

## **A Study on Prediction of L2 Learner's Vocabulary knowledge by Application of Latent Rank Theory to a Vocabulary Test**

Norifumi Ueda<sup>1</sup>, Kazuharu Owada<sup>2</sup> and Eiichiro Tsutsui<sup>3</sup>

<sup>1</sup>Faculty of Arts and Sciences, Komazawa University; <sup>2</sup>Tokyo College of Music and <sup>3</sup>Faculty of Pharmaceutical Sciences, Hiroshima International University

<sup>1</sup>ueda@komazawa-u.ac.jp, <sup>2</sup>qwq03702@nifty.com and

<sup>3</sup>tsutsui@ic.hirokoku-u.ac.jp

### **Abstract**

This study aims to examine (1) what factors of lexical knowledge can affect L2 learner's vocabulary, (2) whether difficulty in lexical acquisition by L2 learners can change according to types of lexical knowledge and (3) whether types of lexical difficulties can be predicted by L2 learners' English proficiency level. The vocabulary test created by Ueda, Tsutsui, Kondo, Oya & Nakano (2010, 2011), and Ueda, Tsutsui, Kondo, Owada & Nakano (2012) were used in the experiment. 658 university students took the vocabulary test. The data were analyzed by Latent Rank Theory (LRT) (Shojima, 2008).

The findings showed (1) lexical information such as synonyms, antonyms, derivational forms, usages, and collocations can affect the L2 learners at all proficiency levels; (2) the L2 learners showed difficulties in answering the items which include more aspects of vocabulary knowledge, even though they are at the high frequency level, and (2) the results of LRT can predict the aspects of learners' difficulties in acquiring vocabulary.

The results of the vocabulary test can suggest developmental process in L2 vocabulary acquisition by L2 learners.

### **Keywords**

Vocabulary Test, L2 vocabulary knowledge, LRT

### **1 Introduction**

In the field of L2 teaching, it is widely accepted that vocabulary is very important in learning L2. Anderson and Freebody (1981), for example, claims "Measures of vocabulary knowledge are potent predictors of a variety of indices of linguistic ability" (p.77). Hence, it is beneficial to develop a vocabulary test which can properly predict L2 learners' vocabulary knowledge.

### **2 Vocabulary knowledge**

Vocabulary knowledge has been examined by many researchers. Nation (2001) describes vocabulary knowledge by two types from the viewpoint of competence of using vocabulary; receptive and productive 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). Moreover, Nation explains lexical knowledge by three areas ('form', 'meaning' and 'use') and some subcategories to each area: 'spoken', 'written' and 'word parts' under the area of 'form'; 'form and meaning', 'concept and referents', and 'associations' under the area of 'meaning'; and 'grammatical function', 'collocations' and 'constrains on use' under the area of 'use'(2001, p27). (Table 1)

Table 1: What is involved in knowing a word (from Nation, 2001, p.27)

Form	spoken	R	What does the word sound like?
		P	How is the word pronounced?
	written	R	What does the word look like?
		P	How is the word written and spelled?
	word Parts	R	What parts are recognisable in this word?
		P	What word parts are needed to express meaning?
Meaning	form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	concepts and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	association	R	What other words does this word make us think of?
		P	What other words could we used instead of this one?
Use	grammatical function	R	In what patters does the word occur?
		P	In what patterns must we use this word?
	collocations	R	What words or types of word occur with this one?
		P	What words or types of words must we use with this one?
	constraints on use	R	Where, when and how often would we meet this word?
		P	Where, when and how often can we use this word?

Note: P represents productive and R, receptive.

Anderson & Freebody (1981), and Reed (1993) explain vocabulary knowledge by classifying 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) and whereas depth of word knowledge means ‘quality of understanding’ words (1981, p.93). Henriksen (1999) depicted vocabulary knowledge with three dimensions: partical-precice 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. Richards (1976) describes lexical competence by eight characteristics:

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.

(p. 83)

The descriptions proposed by Richards (1976) and Nation (2001) cover a broad range of knowledge concerning vocabulary, and commonly include knowledge on syntactic and semantic usage, derivational forms, associations, and frequencies of the words, though some different aspects in lexicon are dealt with.

(Especially Nation uses productive/receptive aspect to define vocabulary knowledge, but Richards not.)

There are many facets of ‘vocabulary knowledge’, but it is common that ‘vocabulary knowledge’ consists of various factors.

### **3 Vocabulary tests**

Many vocabulary tests have been developed, and each vocabulary test aims to evaluate different aspects of vocabulary knowledge. In this section, we review the characteristics of three different kinds of vocabulary tests: Vocabulary Levels Test, Lex 30 and Multidimensional Vocabulary test. Then, we outline the vocabulary test used in the experiment, focusing on the differences from the three tests.

#### **3.1 Vocabulary Levels Test**

Vocabulary levels test (VLT) developed by Nation (1990) estimates L2 learner’s breadth of vocabulary knowledge, or L2 learner’s vocabulary size. In VLT 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 VLT consists of six words and three word definitions. Test-takers are asked to match words and definitions, and the vocabulary level of the test-takers is estimated by the scores.

#### **3.2 Lex 30**

Lex 30, developed by Meara & Fitzpatrick (2000), is a vocabulary test which estimates L2 learner’s depth of vocabulary knowledge. In Lex 30, 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 Lex 30, test-takers are 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.

#### **3.3 Multidimensional Vocabulary test by Mochizuki, Uemura, Aizawa, Sugimori, Ishikawa, Iso & Koizumi (2010)**

Mochizuki et al. (2010) developed a vocabulary test which aims to examine L2 test-taker’s vocabulary knowledge multidimensionally. The test consists of three sub-vocabulary tests: J8 Vocabulary Size Test (J8VST), Lexical Organization Test (LOT) and Lexical Accessibility Test (LEXATT). J8VST estimates the L2 learner’s breadth of vocabulary knowledge. Test items are from the most frequently used 5000 words in the JACET 8000 Word List<sup>1</sup>. In J8VST, test-takers are asked to choose the English word among four choices, which are Japanese stimulus word. LOT measures the L2 learner’s knowledge on vocabulary organization, especially on word associations. In LOT, test-takers are asked to choose two words among three that form the strongest connection among the choices. LEXATT aims to measure word recognition speed. LEXATT is divided into two different tasks. In Part 1, test-takers have to find a common word hidden in a string of letters. In Part 2, test-takers have to choose the meaning of the recognized word as fast as possible.

#### **3.4 The Vocabulary Test developed by Ueda et al.**

The vocabulary test developed by Ueda, Tsutsui, Kodo, Oya & Nakano (2010, 2011) and Ueda, Owada, Kondo, Tsutsui & Nakano (2012) is a test to aim to evaluate L2 learners’ vocabulary knowledge multidimensionally. The main target of this test is Japanese test-takers’ depth of vocabulary knowledge. In this test, the word frequency (or the breadth of vocabulary knowledge) is also taken into consideration. The depth of vocabulary knowledge to be examined in this test is as follows:

---

<sup>1</sup> JACET 8000 Word List (officially, JACET List of 8000 Basic Words) consists of eight levels based on word frequency levels. Level 1, for example, include words from 1 to 1000 word frequency level, and Level 8, from 7001 to 8000 word frequency level.

## Selected Papers of the 18th Conference of Pan-Pan Pacific Association of Applied Linguistics

1. Collocations
2. Synonyms
3. Antonyms
4. Word associations
5. Derivational forms
6. Usages (selectional restriction, etc.)
7. Idioms
8. Polysemys
9. Conceptual differences between L1 (Japanese) and L2 (English)

In the points of examined features of vocabulary knowledge, this test is similar to Multidimensional Vocabulary test by Mochizuki et al (2010), except that the test of Ueda et al examines usage, Idiom and conceptual differences between L1 (Japanese) and L2 (English) but not word recognition speed.

In the vocabulary test, the word frequency levels of each stimulus word, answers and distractor words were carefully controlled according to the word frequency levels in JACET 8000 Word List. (This means that stimulus and distractor words consist of words from Level 1 to Level 8 in JACET 8000.) In the process of creating the vocabulary test, word familiarity as well as word frequency to L2 learners was considered. The distractor words were chosen according to the criteria below:

- Distractor words were at higher frequency level than, or at least as frequent as the stimulus word<sup>2</sup>.
- Distractor words were chosen according to the semantic link (association) to the stimulus words by use of the Edinburgh Associative Thesaurus<sup>3</sup>.
- The familiarity levels of the distractor words was taken into account with reference to English Word Familiarity List for Japanese EFL Learners (Yokokawa, 2006).

The vocabulary test consisted of two types of multiple choice questions: 32 test items with single-answer questions in Part 1, and 21 test items with multiple-answer questions in Part 2. The test in Part 1 aimed to examine L2 learners' knowledge on synonyms, antonyms, collocations and derivational forms of the target words. Item 1 to 11 targeted to examine test-takers' lexical knowledge on synonyms; Item 12 to 21, on antonyms; Item 22 to 27, on collocations; Item 28 to 30, on polysemys, and synonyms; and Item 30 and 31, on polysemys and antonyms. Some of the items for antonyms required test-takers' lexical knowledge on derivational forms (Item 12, 14, 15, 20, and 21). (See Table 2 for targeted features in each test item and Appendix A for test items in Part 1.)

Table 2: Targeted features of each test item in Part 1

Item No	Targeted features
Item 1	Synonyms
Item 2	
Item 3	
Item 4	
Item 5	
Item 6	

<sup>2</sup> In the process of the vocabulary test, this criterion cannot be met when (1) the word frequency of the desirable target words are higher than that of stimulus words as in Item 17, and 22; (2) there is no information about the word frequency of the target words in JACET 8000 as in Item 23; (3) the target words have high familiarity rates even though their word frequency is lower than that of the target words like Item 5; and (4) the stimulus words are polysemys from the high frequency levels in JACET 8000 as in Item 28 and 31.

<sup>3</sup> Read (1993) selected the choices for the test to examine the L2 learner's depth of the vocabulary knowledge in the similar way. Read adopted three types of relationship between stimulus word and choices to make distractors: Paradigmatic type, syntagmatic type and analytic type. The words of the paradigmatic type have similarity between the stimulus and choices. The words of the syntagmatic type have high possibilities to occur together in a sentence. The words of the analytic type have some aspect of the meaning of the stimulus word and appear as a part of the definition in the dictionary. However, Read did not choose distractors with semantic links to the stimulus word. In this regard, the vocabulary test by Ueda et al (2010, 2011) and Ueda et al (2012) is different from the vocabulary test by Read (1993).



	(Vocabulary size) word association, derivational forms, collocations, constraints on use, grammatical function, conceptual differences between L1 and L2	R	Yes
Test of Ueda et al.			

*Note:* R/P represents whether the test evaluate productive (p) or receptive (R) vocabulary knowledge; and Word Freq., whether the test take the information of word frequencies into consideration.

The targeted features of vocabulary knowledge vary among the tests. The test developed by Ueda et al covers various features of vocabulary knowledge. In this study, one of the research questions was what factors of the lexical knowledge can affect L2 learner's vocabulary. Hence, we adopted the test developed by Ueda et al for the experiment.

## 4 Experiment

The purpose of the experiment was to examine (1) what factors of the lexical knowledge can affect L2 learner's vocabulary, (2) whether lexical difficulties for the L2 learner can change according to types of lexical knowledge, and (3) whether types of lexical difficulties can be predicted by L2 learner's English proficiency level.

### 4.1 Participants

658 university students from seven different universities in Japan participated in the experiment. They had various academic backgrounds: pharmacy, sports science, robotics, architecture, business administration, economics, Japanese literature, English literature, sociology, and Engineering. Their English proficiency levels also vary.<sup>5</sup>

### 4.2 Method

The subjects were asked to answer the all the questions in the vocabulary test (Ueda et al., 2010, 2011; Ueda et al., 2012) by either a web-based or paper-based test. There was no time limitation for answering the items. The results were analyzed by Exametrika ver. 5.3, a software program for Latent Rank Theory (Shojima, 2011). The results from each part in the vocabulary test were analyzed separately. The items in Part 2 were multiple-answer questions. Hence, we counted one choice as one test item to analyze the test scores like Yes/No Question type test<sup>6</sup>. For example, Item 1 in Part 2 has four choices. So, Item 1 was treated as four different questions in calculation like Item 1-be, Item 1-become, Item 1-turn and Item 1-make.

### 4.3 Latent Rank Theory

Latent Rank Theory (LRT) is a new testing theory developed where ordinary scale is used to classify examinees into a certain level according to the test results. Shojima (2008) claims that it is difficult to explain the relationship between scores and abilities because the test scores do not have sufficient resolution. As the output of LRT, the rank membership profile (RMP) can be obtained, which is useful for evaluating the possibilities of each examinee belonging to respective ranks.

LRT is useful in educational settings. Some possible applications are proposed: It can construct an ability profile for each achievement level (latent rank) and an achievement progress table (like Can-do statement in CEFR) (Shojima, 2008). LRT is also applicable to the placement tests. Koizumi and Iimura (2010) reported that LRT can produce the same results as Classical Test theory and Rasch modeling can. Kimura (2009)

<sup>5</sup> All the participants did not take the same English proficiency test such as TOEIC: Some of them took TOEIC; some, Assessment Communicative English (ACE) test and some, Standard Test of English Proficiency (STEP). According to the reported TOEIC scores, the range of scores was very wide: from 210 to 825.

<sup>6</sup> The items in Part 2 aimed to check whether the test-taker has acquired lexical knowledge on word associations, collocations and conceptual differences between L1 and L2. Yes/No type questions are more appropriate than questions where single-answer should be chosen of the four choices so that the test-taker's lexical knowledge can be predicted in an exact way.

suggested that it is easy to handle the results by LRT in setting the cutting points of the placement tests.

#### 4.4 Results

##### 4.4.1 Results of Part 1

Table 4 shows that more than 50% of participants could answer the twelve items (Item 1, 3, 4, 5, 6, 9, 13, 14, 17, 18, 19 and 23) correctly. The word frequency levels in these items are mostly from Level 1 to Level 4 in JACET 8000. Moreover, except for Item 14 and 23, these items were simple questions: in other words, it was examined whether test-takers know an antonym or a synonym word to the stimulus. The targeted lexical knowledge in Item 14 was on derivational forms ('direct,' 'indirect') and that of Item 23, on polysemys and collocations (the peripheral meaning of 'green': "not matured").

On the other hand, concerning items with less than 30 % answered correctly (Item 2, 24, 26, and 27), almost all the word frequency levels were Level 4 and Level 5 in JACET 8000. Besides the difficulty that may be caused from the low word frequency of the words in question, this may indicate that some items (Item 24, 26 and 27) were difficult for the test-takers because such items require their lexical knowledge on collocations: 'commit suicide' in Item 24, 'economic sanction' in Item 26 and 'a guilty conscience' in Item 27.

Table 4: Results of the Vocabulary Test in Part 1

	% of correct answer	S. D.
Item 13	0.818	0.386
Item 18	0.801	0.400
Item 8	0.784	0.412
Item 23	0.749	0.434
Item 4	0.655	0.476
Item 2	0.629	0.483
Item 17	0.614	0.487
Item 19	0.599	0.491
Item 3	0.597	0.491
Item 5	0.571	0.495
Item 14	0.508	0.500
Item 31	0.486	0.500
Item 28	0.467	0.499
Item 32	0.460	0.499
Item 30	0.441	0.497
Item 10	0.430	0.495
Item 20	0.400	0.490
Item 16	0.380	0.486
Item 6	0.354	0.479
Item 25	0.348	0.477
Item 22	0.340	0.474
Item 21	0.336	0.473
Item 7	0.330	0.470
Item 12	0.324	0.468
Item 9	0.309	0.462
Item 15	0.301	0.459
Item 29	0.296	0.457
Item 27	0.266	0.442
Item 24	0.264	0.441
Item 1	0.240	0.427
Item 11	0.214	0.411

Item 26	0.169	0.375
---------	-------	-------

## 4.4.2 Results of Part 2

Table 5 shows the percentage of the correct answers for each item<sup>7</sup>. The results in Part 2 show that even though all the items contain basic verbs, of which the word frequency levels are Level 1 and 2, all the participants could not answer the questions perfectly. This suggests that the participants failed to answer the questions with a focus on L2 learner's knowledge on polysemy, usages and distinction of conceptual differences between L1 and L2.

Table 5: Results of the Vocabulary Test in Part 2

Item No.	choices	% of correct answer	S. D.
Item 1	be	0.804	0.397
	become	0.526	0.500
	turn	0.936	0.245
	make	0.093	0.290
Item 2	becoming	0.236	0.425
	getting	0.783	0.413
item 3	becoming	0.211	0.409
	being	0.702	0.458
	getting	0.357	0.480
	turning	0.506	0.500
item 4	became	0.245	0.430
	turned	0.688	0.463
	made	0.804	0.397
Item 5	had	0.672	0.470
	got	0.261	0.440
	took	0.333	0.472
Item 6	get	0.310	0.463
	take	0.562	0.496
	reach	0.774	0.419
Item 7	get	0.430	0.495
	take	0.594	0.491
	reach	0.760	0.427
Item 8	get	0.295	0.456
	take	0.479	0.500
	reach	0.275	0.447
Item 9	get	0.663	0.473
	take	0.666	0.472
Item 10	get	0.658	0.475
	take	0.658	0.475
Item 11	have	0.736	0.441
	get	0.138	0.345
	take	0.663	0.473

<sup>7</sup> Note that in some items, choosing the choice as the answer (or answering that this is a correct choice) is counted, whereas in other items doing so is not. Hence, some of the percentages of items answered correctly represent the percentages that the test-taker can correctly avoid the wrong choice.



Item 12	have	0.527	0.500
	get	0.818	0.386
	take	0.451	0.498
Item 13	get	0.696	0.460
	take	0.702	0.458
Item 14	did	0.708	0.455
	made	0.723	0.448
Item 15	does	0.632	0.483
	makes	0.647	0.478
Item 16	do	0.657	0.475
	make	0.649	0.478
Item 17	do	0.761	0.427
	make	0.758	0.428
Item 18	do	0.611	0.488
	make	0.631	0.483
Item 19	did	0.669	0.471
	made	0.667	0.472
Item 20	did	0.518	0.500
	made	0.509	0.500

#### 4.4.3 Rank setting of LRT

In analyzing data by Exametrika, we have to set the rank setting. We expected that 8 ranks would be appropriate number of ranks considering the 8 word frequency levels in JACET 8000 or that 7 would be appropriate number of ranks because 7 different university students participated in this experiment. However, Akaike's information criteria showed that it was appropriate to set 6 ranks to categorize the test-takers (AIC = 183.050). Hence, 6-rank setting was adopted to analyze the results<sup>8</sup>.

Table 6: Information Criteria in the results of 7 to 8 rank settings

	Rank5	Rank 6	Rank 7	Rank 8
AIC	239.063	183.050	198.163	223.970
CAIC	-3274.028	-3154.386	-2963.619	-2762.158
BIC	-2634.028	-2546.386	-2387.619	-2218.158

*Note:* AIC represents Akaike's information criterion; CAIC, Consistent Akaike information criterion; and BIC, Basian Information Criterion, respectively.

In IRPs, latent ranks represent each achievement level: Rank 1 shows the lowest achievement level, whereas Rank 6, the highest achievement level in this study.

#### 4.4.4 Results of LRT in Part 1

Exametrika produced Item Reference Profiles (IRPs), which represents expectations of each item score at each rank. In IRPs, latent ranks represent each achievement level: Rank 1 shows the lowest achievement level, whereas Rank 6, the highest achievement level in this study.

Table 7 shows IRPs chart of Part 1 in the vocabulary test, where the figures marked with lightest pink showed from 0.40 to 0.59 expectation; those with light pink, 0.60 to 0.79; and those with pink red, more than 0.80. A person in Rank 6 can answer Item 2, 3, 4, 5, 8, 13, 17, 18, 19 and 23 with more than with the probability of 0.8.

Table 7: IRP in Part I in the vocabulary test.

<sup>8</sup> Shojima (n.d.) pointed out CAIC or BIC would be better than AIC in setting the number of ranks. However, it was difficult to interpret the results if the number of ranks more than 6. Hence, we adopted 6 rank settings in this study.

Selected Papers of the 18th Conference of Pan-Pan Pacific Association of Applied Linguistics

Item No.	IRP					
	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
Item 13	0.599	0.741	0.858	0.908	0.928	0.936
Item 18	0.550	0.672	0.813	0.904	0.948	0.969
Item 8	0.589	0.677	0.771	0.853	0.910	0.936
Item 23	0.575	0.681	0.779	0.835	0.855	0.835
Item 4	0.388	0.515	0.663	0.762	0.815	0.835
Item 2	0.319	0.411	0.566	0.733	0.854	0.912
Item 17	0.391	0.451	0.566	0.677	0.766	0.831
Item 19	0.250	0.363	0.550	0.730	0.838	0.890
Item 3	0.364	0.440	0.552	0.669	0.760	0.813
Item 5	0.311	0.406	0.518	0.628	0.742	0.832
Item 14	0.346	0.371	0.427	0.522	0.630	0.720
Item 31	0.226	0.315	0.453	0.585	0.665	0.701
Item 28	0.212	0.266	0.385	0.529	0.658	0.739
Item 32	0.282	0.341	0.432	0.517	0.581	0.620
Item 30	0.220	0.298	0.403	0.511	0.588	0.640
Item 10	0.226	0.285	0.361	0.453	0.572	0.670
Item 20	0.142	0.183	0.260	0.403	0.606	0.762
Item 16	0.260	0.262	0.279	0.345	0.476	0.603
Item 6	0.259	0.261	0.285	0.332	0.424	0.521
Item 25	0.179	0.191	0.239	0.345	0.490	0.606
Item 22	0.284	0.286	0.296	0.326	0.385	0.441
Item 21	0.183	0.184	0.221	0.318	0.466	0.592
Item 7	0.253	0.243	0.244	0.293	0.397	0.500
Item 12	0.171	0.170	0.215	0.320	0.458	0.563
Item 9	0.136	0.147	0.184	0.278	0.445	0.600
Item 15	0.191	0.198	0.206	0.270	0.392	0.508
Item 29	0.210	0.237	0.267	0.307	0.349	0.397
Item 27	0.241	0.271	0.305	0.306	0.269	0.230
Item 24	0.133	0.135	0.157	0.233	0.370	0.506
Item 1	0.244	0.239	0.212	0.197	0.228	0.291
Item 11	0.231	0.221	0.209	0.203	0.200	0.211
Item 26	0.222	0.173	0.121	0.103	0.141	0.201

In the case of Item 8, 13, 18 and 23, the probability of answering this item correctly by a person in Rank 1 was from 0.40 to 0.59, and that of a person in Rank 2, 0.60 to 0.79. On the other hand, the probability of answering this item correctly by a person in Rank 3 was more than 0.80 in Item 13 and 18; but it dropped to 0.60 to 0.79 in Item 8 and 23.

Table 8 shows IRPs from Rank 4 to Rank 6 with the combination of features in each item. We found that items with high probabilities for correct answers contained simple lexical features, whereas those with low probabilities for correct answers had complex lexical features, namely the mixture of two or more lexical features. Table 8 also shows that item with high probabilities for correct answers consisted of relatively high frequency words with comparison to those with low probabilities.

Table 8: IRPs from Rank 4 to Rank 6 with the combination of features in each item

Item No	Rank 4	Rank 5	Rank 6	Targeted features	WL of Stimulus	WL of Target
Item 18	0.908	0.928	0.969	antonyms	1	3
Item 8	0.904	0.948	0.936	Synonyms	5	1
Item 13	0.853	0.910	0.936	antonyms	1	2

Selected Papers of the 18th Conference of Pan-Pan Pacific Association of Applied Linguistics

Item 2	0.835	0.855	0.912	Synonyms		1	1
Item 19	0.762	0.815	0.890	antonyms		4	3
Item 4	0.733	0.854	0.835	Synonyms		4	1
Item 23	0.677	0.766	0.835	collocations			1
Item 5	0.730	0.838	0.832	Synonyms		1	4
Item 17	0.669	0.760	0.831	antonyms		4	4
Item 3	0.628	0.742	0.813	Synonyms		4	2
Item 20	0.522	0.630	0.762	antonyms	derivational form	4	5
Item 28	0.585	0.665	0.739	Synonyms	polysemy	1	2
Item 14	0.529	0.658	0.720	antonyms	derivational form	2	5
Item 31	0.517	0.581	0.701	antonyms	polysemy	1	2
Item 10	0.511	0.588	0.670	Synonyms		5	2
Item 30	0.453	0.572	0.640	Synonyms	polysemy	1	4
Item 32	0.403	0.606	0.620	antonyms	polysemy	4	1
Item 25	0.345	0.476	0.606	collocations			1
Item 16	0.332	0.424	0.603	antonyms		4	1
Item 9	0.345	0.490	0.600	Synonyms		5	1
Item 21	0.326	0.385	0.592	antonyms	derivational form	5	NA
Item 12	0.318	0.466	0.563	antonyms	derivational form	2	0
Item 6	0.293	0.397	0.521	Synonyms		4	2
Item 15	0.320	0.458	0.508	antonyms	derivational form	2	8
Item 24	0.278	0.445	0.506	collocations			4
Item 7	0.270	0.392	0.500	Synonyms		5	1
Item 22	0.307	0.349	0.441	collocations	polysemy		1
Item 29	0.306	0.269	0.397	Synonyms	polysemy		3
Item 1	0.233	0.370	0.291	Synonyms		4	3
Item 27	0.197	0.228	0.230	collocations			5
Item 11	0.203	0.200	0.211	Synonyms		5	4
Item 26	0.103	0.141	0.201	collocations			5

Note: Targeted features represents targeted features in vocabulary knowledge; WL of Stimulus, the word frequency level in JACET 8000 of the stimulus words; and WL of Target, the word frequency level in JACET 8000 of the target words.

#### 4.4.5 Results of LRT in Part 2

Table 9 is Item Reference Profiles (IRPs) in Part 2 in the vocabulary test. The test items in Part 2 were all questions asking knowledge of basic verbs in English. These verbs are all high frequent words. However, the results in Table 9 shows that even the test-takers in Rank 6 had problems in lexical knowledge on basic verbs.

Table 9: IRPs in Part 2.

Item	IRP					
	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
Item-1-turn	0.928	0.936	0.938	0.936	0.937	0.943
Item 12-get	0.723	0.747	0.797	0.837	0.873	0.911
Item1-be	0.734	0.778	0.827	0.844	0.832	0.825
Item 4-made	0.669	0.698	0.756	0.835	0.904	0.943
Item-2-getting	0.609	0.676	0.748	0.822	0.897	0.943
Item 6-reach	0.813	0.809	0.791	0.768	0.743	0.727
Item 17-do	0.612	0.628	0.686	0.783	0.882	0.941
Item 7-reach	0.783	0.794	0.786	0.765	0.736	0.714
Item 17-make	0.571	0.612	0.701	0.805	0.894	0.943
Item 11-have	0.581	0.625	0.698	0.780	0.846	0.877

Selected Papers of the 18th Conference of Pan-Pan Pacific Association of Applied Linguistics

Item 14-made	0.536	0.583	0.643	0.752	0.870	0.936
Item 14-did	0.550	0.577	0.616	0.718	0.842	0.915
Item 3-being	0.610	0.620	0.662	0.727	0.774	0.801
Item 13-take	0.433	0.538	0.674	0.791	0.874	0.913
Item 13-get	0.375	0.514	0.679	0.809	0.893	0.930
Item 4-turned	0.411	0.499	0.630	0.771	0.879	0.931
Item 5-had	0.542	0.627	0.708	0.740	0.732	0.718
Item 19-did	0.614	0.615	0.642	0.681	0.709	0.730
Item 19-made	0.648	0.627	0.633	0.662	0.690	0.714
Item 9-take	0.444	0.555	0.682	0.763	0.789	0.795
Item 9-get	0.410	0.552	0.700	0.776	0.790	0.793
Item 11-take	0.653	0.637	0.630	0.654	0.680	0.701
Item 10-get	0.430	0.511	0.635	0.740	0.798	0.835
Item 10-take	0.395	0.498	0.640	0.751	0.817	0.856
Item 16-do	0.457	0.534	0.623	0.696	0.776	0.841
Item 16-make	0.504	0.548	0.602	0.662	0.745	0.812
Item 15-makes	0.523	0.516	0.555	0.657	0.756	0.827
Item 15-does	0.458	0.483	0.552	0.664	0.763	0.831
Item 18-make	0.485	0.516	0.561	0.631	0.734	0.819
Item 18-do	0.520	0.520	0.531	0.583	0.681	0.775
Item 7-take	0.547	0.558	0.601	0.629	0.624	0.611
Item 6-take	0.574	0.589	0.590	0.573	0.552	0.522
Item 12-have	0.549	0.554	0.548	0.505	0.485	0.510
Item 1-become	0.388	0.443	0.501	0.548	0.612	0.661
Item 20-did	0.442	0.479	0.523	0.537	0.550	0.576
Item 20-made	0.473	0.486	0.504	0.511	0.525	0.548
Item 3-turning	0.223	0.295	0.417	0.557	0.698	0.803
Item 8-take	0.391	0.438	0.493	0.510	0.513	0.531
Item 12-take	0.256	0.337	0.439	0.529	0.577	0.589
Item 7-get	0.378	0.427	0.479	0.479	0.444	0.409
Item 3-getting	0.371	0.384	0.400	0.386	0.335	0.292
Item 5-took	0.227	0.255	0.289	0.328	0.397	0.471
Item 6-get	0.299	0.316	0.320	0.305	0.295	0.318
Item 8-get	0.338	0.312	0.276	0.269	0.288	0.284
Item 8-reach	0.270	0.303	0.313	0.288	0.254	0.242
Item 5-got	0.355	0.322	0.284	0.241	0.200	0.171
Item 4-became	0.327	0.326	0.294	0.229	0.169	0.143
Item 2-becoming	0.383	0.338	0.281	0.211	0.135	0.080
Item 3-becoming	0.235	0.254	0.253	0.220	0.180	0.153
Item 11-get	0.206	0.195	0.162	0.123	0.090	0.068
Item 1-make	0.095	0.114	0.121	0.093	0.071	0.071

To make clear the test-takers' lexical problems in each rank, we reordered IRPs according to Item Number (Table 10). Table 10 clearly shows that higher predictions are found in higher ranks.

Table 10: IRPs reordered according to Item No.

Item	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
Item 1-be	0.734	0.778	0.827	0.844	0.832	0.825
Item 1-become	0.388	0.443	0.501	0.548	0.612	0.661
Item 1-make	0.095	0.114	0.121	0.093	0.071	0.071
Item 1-turn	0.928	0.936	0.938	0.936	0.937	0.943

Selected Papers of the 18th Conference of Pan-Pan Pacific Association of Applied Linguistics

Item-2-becoming	0.383	0.338	0.281	0.211	0.135	0.080
Item-2-getting	0.609	0.676	0.748	0.822	0.897	0.943
Item-3-becoming	0.235	0.254	0.253	0.220	0.180	0.153
Item-3-being	0.610	0.620	0.662	0.727	0.774	0.801
Item-3-getting	0.371	0.384	0.400	0.386	0.335	0.292
Item-3-turning	0.223	0.295	0.417	0.557	0.698	0.803
Item-4-became	0.327	0.326	0.294	0.229	0.169	0.143
Item 4-made	0.669	0.698	0.756	0.835	0.904	0.943
Item 4-turned	0.411	0.499	0.630	0.771	0.879	0.931
Item 5-got	0.355	0.322	0.284	0.241	0.200	0.171
Item 5-had	0.542	0.627	0.708	0.740	0.732	0.718
Item 5-took	0.227	0.255	0.289	0.328	0.397	0.471
Item 6-get	0.299	0.316	0.320	0.305	0.295	0.318
Item 6-reach	0.813	0.809	0.791	0.768	0.743	0.727
Item 6-take	0.574	0.589	0.590	0.573	0.552	0.522
Item 7-get	0.378	0.427	0.479	0.479	0.444	0.409
Item 7-reach	0.783	0.794	0.786	0.765	0.736	0.714
Item 7-take	0.547	0.558	0.601	0.629	0.624	0.611
Item 8-get	0.338	0.312	0.276	0.269	0.288	0.284
Item 8-reach	0.270	0.303	0.313	0.288	0.254	0.242
Item 8-take	0.391	0.438	0.493	0.510	0.513	0.531
Item 9-get	0.410	0.552	0.700	0.776	0.790	0.793
Item 9-take	0.444	0.555	0.682	0.763	0.789	0.795
Item 10-get	0.430	0.511	0.635	0.740	0.798	0.835
Item 10-take	0.395	0.498	0.640	0.751	0.817	0.856
Item 11-get	0.206	0.195	0.162	0.123	0.090	0.068
Item 11-have	0.581	0.625	0.698	0.780	0.846	0.877
Item 11-take	0.653	0.637	0.630	0.654	0.680	0.701
Item 12-get	0.723	0.747	0.797	0.837	0.873	0.911
Item 12-have	0.549	0.554	0.548	0.505	0.485	0.510
Item 12-take	0.256	0.337	0.439	0.529	0.577	0.589
Item 13-get	0.375	0.514	0.679	0.809	0.893	0.930
Item 13-take	0.433	0.538	0.674	0.791	0.874	0.913
Item 14-did	0.550	0.577	0.616	0.718	0.842	0.915
Item 14-made	0.536	0.583	0.643	0.752	0.870	0.936
Item 15-does	0.458	0.483	0.552	0.664	0.763	0.831
Item 15-makes	0.523	0.516	0.555	0.657	0.756	0.827
Item 16-do	0.457	0.534	0.623	0.696	0.776	0.841
Item 16-make	0.504	0.548	0.602	0.662	0.745	0.812
Item 17-do	0.612	0.628	0.686	0.783	0.882	0.941
Item 17-make	0.571	0.612	0.701	0.805	0.894	0.943
Item 18-do	0.520	0.520	0.531	0.583	0.681	0.775
Item 18-make	0.485	0.516	0.561	0.631	0.734	0.819
Item 19-did	0.614	0.615	0.642	0.681	0.709	0.730
Item 19-made	0.648	0.627	0.633	0.662	0.690	0.714
Item 20-did	0.442	0.479	0.523	0.537	0.550	0.576
Item 20-made	0.473	0.486	0.504	0.511	0.525	0.548

This means that there is a general tendency to increase predictions for answering the questions correctly as the rank is higher. This implies that test-takers in a higher rank could acquire lexical knowledge on polysemy, usages and distinction of conceptual differences between L1 and L2 than those in a lower rank could.

## 5 Discussion and Conclusion

Let us turn back our attention to our research questions (RQs): (1) what factors of the lexical knowledge can affect L2 learner's vocabulary, (2) whether lexical difficulties for the L2 learner can change according to types of lexical knowledge, and (3) whether types of lexical difficulties can be predicted by L2 learner's English proficiency level.

For RQ (1), we found that all the lexical knowledge dealt with in the vocabulary test affect the scores of test-takers in all proficiency levels. Especially, word frequencies and vocabulary knowledge on polysemys, collocations, derivational forms can give great effect on lexical acquisition by L2 learners. Zereva (2007) reported that higher proficiency learners provide significantly more word associations than intermediate and beginning level. Our finding can support this tendency reported by Zereva.

Concerning RQ 2 and 3, the finding shows that test-takers in higher ranks have acquired more lexical knowledge on polysemy, usages and distinction of conceptual differences between L1 and L2 than test-takers in lower ranks. In the process of L2 vocabulary acquisition, L2 learners are developing lexical networks in their mental lexicon (Aitchison, 1987; Meara, 2009). Crossley, Salsbury, & McNamara (2009) point out 'Lexical networks are the result of connections between conceptual levels, sense relations, semantic co-referentiality, and word associations' (p.563). Ueda (2011) reported that prototypicality in word meanings and easiness of integrating the conceptual differences between L1 and L2 can affect the feasibility and success in L2 vocabulary acquisition. The findings in the experiment implies that test-takers in higher ranks, (or at higher lexical proficiency levels) could acquire and develop more elaborated lexical networks than those in lower ranks, (or at lower proficiency lexical level). One of the difficulties in development of lexical networks could arise from polysemys, usages and conceptual differences between L1 and L2.

Concerning a new methodology, LRT, it enables us to know what factors of L2 vocabulary can be difficult or easy for L2 learners to acquire; or in other words LRT can produce the achievement progress table in the depth of vocabulary knowledge.

## Acknowledgment

This work was supported by JSPS Grant-in-Aid for Scientific Research (C) (20520516).

## References

- Aichison, J. (1994). *Words in the Mind: an Introduction to the Mental Lexicon. Second Edition*. Oxford: Blackwell.
- Anderson, R. C. & Freebody, P. (1997). Vocabulary Knowledge. In J. H. Guthrie (1997). *Comprehension and Teaching: Research Review*. Newark, DE: International Reading Association
- Crossley, S.A., Slsbury, T., McNamara, D. S., & Jarvis, S. (2010). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing*, 28 (4), 561-580. DOI:10.1177/0265532210378031
- Henriksen, B. (1996). Three dimensions of vocabulary development. *Studies in Second Language Acquisition* 21, 303-17.
- Kimura, T. (2009). Construction and evaluation of an in-house English placement test from a neural test theory perspective. *KATE Bulletin*, 23, 23-34.
- Koizumi, R. & Iimura, H. (2010). Characteristics of Neural Test Theory: Comparison with Classical Test Theory and Rasch Modeling. *JLTA Journal*, 13, 91-109.
- Meara, P. (1996). *The vocabulary knowledge framework*.
- Meara, P. (2009). *Connected Words: Word Associations and Second Language Vocabulary Acquisition*. Amsterdam: John Benjamins Publishing Company.  
Retrieved from <http://www.lognostics.co.uk/vlibrary/meara1996c.pdf>
- Meara, P. and Fitzpatrick, T. (2000). Lex 30: an improved methods of assessing productive vocabulary in an L2. *System* 28, 19-30.
- Mochizuki, M., Uemura, T., Aizawa, K., Sugimori, N., Ishikawa, S., Iso, T., & Koizumi, R. (2010). Development of a Vocabulary Test Battery Estimating English Skills and Proficiency: Integrating Vocabulary Size, Organization, and Access Speed. Report of the Grant-in-Aid for Scientific Research (B) (19320084) supported by Ministry of Education, Science, Sports and Culture.  
Retrieved from <https://kaken.nii.ac.jp/pdf/2010/seika/jsps/32506/19320084seika.pdf>
- Nation, I. S. P. (1990). *Teaching and Learning Vocabulary*. Heinle and Heinle Publishers: Boston.

## Selected Papers of the 18th Conference of Pan-Pan Pacific Association of Applied Linguistics

- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Read, J (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10, 355-371. DOI: 10.1177/026553229301000308
- Shojima., K (2008). *Neural Test Theory: A Latent Rank Theory for Analyzing Test Data*. DNC Research Note, 08-01. Retrieved from <http://antlers.rd.dnc.ac.jp/~shojima/exmk/index.htm>
- Shojima, K. (2011). Exametrika (5.3)[Computer Software]. Retrieved February 1, 2013 from <http://www.rd.dnc.ac.jp/~shojima/exmk/jindex.htm>
- Shojima, K. (n.d.). *Latent Rank Theory*. Retrieved February 1, 2013 from <http://www.rd.dnc.ac.jp/~shojima/ntt/jindex.htm>
- STFC e-Science Word Association Thesaurus. Retrieved from <http://www.eat.rl.ac.uk/>
- Ueda, N. (2011). A Case Study of Developmental Process in L2 Vocabulary Acquisition: Image Schema, Inference and Integration of Concepts in the Mental Lexicon. *Proceedings of the 16<sup>th</sup> conference of Pan-Pacific Association of Applied Linguistics*.
- Ueda, N., Tsutsui, E., Kondo, Y., Oya, M., & Nakano, M. (2010). A Case Study of Developing a Vocabulary Testing: a Progressive Report. *Proceedings of the 15<sup>th</sup> conference of Pan-Pacific Association of Applied Linguistics*.
- Ueda, N., Tsutsui, E., Kodo, Y., Oya, M., & Nakano, M. (2011). A Case Study of Developing a Vocabulary Testing (2): a progressive report. *Proceedings of the 16<sup>th</sup> conference of Pan-Pacific Association of Applied Linguistics*.
- Ueda, N., Tsutsui, E., Kodo, Y., Owada, K., & Nakano, M. (2012). Constructing a vocabulary test to predict learners' proficiency levels. *Proceedings of the 17<sup>th</sup> conference of Pan-Pacific Association of Applied Linguistics*.
- Yokokawa, H. (Ed). (2006). *Nihonjin Eigo Gakushusha no Eitanngogoishinmitsudo Mojihen. (Familiarity in English Vocabulary of Japanese Learners of English (for Letters))*. Tokyo: Kuroshio Shuppan.
- Zareva, A. (2007). Structure of the second language mental lexicon: How does it compare to nativespeakers' lexical organization? *Second Language Research*, 23 (2), 123–153.
- The Edinburgh Associative Thesaurus: <http://www.eat.rl.ac.uk/>

### Appendix A: Test items in Part 1 in the vocabulary test

Item No	Stimulus word	choices			
Item 1	senator	president	congress	representative	minister
Item 2	means	way	meaning	beans	mean
Item 3	proposal	propose	proportion	suggestion	marriage
Item 4	domestic	abuse	foreign	international	home
Item 5	free	wild	liberal	conservative	democrat
Item 6	institute	instruction	construction	organization	nomination
Item 7	commodity	facility	product	essence	common
Item 8	hazard	danger	guess	fog	map
Item 9	profound	deep	foundation	long	impact
Item 10	defect	fault	effect	unnatural	detect
Item 11	grief	sorrow	joy	death	pain
Item 12	active	inactive	unactive	activity	disactive
Item 13	full	empty	fill	complete	employ
Item 14	direct	undirect	disdirect	indirect	bidirect
Item 15	perfect	unperfect	disparfect	imperfect	misparfect
Item 16	vice	main	sin	virtue	president

Selected Papers of the 18th Conference of Pan-Pan Pacific Association of Applied Linguistics

Item 17	internal	inner	terminal	extra	external
Item 18	major	minor	general	captain	colonel
Item 19	rural	urban	country	farm	area
Item 20	relevant	irrelevant	relevant	unrelevant	direlevant
Item 21	competence	uncompetence	discompetence	incompetence	illcompetence

Item 22	The new trainees are still very ( )	pink	green	pale	orange
Item 23	He had been feeling ( ) all week because of his cat's death.	green	blue	pale	red
Item 24	The singer tried to ( ) suicide	make	commit	do	play
Item 25	The software will ( ) your requirements.	meet	see	look	watch
Item 26	The economic ( ) has been lifted	forbid	authority	sanction	function
Item 27	He had a guilty ( ) about what he did.	conscience	consciousness	consequence	constituent
Item 28	I would like to <u>book</u> a table for two people for 8 o'clock.	reserve	read	preserve	observe
Item 29	Kim chose the <u>advanced</u> course of French	upper	intermediate	lower	developed
Item 30	The doctor checked his <u>patient's</u> health.	visitor	customer	guest	client
Item 31	The kitchen on the boat is <u>minute</u> .	tiny	hour	time	huge
Item 32	His father <u>declined</u> his offer.	accept	reject	increase	refuse

Appendix B: Test Items in Part 2 in the vocabulary test

Item No.	Questions	Choices			
Item 1	Becky will ( ) a good doctor.	be	become	turn	make
Item 2	Things are ( ) worse.	becoming	getting		
Item 3	The leaves are ( ) red in fall.	becoming	being	getting	turning
Item 4	The signal ( ) red.	became	turned	made	
Item 5	John ( ) a vacation.	had	got	took	
Item 6	Please ( ) the box for me.	get	take	reach	
Item 7	Please ( ) the box to me.	get	take	reach	
Item 8	Please ( ) me the box.	get	take	reach	
Item 9	Mary will ( ) the degree.	get	take		
Item 10	Mary will ( ) a math course.	get	take		
Item 11	Let's ( ) lunch.	have	get	take	



Selected Papers of the 18th Conference of Pan-Pan Pacific Association of Applied Linguistics

Item 12	Let's ( ) a break.	have	get	take	
Item 13	Naomi will ( ) a high score.	get	take		
Item 14	Sarah ( ) a decision.	did	made		
Item 15	Tom ( ) sports regularly.	does	makes		
Item 16	Mike will ( ) some exercise tomorrow.	do	make		
Item 17	Donald will ( ) an effort to spend more time with his family.	do	make		
Item 18	Victoria will ( ) a speech at the party.	do	make		
Item 19	Blair ( ) some reading.	did	made		
Item 20	Takashi ( ) some research about the college.	did	made		