard to their sense relations. For example, there may have been some words such as “bed” and “traffic” for which it was difficult to produce antonyms. To arrive at these benchmarks, three English native speaker university instructors of EFL with master’s level academic qualifications and an average of 15 years’ teaching experience completed the test items.
The deeper meaning depth tests found that test takers correctly answered 70.7% of the test items related to knowledge of synonyms, antonyms, and paradigmatic words for words known in translation when the receptive and productive results were combined (Table 2). However, receptive word knowledge was 43% higher than productive knowledge with test takers answering 92.2% of the items correctly as compared to 49.2% for the productive items. This difference was statistically significant ($Z = -8.734, p = 0.000$).

**TABLE 2**

<table>
<thead>
<tr>
<th>Word Frequency Level</th>
<th>1,000</th>
<th>2,000</th>
<th>Mean (%)</th>
<th>Mean SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Receptive (R)</td>
<td>94.6/6.5</td>
<td>89.7/8.3</td>
<td>92.2</td>
<td>7.81</td>
</tr>
<tr>
<td>Productive (P)</td>
<td>50.1/16.6</td>
<td>48.2/17.7</td>
<td>49.2</td>
<td>17.1</td>
</tr>
<tr>
<td>% Difference (R - P)</td>
<td>44.5</td>
<td>41.5</td>
<td>43.0</td>
<td>NA</td>
</tr>
<tr>
<td>% Mean (R + P)</td>
<td>72.4</td>
<td>69.0</td>
<td>70.7</td>
<td>NA</td>
</tr>
</tbody>
</table>

Receptive meaning knowledge was strong at both the 1,000 and 2,000 word frequency levels. At the 1,000 word level, participants answered correctly 94.6% of the time, while at the 2,000 word level 89.7% of their answers were correct—a small but statistically significant difference ($Z = -4.867, p = 0.000$). There was not a significant difference between the productive scores at the two word frequency levels with an average score of 49.2%. The results suggest that the participants had very high levels of more precise meaning knowledge of words known in translation when such knowledge was received by them, as on receptive tests. However, when asked to produce synonyms, antonyms, and paradigmatic words, they were able to do so for only approximately half of the test items. Based on these findings, it would seem likely that the participants would have more success on tasks in which English is presented to them, as is the case when reading and listening, than on writing and speaking tasks that require the production of language.

Next, to assess the participants’ ability to use vocabulary knowledge receptively and productively, research question three, the study tested the use of words known receptively and productively on the size tests for collocation knowledge and accurate grammatical use.

5.3. Use-Related Vocabulary Knowledge Tests: Collocation Tests

The collocation tests were administered in both receptive and productive formats. The receptive tests provided possible English collocations for test takers to associate with the
test word, and the productive test asked them to write collocations of the provided test word. For the productive test, it was necessary to once again use English native speaker benchmarks as some words were easier to produce collocations for than others. The benchmarks were based on the scores of the three native speaker university instructors described above. The results in Table 3 indicate that the participants knew 67.3% of the possible collocations associated with the test words when receptive and productive knowledge were combined. Receptive collocation knowledge was 13% higher (73.8%) than productive knowledge (60.8%), a statistically significant difference ($Z = -4.546, p = 0.000$). At the 1,000 word frequency level, a receptive score of 71.1% was found and a productive score of 65.2%. This 5.9% difference in collocation knowledge did not reach statistical significance, but the difference in knowledge at the 2,000 word level of 20% (receptive = 76.4%; productive = 56.4%) was statistically significant ($Z = -4.598, p = 0.000$).

**TABLE 3**

<table>
<thead>
<tr>
<th>Word Frequency Level</th>
<th>Mean (%)</th>
<th>Mean SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,000</td>
<td>71.1/11.3</td>
<td></td>
</tr>
<tr>
<td>2,000</td>
<td>76.4/11.3</td>
<td></td>
</tr>
<tr>
<td>Receptive (R)</td>
<td>73.8</td>
<td>11.5</td>
</tr>
<tr>
<td>Productive (P)</td>
<td>65.2/24.8</td>
<td>25.5</td>
</tr>
<tr>
<td>% Difference (R - P)</td>
<td>5.9</td>
<td>20.0</td>
</tr>
<tr>
<td>% Mean (R + P)</td>
<td>68.2</td>
<td>66.4</td>
</tr>
</tbody>
</table>

The results of the collocation tests suggest a 67.3% accuracy rate for syntagmatic knowledge accompanies translation knowledge when the receptive and productive scores are combined. The higher level of receptive knowledge suggests that reading and listening tasks would be more strongly supported than writing and speaking tasks. Counter-intuitively, receptive collocation knowledge was statistically significantly higher at the 2,000 word frequency level than the 1,000 word level ($Z = -2.827, p = 0.005$) as opposed to productive knowledge, which was higher at the 1,000 word level than the 2,000 word level ($Z = -3.007, p = 0.003$). As one would expect that the greater exposure to more frequently appearing words would promote syntagmatic associations, it is unclear why the words tested receptively at the 2,000 word level produced a higher collocation score than those at the 1,000 word level.
5.4. Use-Related Vocabulary Knowledge Tests: Grammatical Use Tests

The receptive grammatical use test asked the participants to locate possible grammatical errors in English sentences, while the productive test instructed them to write a grammatically accurate English sentence that included the test word. To score the receptive test, both unidentified and misidentified errors were counted. That is, if test takers failed to find the error in a sentence or indicated that a correct sentence was wrong, the item was marked incorrect. For the productive test, only errors related directly to the test word were counted. For example, in the sentence “I received a letter from a very close friends,” if the test word were “letter,” the item would be marked correct because the test word is used in a grammatically correct clause with the only error being in the related prepositional phrase.

| TABLE 4 |
| Mean Receptive and Productive Recall Grammatical Use Test Scores |
| Word Frequency Level | 1,000 | 2,000 | Mean (%) | Mean SD |
| (%)/SD |
| Receptive (R) | 56.6/23.4 | 43.9/19.6 | 50.3 | 22.4 |
| Productive (P) | 63.7/23.0 | 48.3/23.1 | 56.0 | 24.2 |
| % Difference (R - P) | -7.1 | -4.4 | -5.7 | NA |
| % Mean (R + P) | 60.2 | 46.1 | 53.2 | NA |

The results of the grammatical use tests (Table 4) found that the test takers were able to accurately use the target vocabulary in sentences only 53.2% of the time when the productive and receptive test scores were averaged. While the raw data suggest that they were able to produce sentences somewhat more correctly than identify inaccuracies in sentences (productive = 56.0%; receptive = 50.3%), this difference did not reach the 95% probability level of statistical significance (Z = -1.905, p = 0.057). Similarly, there were no significant differences in results at the 1,000 or 2,000 word frequency levels when the receptive and productive scores were compared. However, it was found that the participants were better able to produce accurate sentences using words from the 1,000 word frequency level both receptively and productively as compared to 2,000 level words; these differences were statistically significant (receptive [1,000 vs. 2,000]: Z = -3.348, p = 0.001; productive [1,000 vs. 2,000]: Z = -3.915, p = 0.000). These results suggest that the participants would be successful using words known in translation grammatically correctly only about one half of the time, and that this success rate would be the same for both productive and receptive tasks. They would, however, enjoy 14.1% more success when
working with words at the 1,000 word frequency level as compared to using words at the 2,000 word frequency level. Furthermore, this difference in use-related knowledge supports the view of vocabulary knowledge development being incremental in nature as higher frequency words were better known grammatically than lower frequency words.

5.5. Vocabulary Knowledge Continuum

An analysis of the vocabulary knowledge results suggests that when learners have learned receptive and productive form-meaning word knowledge, they also know a substantial amount of deeper meaning knowledge, as they answered 70.7% of the test items correctly (Table 2). However, this knowledge is primarily receptive as the participants had a high level of receptive deeper meaning knowledge (92.2%), but a relatively low level of productive deeper meaning knowledge (49.2%). The results indicate that the acquisition of productive deeper meaning knowledge considerably lags receptive knowledge.

When the use-related knowledge test scores were averaged, it was found that only 60.3% of the words known receptively and productively on the size tests could also be used. In addition, this use amount was not significantly different for receptive and productive use, which were 62.0% and 58.4%, respectively ($Z = -1.891, p = 0.059$). Intuitively, one would have thought that the greater receptive deeper meaning knowledge would have facilitated higher scores on the receptive use tests, as it is use-enabling knowledge. The reason this did not occur may be due to the limited context of the use tests. Perhaps if paragraph-length use test items had been used, with their attendant challenges, knowledge of synonyms, antonyms, superordinates, and co-hyponyms of the test word might have led to higher receptive use scores. The results of the use tests suggest that receptive and productive vocabulary use remain highly similar despite the much greater focus on receptive skills in the participants’ educational backgrounds. Of interest with regard to the vocabulary knowledge continuum is that deeper meaning knowledge overall was greater at 70.7% than use-related knowledge at 60.3%, a statistically significant difference ($Z = -5.283, p = 0.000$). This finding supports the hypothesis of vocabulary learning being a developmental process with knowledge such as form-meaning knowledge being acquired before deeper meaning knowledge, which in turn is learned before use-related knowledge. In psycholinguistic terms, this incremental development of knowledge would be explained by the increasing difficulty of acquiring the different types of knowledge. Form-meaning knowledge requires straightforward knowledge of a concept-form link, whereas deeper meaning knowledge can add the burden of a logical operation, as sometimes happens in synonymy and antonymy (Aitchison, 1994). Finally, use-related knowledge involves knowledge of individualistic language conventions in the case of
collocations and grammatical knowledge for sentence-length items.

However, when receptive and productive deeper meaning and use-related knowledge were compared separately, it was found that the continuum hypothesis held for receptive knowledge but not productive knowledge. Specifically, total receptive deeper meaning knowledge (92.2%) was significantly higher than total receptive use-related knowledge (62.4%) \((Z = -8.762, p = 0.000)\), a significant difference which held for both the 1,000 and 2,000 word frequency levels. However, total productive use-related knowledge (58.4%) was significantly higher than total deeper meaning knowledge (49.2%) \((Z = -4.270, p = 0.000)\), and this significant difference was also found at both the 1,000 and 2,000 word levels. Furthermore, there was not a significant difference between the two use components, collocation and grammatical use, for productive knowledge, which strengthens the argument that productive use-related knowledge is higher than productive deeper meaning knowledge.

This finding highlights the difficulty the participants had producing word associations as opposed to discerning associations when presented with an L2 word. A possible explanation for this result is that when producing associated words the learner may search for them in the L1 lexicon and then experience difficulty translating the word to the L2 due to the greater complexity of L1 lexicon knowledge. Alternatively, the learner may move from the L1 lexicon to the L2 lexicon, and unsuccessfully search for associated words in this presumably less extensive word network. However, when receiving words the learner could move from the L2 word to the L1 lexicon, which because of its more extensive interconnectedness would promote more successful word association. This argument is supported by the collocation test results in that the participants were less successful with the receptive collocation items (73.8%) than the receptive word association items (92.2%), and more successful with the productive collocations items (60.8%) than the productive word association items (49.2%). The receptive collocation items would have been more difficult because if the learner moved from the L2 lexicon to the L1 lexicon collocations would be those of the L1 lexicon, which would not necessarily match the conventions of the L2 lexicon. Associated words (e.g., synonyms and antonyms), on the other hand, would likely be similar in both lexicons as they are based on logical relations not conventions of language. Productive collocation knowledge would be easier to obtain than productive associated word knowledge because the learner would move from the L1 lexicon to the L2 lexicon where collocations would be closely linked due to frequent proximal co-occurrence, a phenomenon less common with associated word knowledge. On the other hand, if the test taker were to seek productive associated words first in the L1, and then translate to the L2, the greater complexity of the L1 would increase the likelihood of selecting L1 words without L2 translations.

Also of importance with regard to the vocabulary knowledge continuum is that when the
two use-related knowledge test results were compared (Tables 3 and 4), it was shown that the collocation tests resulted in statistically significantly higher use scores than the grammatical use tests, 67.3% compared to 53.2%, perhaps reflecting the greater contextualization of the grammar tests ($Z = -6.053, p = 0.000$). That is, less cognitive demand is likely to have been placed on the participants when answering the single-word collocation items as compared to the sentence-length grammatical use items due to the greater linguistic complexity of the sentences.

6. CONCLUSION

This study examined the receptive and productive recall vocabulary size of 169 EFL students, and the receptive and productive recall depth knowledge of 51 of the 169 study participants in order to compare aspects of word knowledge along a continuum from basic form-meaning knowledge to deeper meaning knowledge to use-related knowledge. First, with regard to the form-meaning vocabulary knowledge results, total receptive and productive size was 70.2% at the 1,000 and 2,000 word levels combined. In addition, receptive vocabulary knowledge was found to be 76.2% and productive, 64.1%, a difference of 12.1%. These results are supportive of researchers who claim a relatively small advantage for receptive versus productive vocabulary in EFL contexts. On the deeper meaning tests, which measured knowledge of synonyms, antonyms, and paradigmatic words, the test takers correctly answered 70.7% of the test items. However, receptive knowledge was much higher than productive at 92.2% and 49.2%, respectively, which suggests that receptive associated semantic word knowledge on average is learned before productive knowledge. The use-related knowledge test results showed a 67.3% accuracy rate on the collocation tests when the receptive and productive scores were combined, and a 13.0% advantage for receptive knowledge when the two test types were compared. The grammatical use tests resulted in a 53.2% accuracy rate when the receptive and productive scores were averaged for the two word frequency levels, and the results were essentially the same for both the receptive and productive tests. Based on the averaged use-related knowledge test results, it was suggested that it is inappropriate to equate form-meaning word knowledge with the ability to use a word. This is because the study found that only 60.3% of words known on the size tests were correctly used on the use tests, and that this level of use was similar for both receptive and productive test items. Finally, the test results suggest support for a vocabulary knowledge continuum in that higher amounts of deeper meaning knowledge were found compared to use-related knowledge when the receptive and productive scores were averaged. It was argued that the more cognitively demanding nature of use-related knowledge may have been responsible for this difference, as
suggested by the higher scores for the lower context use tests as compared to the higher context tests. However, though this finding held for receptive knowledge when the two types of vocabulary knowledge were compared individually, it was found that productive use-related knowledge was greater than productive deeper meaning knowledge. It was suggested that this difference may have been caused by the differing L1 and L2 lexicon transitions related to receptive and productive knowledge, and this, in turn, suggests that a model of the vocabulary knowledge continuum would benefit from considering receptive and productive knowledge acquisition separately.

The study results emphasize the difference between form-meaning vocabulary knowledge, deeper meaning knowledge, and use-related word knowledge, and caution against assuming that knowledge of a word’s form and meaning implies that the word can also be used in everyday communication.

7. LIMITATIONS AND FUTURE RESEARCH

The main limitation of this study resides in the generalizability of the findings. Because the research was conducted with participants from only one university in South Korea, it may be that the findings are site specific and may not apply to other contexts. Additionally, while the study aimed to measure a large sample of the student population in order to strengthen the validity of the findings, it should be noted that the 169 participants who completed the vocabulary size tests were drawn from majors based on convenience of access, and that the participants came from intact classes. These two factors again limit the generalizability of the results as it may be that a study conducted with students drawn randomly from a larger number of majors would produce different findings. In order to address these concerns, it is suggested that it would be beneficial to conduct future studies in a variety of contexts with participants randomly selected from a wider range of majors.

REFERENCES

Cambridge University Press.


Mokhtar, A. (2010). Vocabulary knowledge of adult ESL learners. *English Language...


**APPENDIX A**

Receptive Depth Tests: Meaning and Collocation Items

Instructions: (A) Put a checkmark (√) beside words that have the same or opposite meaning as the test word and words that are in the same meaning category as the test word (e.g., vehicle-car). (B) Put a checkmark beside words that are commonly used near the test word in a sentence. If the words are not the same or opposite, or not in the same meaning category or are not used with the test word in a sentence, do not write anything. (A) 주어진 문제의 단어와 동의어, 반의어 또는 같은 범주 (예: 차 ↔ 자동차)에 속하는 단어 옆에 (√) 표 하시오. (B) 문제의 단어와 연관 하여, 문장 속에서 자주 함께 인용되는 단어 옆에도 같은 표시를 하시오. 그러나 답이 아닌 단어에는 아무런 표시도 하지 마십시오.

**Example: night (문제 보기: 밤)**

(A) Words with the same or opposite meaning, or in the same meaning category. 문제의 단어와 동의어, 반의어 또는 같은 범주에 속하는 단어 옆에 (√) 표 하시오.

(same meaning) (동의어) (i) tree ___ (ii) animal ___

(opposite meaning) (반의어) (i) day √ (ii) black ___

(same meaning category) (동의어의 범주) (i) dusk √ (ii) evening √

(B) Words used together in sentences. 문장 속에서 함께 자주 인용되는 단어에 (√) 표 하시오.

(i) last √ (ii) red ___ (iii) sleepless √
Applicable levels: Tertiary, K-12

Donald W. Makarchuk
Department of English Language and Literature
Kyonggi University
San 94-6, Iui-dong, Yeongtong-gu, Suwon-si
Gyeonggi-do 443-760, Korea
Phone: 031-249-9144
E-mail: dmak@kgu.ac.kr

Received in September 2013
Reviewed in October 2013
Revised version received in November 2013