

Exploring L2 learners' vocabulary knowledge of primary and peripheral meanings

AIZAWA, Kazumi

I Introduction

Recently an increasing emphasis has been put on developing learners' lexical competence. As Wilkins (1972, p.111) says, "without grammar very little can be conveyed, without vocabulary nothing can be conveyed." This being the case, researchers direct an increasing amount of attention to vocabulary acquisition in a wide variety of second or foreign language learning settings around the world (Coady & Huckin, 1997; Nation, 2001; Milton, 2009, etc.). For example, a number of studies have been made not only on the development of vocabulary tests to measure learners' receptive knowledge, but also on tracking learners' longitudinal development of vocabulary knowledge.

It is often argued that there are two aspects of vocabulary knowledge: the breadth and depth of word knowledge. Breadth tests of vocabulary provide rough comparative estimates of an individual's overall vocabulary knowledge, while depth tests are defined either in terms of kinds of knowledge of specific words or in terms of the degree of knowledge of such words. However, this metaphor has not been explored empirically enough. We still do not know whether the knowledge of words with multiple meanings should be categorized as width or depth of vocabulary knowledge. This study investigated learners' knowledge of high frequency polysemous words.

II Background

I Vocabulary Tests

Several tests have been developed to measure learners' vocabulary knowledge. Among breadth tests of vocabulary knowledge, the Vocabulary Levels Test (hereafter VLT) (Nation, 2001; Schmitt, Schmitt & Clapham, 2001), the EFL Vocabulary Tests (Meara, 1992) and Eurocentres Vocabulary Size Tests (Meara & Jones, 1988) (hereafter Yes/No test) are commonly used not only for research purposes but also for pedagogical assessment purposes. These tests have two advantages: the test

user can choose the levels of vocabulary frequency band appropriate for the students; and an adequate level of reliability can be established.

However, there are two concerns that suggest that these tests should be treated with caution. The first concern is whether the breadth test of vocabulary knowledge can correctly estimate the number of words a learner knows. A preliminary study (Kosuge, 2003) conducted at a Japanese junior high school compared the vocabulary size computed using the scores of the Vocabulary Size Test (Mochizuki, 2003) and the number of words learners correctly translated. The results showed that the number of known words estimated from the scores of the Vocabulary Size Test were about 10% more than learners actually translated correctly.

The second point is that some studies treat breadth of vocabulary knowledge as equivalent to English proficiency; however, few empirical studies have been conducted to confirm the validity of this assumption.

Along with these issues common to all three of the tests shown in Table 1, there are more specific limitations to each of the breadth tests. In most of the cases, a word has multiple meanings, yet the VLT test only deals with one meaning of a word and does not test other meanings. As for the Yes/No test, it does not ask test takers the meaning of a word. Test takers are presented with a sequence of words and non-words. If they claim to know non-words, this is taken as evidence that they are overstating their vocabulary knowledge and their scores are adjusted accordingly.

The depth tests of vocabulary knowledge are not as popular as breadth tests. As shown in Table 2, the two most popular tests are the Word Association Test (hereafter WAT) (Read, 1993) and the Vocabulary Knowledge Scale (hereafter VKS) (Paribakht & Wesche, 1993). However, each of these tests has limitations. The WAT uses 40 target words and presupposes that test takers know at least one of the meanings of the words. However, in the case of Japanese students, whose vocabulary size is not very large, it is important to verify this supposition. More importantly, if we cannot assume that test takers know all the target words, it is difficult to interpret the test data because it is not altogether clear whether the test actually measures the breadth aspect or the depth aspect of the

Table 1. Breadth tests of vocabulary knowledge

| | Test | Advantage | Disadvantage |
|---|--------------------------------------|---|-------------------------------|
| Nation (1990) Schmitt, Schmitt & Clapham (2001) | Vocabulary Levels Test | 1) Can choose the level appropriate for learners | Tests just one meaning |
| Meara (1992) | EFL Vocabulary Test | | |
| Meara & Jones (1990) | Eurocentres Vocabulary Size Tests | 2) Adequate reliability | Doesn't test word meanings |

Table 2. Depth tests of vocabulary knowledge

| | Test | Advantage | Disadvantage |
|---------------------------|----------------------------------|---|--|
| Read (1993) | Word Association Test (WAT) | 1) Can measure advanced learner's knowledge 2) Adequate validity | 1) Not suitable for lower proficiency students 2) If learners do not know the target words, what does the test measure? |
| Paribakht & Wesche (1993) | Vocabulary Knowledge Scale (VKS) | Can measure breadth and depth | 1) Limited number of words can be tested 2) Is this a test or a profile? |

vocabulary knowledge.

In the case of the VKS, its use in the classroom to measure learners' vocabulary knowledge is also problematic. As Wesche & Paribakht (1996) admit, "the purpose of the VKS is not to estimate general vocabulary knowledge, but rather to track the early development of specific words in an instructional or experimental situation (p.33)." For example, at level 5, the VKS requires students to write sentences using the target words. Read (2000) advises caution about accepting a sentence made by a student as evidence of word knowledge, as advanced learners can compose acceptable sentences without having a good understanding of what the target word means.

However, the most crucial point to be raised here is that there are only a few studies, at most, about which particular aspect of vocabulary knowledge is actually tested by a certain test. To make this point clearer, it is useful here to refer to Ringbom's description (1987) of vocabulary knowledge (Figure 1). He classified lexical knowledge into six sub-elements: accessibility, morphophonology, syntax, semantics, collocation and association. He also postulates that an educated native speaker's knowledge of a word usually means that it is near the very top of the continuum of each dimension, while an L2 learner has to work his/her way up from the bottom.

If we compare the construct that each test measures and Ringbom's description of vocabulary knowledge, it is clear which construct some tests measure, while other tests are rather difficult to categorize by construct (see Table 3). For example, the VLT and the Yes/No test measure "Semantics: Knowing one meaning only" and "Morphophonology: Knowing one form only" respectively. On the other hand, it is not clear which aspects of vocabulary knowledge the WAT and the VKS measure. If the construct is not clear, it is understandably difficult to interpret the test data.

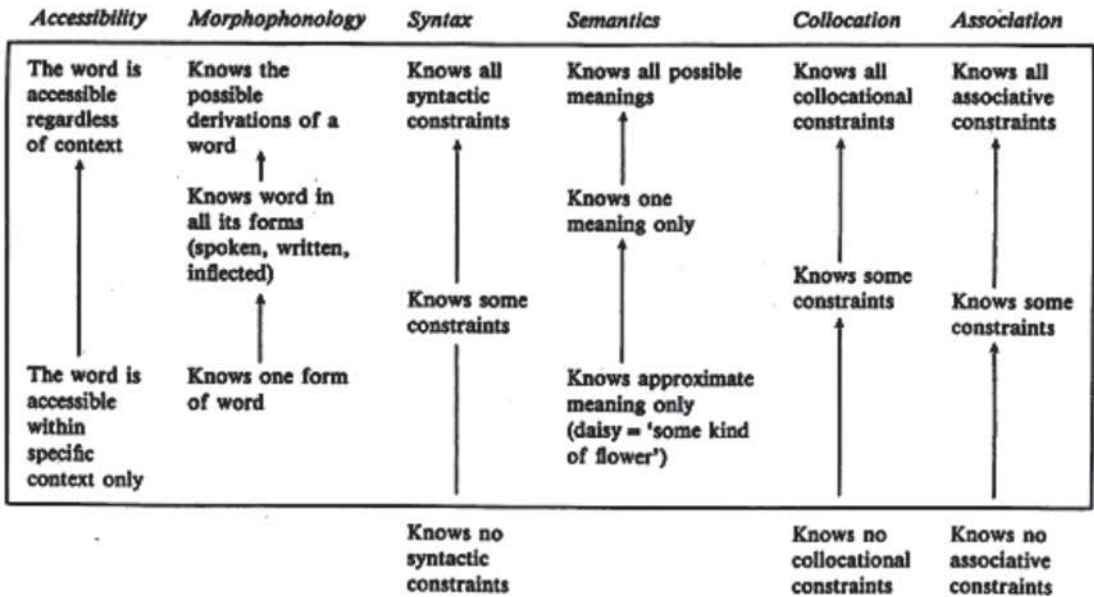


Figure 1. Ringbom’s description of Vocabulary knowledge (1987, p.37)

Reviews of the literature showed that, based upon Ringbom’s model, no tests have been developed to measure the aspects of “Semantics: Knows all possible meanings.” This aspect of vocabulary knowledge seems to be a deeper aspect of the constructs which the VLT and the Yes/No tests measure. They are also very important aspects of vocabulary knowledge from a pedagogical point of view.

2 Defining polysemous words

A polysemous word is a word that has two or more meanings. These multiple meanings can be classified into two categories: (1) a core sense, and (2) a peripheral sense. As Verspoor and Lowie (2003) explain, while there is only one core sense, there are often multiple peripheral senses with different degrees of abstractness in meaning. Core meaning is the most literal meaning or the logical

Table 3. Comparison of test constructs and Ringbom’s dimensions of word knowledge

| Test | Dimensions |
|----------------------------|--|
| Vocabulary Levels Test | Semantics: “Knows one meaning only” |
| Yes/No Test | Morphophonology: “Knows one form only” |
| Word Association Test | Collocation: “Knows some constraints” |
| | Association: “Knows some constraints” |
| Vocabulary Knowledge Scale | Semantics: “???” |
| | ??? |

central application that can be related to the central applications by relatively clear semantic relation. To sum up, the relationship between a core sense and peripheral senses is one of meaning extension.

Based upon the review in the previous sections, there is a question to be raised. It is not clear which contributes more to English proficiency, to know two meanings of one particular word, or to know one meaning of two words respectively. It seems to be very important to clarify whether the knowledge of multiple senses of the polysemous word is the breadth aspect or the depth aspect of knowledge. If the correlation between a learner's scores on a polysemous test and an English proficiency test is high enough, we can assume that a polysemous word knowledge test is more of a vocabulary size test. On the other hand, if the correlation between a learner's scores on a polysemous test and an English test is not as high as scores of a vocabulary size test, polysemous word knowledge should be distinguished from breadth knowledge of words.

Therefore, in this study, two new sets of vocabulary tests were designed to measure these particular aspects of vocabulary knowledge.

III Study

1 Purpose

The purposes of this study were (1) to investigate the relationship of test scores of the primary meaning-based vocabulary size test and those of the English proficiency test, (2) to measure learners' knowledge of primary and peripheral meanings, and (3) to compute the correlation among these scores to find which test score can best predict those of English proficiency tests.

2 Participants

A total of 59 Japanese university students participated in the study on a voluntary basis. They were taking English classes for a partial fulfillment of their course requirements in engineering.

3 Instruments

Vocabulary test: As a measurement of vocabulary knowledge, the first four sub-tests (1K through 4K) of the TDU VLT (Aizawa, 2006) were used (see Figure 2). This set of tests (hereafter VLT) is a paper-pencil receptive vocabulary test that estimates the sizes of overall vocabulary knowledge up until the 4000-frequency band. Thirty randomly chosen vocabulary items were used for that particular frequency band, totaling 120 test items. At each frequency band, three Japanese translations of the target words were presented together with multiple choices of six English words. Test takers were asked to match the three Japanese words with their English equivalents out of six

| | | | | | | | |
|--|--|---|---|---|--|----|--|
| ① art ② light ③ meeting ④ pound ⑤ technology ⑥ year | (1 技術, 工業技術 2 光, 光線 3 会うこと, 集まること) | ① care ② corner ③ leader ④ record ⑤ today ⑥ wife | (16 今日, 現代 17 記録, 登録 18 先導者, 指揮者) | 1 | | 16 | |
| ① face ② house ③ moment ④ purpose ⑤ raise ⑥ remain | (4 家 5 顔, 顔面 6 目的, 意図) | ① argue ② join ③ provide ④ see ⑤ successful ⑥ tend | (19 ~する傾向がある 20 見える, 会う 21 ~を与える, 用意する) | 2 | | 17 | |
| | | | | 3 | | 18 | |
| | | | | 4 | | 19 | |
| | | | | 5 | | 20 | |
| | | | | 6 | | 21 | |

Figure 2. Sample questions of TDU VLT

choices. The maximum score of 120 is the equivalent of the estimated value of 4000 words.

Polysemous words test: Two sets of target words (50 words and 52 words) with multiple meanings (nouns, verbs, and adjectives) were carefully chosen from the most frequent 3000 words of the JACET 8000 word list. The first group of words has at least three distinguishable meanings, while the latter has two different meanings. In order to have a large enough number of words, not only polysemous words but also homonyms were included in this list (*bank, nail, temple, wind, lead, and tear*).

The first task (PT1) asked learners to choose the incorrect definition for each target word from the four L1 definitions given (50 items). If all of the four meanings are correct, participants were asked to choose “all correct.” Therefore, each item has five different choices (A, B, C, D + E). This task tested whether the learners knew the three or four correct meanings of the target words. For example, in Question No.1 “due” has meanings of A, B, and D,” and does not have the meaning of C, therefore learners were expected to answer “C.”

The second task (PT2) asked learners to choose the primary and secondary meanings of the target words in L1 (52 words). This task was focused on testing if learners know the primary and

| | Word | A | B | C | D | E | 解答欄 |
|---|------|------------|--------|-----|------|-------|-----|
| 1 | due | ~によって, せいで | しかるべき | 続く | 当然の | すべて正解 | |
| 2 | even | ~でさえ | 偶数の | 互角の | 平らな | すべて正解 | |
| 3 | fair | 公正な | 色白の | 博覧会 | 晴れた | すべて正解 | |
| 4 | fast | ぐっすり | しっかりした | 断食 | はやい | すべて正解 | |
| 5 | fine | 細かい | 罰金 | 晴れた | りっぱな | すべて正解 | |
| 6 | flat | アパート | 退屈な | 平らな | ちょうど | すべて正解 | |

Figure 3. Sample questions of polysemous test 1

| | Word | A | B | C | D | Primary | secondary |
|---|---------|------|------|-----|------|---------|-----------|
| 1 | bank | 川底 | 銀行 | 商店 | 堤防 | | |
| 2 | choice | 上等の品 | 整理 | 選択 | 粗品 | | |
| 3 | copy | 原本 | 筆記用具 | 本一部 | 模写 | | |
| 4 | current | 過去の | 現在の | 沈黙 | 流れ | | |
| 5 | desert | 砂漠 | 拾う | 湖 | 見捨てる | | |

Figure 4. Sample questions of polysemous test 2

secondary meanings of the target words. For example, in Question No.1 “bank” has meanings of “B” and “D,” and does not have the meaning of “A” and “C,” therefore learners were expected to choose “B” and “D.” Furthermore, “bank” is more often used as the meaning of “B” than “D,” therefore learners were supposed to answer “B” as primary and “D” as secondary.

The scores from those two kinds of tests were compared with the vocabulary size test scores and TOEIC scores they had previously taken.

4 Procedure

The TDU VLT, the polysemous word test 1 and 2, were given to students one by one over three English class sessions. Before these tests, they had had a TOEIC institutional test. Four among the 59 participants had not taken this TOEIC test, so the number of participants amounted to 55.

5 Results

Table 4 shows the basic statistics of scores from the VLT and the TOEIC. The mean total score of the VLT turned out to be 95.4, while the mean scores on each frequency-band ranged from 17.2 (4K) to 27.9 (1K). There were significant differences between each neighboring level ($MSe=7.7826$, $*p<.05$, $LSD=1.6808$). Only between 2K and 3K, the mean scores were controverted against frequency (1K > 2K*, 1K > 3K*, 1K > 4K*, 2K < 3K*, 2K > 4K*, 3K > 4K*). We will come back later to discuss this contradiction in the discussion section.

Table 5 shows the basic statistics for the four tests: VLT, PT1, PT2 and TOEIC. As for the PT1, 8

| Var. | Mean | S.D. | Min. | Max. |
|------|------|------|------|------|
| 1K | 27.9 | 2.62 | 20 | 30 |
| 2K | 24.4 | 3.99 | 13 | 30 |
| 3K | 25.9 | 3.43 | 16 | 30 |
| 4K | 17.2 | 5.59 | 5 | 28 |
| VLT | 95.4 | 13.3 | 64 | 117 |

| Var. | Mean | S.D. | Min. | Max. |
|-------|------|------|------|------|
| VLT | 95.4 | 13.3 | 64 | 117 |
| PT1 | 24.6 | 3.66 | 16 | 33 |
| PT2 | 31.6 | 4.61 | 13 | 40 |
| TOEIC | 324 | 63.8 | 210 | 480 |

items (*copy, force, letter, nail, stage, train, start, and taste*) were deleted from the data analysis because fewer than 10 students answered correctly. So the maximum possible score for the PT1 decreased from 50 to 44. The mean score of the PT1 was 24.6 while that of PT2 was 31.6. The total scores at the two tests were different (PT1, 44; PT2, 50), so it is impossible to compare. It might, however, be safely concluded that the PT1 was more difficult than the PT2, considering the maximum possible scores of those tests.

Table 6 shows the correlations among scores of each frequency band of the VLT and score on the TOEIC. The correlation between each frequency band on the VLT and that of the TOEIC ranged from .485 (1K) to .676 (2K). It is interesting to know that the discrepancy between TOEIC scores and scores of 1K and 2K was quite large, compared to other neighboring levels (1K, $r=.485$; 2K, $r=.676$).

Table 7 shows the correlations among four test scores: VLT, PT1, PT2 and TOEIC. The highest correlation was found between scores of the VLT and the TOEIC ($r=.735$). It is striking to find out that the correlation between the scores on the PT2 and the TOEIC was quite low ($r=.294$). Furthermore, the correlation between scores of the PT2 and those of the TOEIC was not significant. This result clearly shows that the PT2 measures a different aspect of vocabulary knowledge, compared to the VLT and the PT1.

Table 6. Correlation table of VLT and TOEIC

| | 1K | 2K | 3K | 4K | TOEIC |
|-------|----|---------|---------|---------|---------|
| 1K | – | 0.591 * | 0.457 * | 0.524 * | 0.485 * |
| 2K | | – | 0.733 * | 0.629 * | 0.676 * |
| 3K | | | – | 0.708 * | 0.632 * |
| 4K | | | | – | 0.652 * |
| TOEIC | | | | | – |

* $p<.05$

Table 7. Correlation table of four tests

| | VLT | PT1 | PT2 | TOEIC |
|-------|-----|---------|----------|----------|
| VLT | – | 0.539 * | 0.294 * | 0.735 * |
| PT1 | | – | 0.208 ns | 0.438 * |
| PT2 | | | – | 0.225 ns |
| TOEIC | | | | – |

* $p<.05$

IV Discussion

One striking result of this study is that a high correlation was found between the VLT and the TOEIC. More specifically, high ($r > .7$) correlations were found between 2K & 3K, 3K & 4K, and moderately high correlations ($.7 > r > .4$) were found between the TOEIC & 2K, 3K & 4K. There might be a ceiling effect on 1K scores. This result may lead us to conclude that the breadth of vocabulary knowledge is the most important for English proficiency.

The results of the VLT showed that a knowledge of the primary meaning correlated with the TOEIC scores higher than that of the peripheral meanings. These results showed that whether learners knew the primary meaning was the most important predictor of English proficiency. This finding may explain why the bilingual vocabulary size test developed in Japan are being used to estimate learners' English proficiency. As far as this result is concerned, a knowledge of polysemous words can be considered somewhere between the breadth and width of vocabulary knowledge.

An interesting result was that scores of 2K and 3K were reversed. This was probably because participants had been given quizzes on 3K frequency words (JACET 8000) as a part of class assignments in their regular class. Intentional vocabulary learning may explain the lower correlation between the score of the 3K section and that of TOEIC score.

One of the polysemous tests did not correlate either with primary knowledge of the words or TOEIC scores. One possibility was that the test itself had flaws. There were no contexts given. Another possibility was that learners might have used test taking strategies of associating core meanings and peripheral meanings by comparing multiple-choice translations in Japanese.

The last but most important point to be discussed is the relationship between vocabulary knowledge and English proficiency. Among the three vocabulary tests used in this survey, the scores of the VLT had a higher correlation score with those of the TOEIC than the other two polysemous tests. As most of the learners in this survey were lower-intermediate students, it is safely assumed that they were below the 5000-word level. Based upon this assumption, the higher correlation between the VLT and English proficiency test scores may support Meara's assertion (1996): breadth knowledge is more important for the learners with less than 5000 words, yet for those who have over 5000 words, depth knowledge is more important.

V Conclusion

It may be concluded that knowledge of the peripheral meanings of words may be somewhat

different from breadth of vocabulary knowledge. It could be somewhere between breadth and depth of vocabulary knowledge. The limitation of this study remains with the polysemous words tests. For the next stage of this study, it is necessary to develop a new test format to measure learners' polysemous knowledge.

Reference

- Aizawa, K. (2006). Rethinking frequency markers for English-Japanese dictionaries. In Murata, M., Minamiide, K., Tono, Y., & Ishikawa, S. (Eds.) *English lexicography in Japan* (pp.108-119). Tokyo, Taishukan.
- Carter, R. & McCarthy, R. (1988). *Vocabulary and language teaching*. London: Longman.
- Coady, J. & Huckin, T. (Eds.). (1997). *Second language vocabulary acquisition*. New York: Cambridge University Press.
- Kosuge, A. (小菅敦子) (1993). 「語彙サイズテストからみた語彙の習得」太田洋・金谷憲・小菅敦子・日臺滋之 (編著) 『英語力はどのようにのびてゆくか』 81-115, 大修館書店
- Meara, P. (1992). EFL Vocabulary Test. Centre for Applied Language Studies, University of Wales Swansea.
- (1996). The dimensions of lexical competence. In Brown, G., Malmkjaer, K., & Williams, J. (eds.) *Performance & Competence in Second Language Acquisition* (35-53). Cambridge: Cambridge University Press.
- Meara, P. & Jones, G. (1988). Vocabulary size as a placement indicator. In Grunwell (ed.), *Applied Linguistics in Society* (pp. 80-87). London: Centre for Information on Language Teaching and Research.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol, UK: Multilingual Matters.
- Mochizuki, M. (望月正道) (2003). 「英語の語彙をどう評価するか」望月正道・相澤一美・投野由紀夫 (編著) 『英語語彙の指導マニュアル』 大修館書店 . 181-209.
- Nation, I.S.P. (2001). *Vocabulary in another language*. Cambridge: Cambridge University Press.
- Paribakht, T. & Wesche, M. (1993). The relationship between reading comprehension and second language development in a comprehension- based ESL program. *TESL Canada Journal* 11, 9-29.
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing* 10, 355-371.
- (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Ringbom, H. (1987). *The role of first language in the second language learning*. Clevedon/ Philadelphia: Multilingual Matters.
- Schmitt, N., Schmitt, D. & Clapham, (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test, *Language Testing* 18, 55-88.
- Verspoor, M. & Lowie, W. (2003). Making Sense of Polysemous Words. *Language Learning* 53, 547-586
- Wesche, M. & Paribakht, T. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. *The Canadian Modern Language Review* 53, 13-40.
- Wilkins, D. (1972). *Linguistics in language teaching*. London: Arnold.