Depth of lexical knowledge in L2 Acquisition: Problems of predictability of L2 proficiency in vocabulary tests

Norifumi UEDA

Abstract

Researchers in the field of language teaching think that it is good for L2 learners to know more words. Moreover, it is widely accepted that as L2 leaners are getting higher L2 proficiency in L2, they are expected to acquire larger vocabulary. Many vocabulary tests have been developed for evaluating learner’s vocabulary knowledge. However, these tests can evaluate only small parts of vocabulary knowledge in L2 leaners such as vocabulary size and word association. In this paper, we investigate what vocabulary knowledge is in L2 learners, and what kinds of items vocabulary tests should include. We also report the results of our newly developed vocabulary tests, and suggest new types of vocabulary tests.

0 Introduction

Vocabulary is very important for L2 language learning. Among researchers, it is widely accepted that, as L2 leaners are getting higher proficiency in L2, they are expected to acquire larger vocabulary. (Palmer 1972). Many types of vocabulary tests for second language learners have been developed: for example, Nation’s Vocabulary Levels Test (Nation, 1990), Lex 30 (Meara & Fitzpatrick, 2000), and V-Links (Meara & Wolter, 2004). Though these aim at evaluating L2 learner’s vocabulary knowledge, these tests evaluate different aspects of lexicon. Moreover, there is a problem that the validity of these vocabulary tests has not been discussed and evaluated.
In the following section, we discuss (1) definition of vocabulary knowledge (what vocabulary knowledge is in L2 learners), and (2) validity of vocabulary tests (what kinds of vocabulary tests are appropriate to evaluate vocabulary knowledge in L2 learners). In Section 2, we investigate two main types in L2 vocabulary tests, and clarify shortcomings in them. In Section 3, we report newly developed vocabulary tests to improve the shortcomings outlined in Section 2. The validity of these vocabulary tests is examined and the results suggest the possibility of a new type of vocabulary test to evaluate L2 learner’s lexical knowledge.

1 Definitions of vocabulary knowledge

Many researchers discuss the characteristics of vocabulary knowledge. Nation (2001) describes vocabulary knowledge by two types from the viewpoint of competence of using vocabulary; receptive, or active, and productive, or passive, vocabulary. Nation defines receptive vocabulary as lexical knowledge for ‘perceiving the form of a word while listening and retrieving its meaning’; and productive vocabulary, lexical knowledge for ‘expressing a meaning through speaking or writing and retrieving and producing the appropriate spoken or written word form’ (2001, pp. 24-25). Anderson & Freebody (1981) and Read (1993) explain vocabulary knowledge by distinguishing it with two notions: breadth of word knowledge and depth of word knowledge. Breadth of word knowledge refers to lexical knowledge ‘by which we mean the number of words for which the person knows at least some of the significant aspects of meaning’ (Anderson & Freebody, 1981, p.93) (or to say more simply, the size of a learner’s vocabulary), whereas depth of word knowledge means ‘quality of understanding’ words (1981, p.93). Henriksen (1999) depicts vocabulary knowledge with three dimensions: partial-precise knowledge dimension, a depth of knowledge dimension, and a receptive-productive dimension. Meara (1996) explains vocabulary knowledge from the practical viewpoint; ‘size’ and ‘organization’. ‘Organization’ means associations between words.
Depth of lexical knowledge in L2 Acquisition: 
Problems of predictability of L2 proficiency in vocabulary tests

The dichotomic definitions on vocabulary knowledge mentioned above are incomplete to depict vocabulary knowledge (Milton, 2009). Some researchers have been trying to describe vocabulary knowledge in more detailed descriptive ways. Richards (1976) describes vocabulary knowledge as follows:

1. The native speaker of a language continues to expand his vocabulary in adulthood, whereas there is comparatively little development of syntax in adult life.
2. Knowing a word means knowing the degree of probability of encountering that word in speech or print. For many words we also know the sort of words most likely to be found associated with the word.
3. Knowing a word implies knowing the limitations on the use of the word according to variations of function and situation.
4. Knowing a word means knowing the syntactic behaviour associated with the word.
5. Knowing a word entails knowledge of the underlying form of a word and the derivations that can be made from it.
6. Knowing a word entails knowledge of the network of associations between that word and other words in the language.
7. Knowing a word means knowing a semantic value of a word.
8. Knowing a word means knowing the different meanings associated with a word.

(1976, p.83)

He claims that learners should learn syntactic and semantic usage, derivational forms, associations, and word frequencies of the words.

Nation also outlines vocabulary knowledge on the basis of the distinctive categories, receptive and productive. He sets three areas (form, written, and meaning), under which he puts each three subcategories, then explains vocabulary knowledge according to these subcategories and a receptive/productive distinction. (Table 1).
The descriptions proposed by Richards and Nation cover broad range of knowledge concerning vocabulary, and commonly include knowledge on syntactic and semantic usage, derivational forms, associations and word frequencies of the words, though there are some different aspects in lexicon dealt with.

There are other ideas about vocabulary knowledge in psycholinguistics. There are many psycholinguistic studies on vocabulary relations between concepts, and words of L1 and L2 in L2 acquisition process in bilingual studies. It is hypothesized that in the process of L2, L2 learners should integrate the conceptual differences between two languages. And also, it is proposed that when bilinguals acquire L2 words, which have

---

**Table 1**

*What is involved in knowing a word*

<table>
<thead>
<tr>
<th>Form</th>
<th>spoken</th>
<th>R</th>
<th>What does the word sound like?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>P</td>
<td>How is the word pronounced?</td>
</tr>
<tr>
<td>written</td>
<td>R</td>
<td>What does the word look like?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>How is the word written and spelled?</td>
<td></td>
</tr>
<tr>
<td>word parts</td>
<td>R</td>
<td>What parts are recognizable in this word?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>What word parts are needed to express the meaning?</td>
<td></td>
</tr>
<tr>
<td>meaning</td>
<td>R</td>
<td>What meaning does this word form signal?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>What word form can be used to express this meaning?</td>
<td></td>
</tr>
<tr>
<td>concept and referents</td>
<td>R</td>
<td>What is included in the concept?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>What items can the concept refer to?</td>
<td></td>
</tr>
<tr>
<td>associations</td>
<td>R</td>
<td>What other words does this make us think of?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>What other words could we used instead of this one?</td>
<td></td>
</tr>
<tr>
<td>Use</td>
<td>grammatical function</td>
<td>R</td>
<td>In what patterns does the word occur?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P</td>
<td>In what patterns must we use this word?</td>
</tr>
<tr>
<td>collocations</td>
<td>R</td>
<td>What words or types of words occur with this one?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P</td>
<td>What words or types of words must we use with one?</td>
<td></td>
</tr>
<tr>
<td>constraints on use</td>
<td>R</td>
<td>Where, when, and how often would we expect to meet this word?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(register, frequency ...)</td>
<td>P</td>
<td>Where, when, and how often can we use this word?</td>
</tr>
</tbody>
</table>

Source: Nation (2001, p.27)

Note. R = receptive, P = productive
concepts roughly equivalent to L1 but have different conceptual boundaries, or where one conceptual domain of L1 words is divided into two or more in L2, bilinguals have to restructure the L1 concepts to L2, or adjust the concepts of L2 to L1. (Ameel, Malt, Storms & Assche, 2009; Dong, Gui, & MacWhinney, 2005; Ijaz, 1986; Ueda, 2007).

2 Vocabulary Tests

Many vocabulary tests have been developed for evaluating learner’s vocabulary knowledge. Different researchers developed different vocabulary tests to measure vocabulary knowledge according to their interests in a particular dimension of vocabulary knowledge. (Laufer & Goldstein, 2004). Generally, these tests can be categorized into two types: One type is for testing the width of vocabulary knowledge like Nation’s Vocabulary Levels Test (Nation, 1990) and the other, for examining the depth of vocabulary knowledge such as Lex 30 (Meara & Fitzpatrick, 2000). We will review how these two types of vocabulary tests have been made to evaluate breadth and depth of vocabulary knowledge in the next section.

2.1 Vocabulary tests for Breadth of vocabulary knowledge

Tests for breadth of vocabulary knowledge intend to evaluate how much words the learners retain in their mental lexicon. The breadth of vocabulary knowledge is estimated by word frequencies.

In Vocabulary Levels test (Nation, 1990), for example, items are randomly selected from each word frequency level: 2000-word level, 3000-word level, 5000-word level, the university word level, and 10000-word level.

Each section of Vocabulary Levels test consists of six words and three word definitions. The definitions in one section come from the words which are included in the higher level of the word frequencies: for example, the words from the 2000-word level use words in the first 1000 words for the definitions. In this test, the vocabulary levels of testees are estimated by the scores.
2.2 Vocabulary tests for depth of vocabulary knowledge

Tests for depth of vocabulary knowledge intend to evaluate paradigmatic (synonyms), syntagmatic (collocations) and analytic knowledge (associations which represent one aspect or components of the meaning of the stimulus word and is likely to form part of its dictionary) of vocabulary. (Read, 1993). In Lex 30, for example, all test items are selected based on the following criteria:

1. All the stimulus words are highly frequent.
2. None of the stimulus words typically elicits a single, dominant primary response.
3. Each of the stimulus words typically generates responses which are not common words.

(Meara & Fitzpatrick, 2000)

In this test, testees were asked to write words using free word association. The collected data were analyzed according to the word frequencies. Scores were given by word frequency level. (For example, Words in Level 0 and 1 were given zero points, and those in other levels than Level 0 and 1, one point.)

The main issue of vocabulary tests to examine the depth of vocabulary knowledge like Lex 30 put a focus mainly on evaluating the learner’s knowledge on word association. (The word frequency is used as one of the criteria to evaluate the learner’s vocabulary knowledge, so it can be said that Lex 30 examines the learner’s knowledge on the word frequency1.) It is clear to utilize mainly word frequency (in order to evaluate the depth of vocabulary knowledge) though vocabulary knowledge consists of many other types of knowledge such as knowledge on syntactic and semantic usage as discussed in the previous section.

---

1 Effects of word frequency on lexical decision are reported in many researches in psycholinguistics field: for example, Foster (1981) and Gordon (1983). In this sense, it is inevitable to include the word frequency information as a criterion in vocabulary tests.
3 New developed vocabulary tests to examine the depth of vocabulary knowledge.

As we saw in the previous sections, the tests to evaluate width of knowledge use word frequencies to estimate learners’ vocabulary size, whereas the tests for evaluation of depth of knowledge utilize word associations task. The existing tests for the depth of vocabulary knowledge deal with a few one of various aspects of vocabulary knowledge. It is better to evaluate more aspects in lexical knowledge in a vocabulary test. We have been developing new type of vocabulary tests to evaluate as many aspects of lexical knowledge as possible. We developed three sets of vocabulary tests supported by Grant-in-Aid for Scientific Research\(^2\). The aim of the vocabulary test is to examine the depth of knowledge in the second language learners of English. In the tests, the subjects were required to distinguish synonyms by using syntactic and semantic knowledge (Ueda, Tsutsui, Kondo, & Nakano, 2009; Ueda, Tsutsui, Kondo, Oya, and Nakano, 2010). Test items in the tests can also trace L2 learner’s developmental process in conceptual representations in the mental lexicon. The detailed information of the two tests is in Table 2 and 3. (See also Appendices.)

Table 2

Examples of tests items in Test 1

<table>
<thead>
<tr>
<th>[hear/ listen]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choose the proper word/words.</td>
</tr>
<tr>
<td>1. Did you (   ) that noise? [hear/ listen]</td>
</tr>
<tr>
<td>2. I (   ) to music. [hear/ listen]</td>
</tr>
<tr>
<td>3. I can (   ) somebody knocking the door. [hear/ listen]</td>
</tr>
<tr>
<td>4. I told him not to go, but he wouldn't (   ). [hear/ listen]</td>
</tr>
<tr>
<td>5. She doesn't (   ) very well. [hear/ listen]</td>
</tr>
<tr>
<td>6. You can (   ) that news again. [hear/ listen]</td>
</tr>
<tr>
<td>7. (   ), I'm sure you can work this out. [hear/ listen]</td>
</tr>
<tr>
<td>8. Do you (   ) me? [hear/ listen]</td>
</tr>
</tbody>
</table>

---

2 This project was supported by Grant-in-Aid for Scientific Research (C) (20520516) (2008-2010)
In Test 2, as the test items, we chose basic verbs: hear and listen (Group 1); see, look, watch, gaze, and stare (Group 2); say, talk, speak, tell and utter (Group 3). The verbs in each group have similar meanings but occur in different syntactic or semantic circumstances. The subjects were requested to know syntactic and semantic knowledge to use the synonyms. This vocabulary test aimed to estimate how much knowledge the learners have to use synonyms properly. We made 8 items for Group 1, 20 items for Group 2 and 23 items for Group 3. In each test item, there is a blank to be filled in by the most appropriate verb(s). The total number of the test items was 53. (See Appendix 1 and 2.)

In Test 2, test items consist of the structure; basic verb + noun. The verbs in each item have similar meanings but occur in different syntactic or semantic circumstances. The subjects were required to know the rules of selectional restriction. The verbs and nouns are high frequent words and familiar to the L2 learners: for example, verbs are make, take, do, get, become and so on. 22 items were developed. (See Table 3 for example items.)

We also developed one more test, focusing on conceptual changes. The test items include adjectives. In this paper, the former two tests are only used, so we do not explain the detailed about this latter test.

The CGI on the internet was made for the two tests, and subjects accessed the
internet to answer all the items. Subjects were asked to answer the all the questions. There is no time limitation for answering the items. All the items were analyzed to calculate infit and outfit by Winsteps 3.68.1., an IRT software, to select good items. All the test items, in fact, consisted of multiple choices (or the test items developed here were not simple one answer to one test item). This type of test items could not calculated by Winsteps. Hence, we counted one choice as one test item: For example, in a question in Test 1 (Did you (         ) that noise? [hear/ listen]), we regarded this as two test items. From the results of calculating infit and outfit in two tests, we checked adequacy as the test items according to McNamara (1996). We found that all the items were proper as the test items. (See Appendix 1 and 2 for all the items in the two tests.)

4 Experiment

4.1 Outline

By using the refined two tests, we conducted an experiment. The aim of the experiment is to investigate how effectively our developed vocabulary tests can predict the learner’s L2 proficiency level. For this purpose, we gathered data from university students by using the originally developed vocabulary test and also the students were asked to take other English test to examine their English proficiency level. For this, we used Assessment of Communicative English (ACE) test, which is developed on the basis of item response theory. This test consists of three parts: the listening section, the grammar and vocabulary section, and the reading section.

Participants were university students majoring in various academic fields: robotics, information communication technology, architecture technology, socio-environmental design, Pharmaceutical Sciences, nursing, psychology, social welfare, social information and media. Different groups participated in the two tests. 100 students participated in Test 1 (Group 1), and 134 students, in Test 2 (Group 2). In Group 1, 22 of the participants in Group 1 and 23 of the in Group 2 did not take ACE test, so available data is from 88 students in Group 1 and from 111 students in Group 2. As
for the difference in quality between two groups, there was no statistical difference in scores between two groups ($t = .87$, $df = 197$). (See Table 4). Hence, we can say that these two samples come from the same population: that is to say, it is possible to hypothesize that these two samples have the same distribution.

### 4.2 Results of the Experiment

Correlation coefficient between the results of ACE test and those of the developed vocabulary tests were calculated within the groups. Statistically significant correlation coefficients were found between the scores of ACE and those of the vocabulary tests ($r = .459$, $p < .01$ between ACE Test and Test 1; $r = .485$, $p < .01$ between ACE test and Test 2).

### 5 General Discussion and Conclusion

Regarding vocabulary knowledge, many researchers implicitly hypothesize that the more words the learner knows, the higher the possibility becomes that these words can belong to the high frequency levels. Considering the words in the developed two tests, almost all of them are high frequency ones. According to JACET 8000, almost all of the words belong to Level 1 or Level 2 (Level 1 represents 1000 word frequency level, and Level 2, 2000 word frequency level). Only two words, ‘utter’ and ‘browse’ are low frequency words (‘utter’ is in 6000 word level and ‘browse’ is not included in JACET 8000).
Depth of lexical knowledge in L2 Acquisition: Problems of predictability of L2 proficiency in vocabulary tests

In the experiment, we examined how effectively our developed vocabulary tests can predict the learner’s L2 proficiency level. The results showed high correlation coefficient between the participant’s proficiency level and the scores in vocabulary tests. The results can support the presupposition that the more words the learner knows, the higher the possibility becomes that these words can belong to the high frequency levels. It also implies more advance learners tend to have bigger vocabularies than less advanced ones concerning the depth as well as breadth of vocabulary knowledge.

The vocabulary size has been considered as one criterion to indicate the learner’s proficiency level since Palmer (1917). As Milton (2007) pointed out about frequency-based model of lexical learning, information on word frequency cannot necessarily be relevant to foreign language learner’s vocabulary knowledge, and frequency is not only factor to explain learner’s vocabulary knowledge. It follows that most vocabulary tests like Vocabulary Levels Test, in this sense, cannot evaluate the learner’s vocabulary knowledge in a comprehensive manner. In this paper, we investigated the possibility to develop a new type of vocabulary tests. Our developed vocabulary tests aims at evaluation of depth of vocabulary knowledge, and cover wide range of vocabulary knowledge not like other tests for vocabulary tests. Hence, our test style can guarantee validity of the tests. Vocabulary knowledge consists of various layers, so it is difficult to comprehensibly depict it by one vocabulary test to evaluate one aspect of lexical knowledge. From now on, many vocabulary tests have been developed to evaluate depth of lexical knowledge. To improve validity of vocabulary tests, the researchers should add more aspects into the evaluation criteria (word association and word frequency) in vocabulary tests.

References


Norifumi UEDA


Depth of lexical knowledge in L2 Acquisition:
Problems of predictability of L2 proficiency in vocabulary tests


*Appendix 1*

Test items in Test 1 after the selection by Winsteps.

<table>
<thead>
<tr>
<th>hear/ listen</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Did you ( ) that noise?</td>
</tr>
<tr>
<td>2. I ( ) to music</td>
</tr>
<tr>
<td>3. I can ( ) somebody knocking the door.</td>
</tr>
<tr>
<td>4. I told him not to go, but he wouldn't ( ).</td>
</tr>
<tr>
<td>5. She doesn't ( ) very well.</td>
</tr>
<tr>
<td>6. You can ( ) that news again.</td>
</tr>
</tbody>
</table>
7. (         ), I'm sure you can work this out.
8. Do you (         ) me?

[Tell, say, speak, talk, utter.]
9. You should (         ) me your name and address.
10. I wouldn't (         ) so
11. Who was that you were (         ) to at the party?
12. He refused to (         ) his name.
13. Do you (         ) English
14. He never (         ) a word.
15. We need to (         ) before meeting.
16. She (         ) a sigh.
17. The red light (         ) you when the machine is ready to use.
18. I would like to (         ) about our culture.
19. The stress of work (         ) on their marriage.
20. The clock (         ) it is five now.
21. Most babies start to (         ) by 18 months.

[see, watch, look at, view, stare, gaze, browse]
22. I can(         ) that you're not very happy with the situation.
23. Jack (         ) them slowly crime up the wall.
24. Sorry, I wasn't (         ).
25. Tourists came to (         ) the gardens every year.
26. She doesn't like being (         ) at.
27. She (         ) at me in disbelief when I told her the news.
28. She (         ) through the magazine.
29. I can (         ) why.
30. I (         ) the baseball game.
31. (         ) at me.
32. The building is more beautiful when you (         ) it from the river.
33. It is rude to (         ).
34. He (         ) in a library.
35. Can I (         ) your ticket?
36. Can you (         ) my bag?
37. She was (         ) vacantly into the space.
Appendix 2

Test items in Test 2 after the selection by Winsteps

1. Becky will ( ) a good doctor.
   [be, become, turn, make]

2. Things are ( ) worse.
   [becoming, getting]

3. The leaves are ( ) red in fall.
   [becoming, being, getting, turning]
   [becoming, being, getting, turning]

4. The signal ( ) red.
   [became, turned, made]

5. Kim ( ) a picture of the castle.
   [got, took]

6. John ( ) a vacation.
   [had, got, took]

7. Please ( ) the box for me.
   [get, take, reach]

8. Please ( ) the box to me.
   [get, take, reach]

9. Please ( ) me the box.
   [get, take, reach]

10. Mary will ( ) the degree.
    [get, take]

11. Mary will ( ) a math course.
    [get, take]

12. Let’s ( ) lunch.
    [have, get, take]

13. Let’s ( ) a break.
    [have, get, take]

14. Naomi will ( ) a high score.
    [get, take]

15. Sarah ( ) a decision.
    [did, made]

16. Tom ( ) sports regularly.
    [does, makes]

17. Mike will ( ) some exercise tomorrow.
    [do, make]
18. Donald will ( ) an effort to spend more time with his family.
   [do, make]
19. Victoria will ( ) a speech at the party.
   [do, make]
20. Blair ( ) some reading.
    [did, made]
21. Takashi ( ) some research about the college.
    [did, made]