

RECEPTIVE AND PRODUCTIVE VOCABULARY SIZE OF
JAPANESE UNDERGRADUATE STUDENTS:
A CORRELATION WITH READING AND WRITING PROFICIENCY



Presented in partial fulfillment of the requirements for the
Master of Arts Degree in Teaching English as a Foreign Language
at Srinakharinwirot University

March 2009

Copyright 2009 by Srinakharinwirot University

RECEPTIVE AND PRODUCTIVE VOCABULARY SIZE OF
JAPANESE UNDERGRADUATE STUDENTS:
A CORRELATION WITH READING AND WRITING PROFICIENCY



AN ABSTRACT
BY
RIE SAKAI

Presented in partial fulfillment of the requirements for the
Master of Arts Degree in Teaching English as a Foreign Language
at Srinakharinwirot University

March 2009

Rie Sakai. (2009). *Receptive and Productive Vocabulary Size of Japanese Undergraduate Students: A Correlation with Reading and Writing Proficiency*. Thesis, M.A. (Teaching English as a Foreign Language). Bangkok: Graduate School, Srinakharinwirot University. Advisor Committee: Dr. Saengchan Hemchua, Dr. Aurapan Weerawong.

The purpose of the study was to investigate the relationship between Japanese undergraduate students' receptive and productive vocabulary size at different frequency levels and to examine the correlation with their performance in reading and writing. Participants were 49 Japanese undergraduate students who took the Receptive and Productive Translation Tests (Webb, 2008) that compared their receptive and productive vocabulary size. First, the scores on vocabulary tests were compared, and the relationship between receptive and productive vocabulary size at different word frequencies was examined. Then, scores were analyzed to investigate the relationship between receptive and productive vocabulary size for learners with different receptive vocabulary size. Further, scores on the receptive and productive vocabulary tests were analyzed to determine how loanwords influenced the participants' performance on the vocabulary test. Finally, vocabulary size scores were correlated with scores on a reading comprehension test and free writing.

It was revealed that receptive vocabulary size was larger than productive vocabulary size regardless of word frequency and learners' receptive vocabulary size. It was also found that the learners recognized and retrieved the meaning of loanwords more successfully than non-loanwords. Finally, the correlation analysis revealed that vocabulary size correlated more closely with reading comprehension than writing ability.

ปริมาณคำศัพท์การรับสารและการผลิตสารของนักศึกษาระดับปริญญาตรีชาวญี่ปุ่น:
ความสัมพันธ์ต่อความสามารถทางการอ่านและการเขียน



เสนอต่อบัณฑิตวิทยาลัย มหาวิทยาลัยศรีนครินทรวิโรฒ เพื่อเป็นส่วนหนึ่งของการศึกษา
ตามหลักสูตรปริญญาศิลปศาสตรมหาบัณฑิต
สาขาการสอนภาษาอังกฤษในฐานะภาษาต่างประเทศ
มีนาคม 2552

ริเอะ ซาโก. (2552). ปริมาณคำศัพท์การรับสารและการผลิตสารของนักศึกษาระดับปริญญาตรีชาว

ญี่ปุ่น:ความสัมพันธ์ต่อความสามารถทางการอ่านและการเขียน. ปริญญาโท, ศศ.ม.

(การสอนภาษาอังกฤษในฐานะภาษาต่างประเทศ). กรุงเทพฯ: บัณฑิตวิทยาลัย มหาวิทยาลัย

ศรีนครินทรวิโรฒ. คณะกรรมการควบคุม: ดร.แสงจันทร์ เหมเชื้อ, ดร.อรพรรณ วีระวงศ์.

งานวิจัยชิ้นนี้มีวัตถุประสงค์เพื่อศึกษาความสัมพันธ์ระหว่างการรับและการผลิตคำศัพท์ของนักศึกษาระดับปริญญาตรีชาวญี่ปุ่นในระดับความถี่ที่แตกต่างกัน รวมทั้งศึกษาความสัมพันธ์ต่อความสามารถในการอ่านและการเขียน กลุ่มตัวอย่างเป็นนักศึกษาระดับปริญญาตรีชาวญี่ปุ่น จำนวน 49 คน ที่เข้ารับการทดสอบการรับสารและการผลิตสารของ Webb (2008) เพื่อเปรียบเทียบปริมาณในการรับและผลิตคำศัพท์ งานวิจัยนี้เริ่มจากการเปรียบเทียบคะแนนจากการทดสอบคำศัพท์ และการวิเคราะห์ความสัมพันธ์ระหว่างการรับและการผลิตคำศัพท์ในระดับความถี่ที่แตกต่างกัน จากนั้นนำคะแนนจากการทดสอบมาประเมินเพื่อสำรวจความสัมพันธ์ระหว่างการรับและการผลิตคำศัพท์ของผู้เรียนในระดับการรับสารที่แตกต่างกัน นอกจากนี้จะนำคะแนนจากการทดสอบมาประเมินอิทธิพลของคำยืม (loanwords) ต่อประสิทธิภาพในการทำแบบทดสอบ ท้ายสุดจะนำคะแนนจากการสอบคำศัพท์ การสอบความเข้าใจในการอ่าน และการเขียนอิสระมาวิเคราะห์ความสัมพันธ์

จากการวิจัยพบว่า ปริมาณคำศัพท์ในการรับสารมีมากกว่าการผลิตสาร ในทุกระดับความถี่การใช้คำและการรับคำศัพท์ของผู้เรียน นอกจากนี้ พบว่าผู้เรียนเข้าใจและเรียกคืนความหมายของคำยืม (loanwords) ได้อย่างสัมฤทธิ์ผลมากกว่าคำศัพท์โดยทั่วไป ท้ายสุด จากการวิเคราะห์ความสัมพันธ์แสดงให้เห็นว่าปริมาณคำศัพท์มีความสัมพันธ์กับความเข้าใจในการอ่านมากกว่าความสามารถในการเขียน

The Thesis titled

“Receptive and Productive Vocabulary Size of Japanese Undergraduate Students:
A Correlation with Reading and Writing Proficiency”

by

Rie Sakai

has been approved by The Graduate School as partial fulfillment of the requirements for the
Master of Arts degree in Teaching English as a Foreign Language
of Srinakharinwirot University.

..... Dean of Graduate School
(Associate Professor Dr. Somchai Santiwatanakul)

March 2009

Oral Defense Committee

..... Chair
(Dr. Saengchan Hemchua)

..... Co-advisor
(Dr. Aurapan Weerawong)

..... Reader
(Dr. Wanlapa Thaijinda)

..... Reader
(Dr. Walaiporn Chaya)

Acknowledgements

First and foremost, I am extremely grateful to Dr. Saengchan Hemchua, my thesis advisor, who recommended the idea for this project, encouraged me to adapt the focus and organization of my thesis, and provided valuable suggestions throughout the study.

I would like to express my gratitude to Dr. Aurapan Weerawong, my co-advisor, who has been supportive in every way and has guided me through my full course of study.

Dozens of people on campus have helped me on this thesis. I would especially like to thank Mr. Sompong Wattanapak for his help on the pilot study and completing the pile of documents. I wish to extend my gratitude to Ms. Pranee Burakum, Dr. Panpim Cheaupalakit, Ms. Malin Dejtsak, and other staff members in the International Relations Division and International College for Sustainability Studies. Their support has been a tremendous source of motivation.

I would like to dedicate my thesis to Eisuke Uemoto, who aided me in carrying out the study, assisted in recruiting participants, and facilitated the data collection process in Japan. I am particularly grateful to the research participants who made this work possible. I also acknowledge the statistical help of Yuichiro Kikuchi. I also wish to thank Keiko Yamada for her help in grading the tests.

Finally, special thanks to my family in Japan for their understanding and support.

Rie Sakai

TABLE OF CONTENTS

Chapter	Page
1 INTRODUCTION.....	1
Background	1
Objective of the Study	3
Research Questions	3
Significance of the Study	4
Scope of the Study	4
Definition of Terms	5
2 REVIEW OF THE LITERATURE.....	6
Vocabulary Learning.....	6
Testing Vocabulary Size and Knowledge.....	12
Vocabulary Size Test	16
Writing Assessment	22
Correlation Studies	25
3 METHODOLOGY.....	30
Participants	30

TABLE OF CONTENTS (CONTINUED)

Chapter	Page
3 (CONTINUED)	
Instruments	31
Scoring.....	35
Procedure of the Study	40
Data Analysis	43
4 FINDINGS	46
5 CONCLUSION AND DISCUSSION	59
Objective of the Study	59
Research Questions	59
Procedures	60
Research Findings.....	61
Discussion of the Findings.....	62
General Recommendations	72
Limitations of the Study.....	74
Suggestions for Further Research	75
REFERENCES.....	76

TABLE OF CONTENTS (CONTINUED)

Chapter	Page
APPENDICES.....	81
APPENDIX A.....	82
APPENDIX B.....	84
APPENDIX C.....	91
APPENDIX D.....	93
APPENDIX E.....	95
APPENDIX F.....	101
APPENDIX G.....	103
APPENDIX H.....	105
APPENDIX I.....	107
VITAE.....	109

LIST OF TABLES

Table	Page
1 What is involved in knowing a word	9
2 Two types of format for vocabulary knowledge test	18
3 Comparison of Vocabulary Levels Text and Productive Vocabulary Levels Test	19
4 Comparison with the TOEFL score	33
5 EIKEN Grades and word levels	34
6 Number of words and questions in each passage in EIKEN Grade 2	34
7 Means, standard deviations and ratios for the total receptive and productive scores	46
8 Correlation between receptive and productive scores	48
9 Means, standard deviations and ratios for receptive and productive scores at different word frequency bands	49
10 Means, standard deviations and ratios for the participants with the highest and lowest receptive scores at different word frequency bands using sensitive scoring	51
11 Means, standard deviations and ratios for the participants with the highest and lowest receptive scores at different word frequency bands using strict scoring	52
12 Number and percentage of loanwords at each word frequency band	53

LIST OF TABLES (CONTINUED)

Table	Page
13 Percentage of correct responses for loanwords and non-loanwords	54
14 Percentage of correct responses for loanwords and non-loanwords at each word frequency band.....	55
15 Correlation of receptive and productive vocabulary with reading comprehension	56
16 Correlation of receptive and productive vocabulary with reading comprehension for the participants with the highest and lowest receptive scores.....	57
17 Correlation of receptive and productive vocabulary with free writing	57
18 Correlation of receptive and productive vocabulary with free writing for the participants with the highest and lowest receptive scores.....	58
19 Percentage of correct responses for total receptive and productive scores	63
20 Comparison of correlations between vocabulary size and reading comprehension.....	70
21 Comparison of correlation between vocabulary size and writing ability	71

LIST OF FIGURES

Figure		Page
1	A mastery of basic spelling.....	10
2	A tentative model showing factors affecting the holistic scores of written compositions....	25
3	Word bands on the Receptive and Productive Translation Tests.....	31
4	Upper and lower groups of the receptive translation test.....	44
5	Receptive and productive scores at each frequency band	49
6	Ratio of productive to receptive scores at each frequency band	49
7	Receptive and productive scores for the participants with the highest and lowest receptive scores at different word frequency bands using sensitive scoring	51
8	Receptive and productive scores for the participants with the highest and lowest receptive scores at different word frequency bands using strict scoring	52
9	Percentage of correct responses for loanwords and non-loanwords	54
10	Percentage of correct responses for loanwords and non-loanwords at each frequency band using the sensitive scoring method.....	55
11	Percentage of correct responses for loanwords and non-loanwords at each frequency band using the strict scoring method	55

Chapter 1

Introduction

1.1. Background

Vocabulary size has been of considerable interest to researchers in the field of English as a second language/English as a foreign language (ESL/EFL) for over three decades. There have been many studies trying to determine the vocabulary size of people. Some of the studies have focused on the vocabulary size of native speakers (Goulden, et al., 1990) and second language (L2) learners (Nonaka, 2004); others have focused on aspects of receptive vocabulary and productive vocabulary (Laufer, 1998; Laufer & Nation, 1995; 1999; Laufer & Paribakht, 1998, Schmitt & Meara, 1997; Waring, 1997, Webb, 2008) and the relationship of vocabulary size with reading and writing proficiency (Beglar & Hunt, 1999).

Vocabulary knowledge can be viewed from various aspects, especially from quantitative and qualitative perspectives – *breadth* and *depth*. Vocabulary *breadth* refers to how many words learners know, while *depth* refers to how well those words are known (Nation, 2001). Vocabulary breadth is generally regarded as a good indicator for overall language performance, and in general, the larger the students' L2 vocabulary, the better they are likely to perform.

An average native speaker of English is estimated to know 15,000-20,000 word families. Approximately 2,000 most frequent word families cover more than 95% of spoken language (Schonell et al., 1956, as cited in Nation, 2001). These 2,000 word families cover about 87% of written texts and about 80% of academic texts (Nation, 1990). Therefore, learners of L2 need to know about 3,000-5,000 word families to comprehend most authentic texts (Nation & Waring, 1997) and about 4,000-5,000 word families for academic texts

(Sutarsyah, et al., 1994). Moreover, a minimum of 10,000 word families is required to study at the university level (Hazenberg & Hulstijn, 1996).

A number of studies have revealed that receptive vocabulary is larger than productive vocabulary by using the Vocabulary Levels Test (Nation, 1983) and the Productive and the Vocabulary Levels Test (Laufer & Nation, 1999) to compare receptive and productive vocabulary size (Laufer, 1998; Schmitt, 1997; Waring, 1997). By making use of these tests, the size of an L2 learner's vocabulary has been found to correlate closely with reading comprehension (Beglar & Hunt, 1999; Laufer, 1992; Qian, 1999) as well as with writing ability (Beglar & Hunt, 1999; Laufer, 1998; Laufer & Nation, 1995).

Recently, an article examining a unique test appeared in the Japanese context. The test was designed to measure the receptive and productive vocabulary size of Japanese learners of English. The Receptive and Productive Translation Tests (Webb, 2008) feature a translation format for both receptive and productive vocabularies, containing 180 target words from three different frequency bands. One may be puzzled about what is new with this test, but the target words were selected carefully to tackle the disadvantages of traditional translation tests. However, no study has been conducted so far to investigate this test as it is quite new.

Although considerable studies have been done to date, more studies need to be conducted to investigate the relationship between receptive and productive vocabulary size as well as to determine the correlation between vocabulary size and overall reading comprehension and writing proficiency. The current study attempts to confirm some of the findings of previous studies regarding the relationship between learners' receptive and productive vocabulary size at different frequency levels. In addition, this study applies the

Receptive and Productive Translation Tests (Webb, 2008) to examine the correlation with learners' performance in reading and writing.

It is expected that an outcome of the study will contribute to providing practical measurements for EFL teachers to determine how much their students know and what more they need to learn. Moreover, it will help teachers and learners in considering how receptive/productive vocabulary size corresponds with reading and writing ability.

1.2. Objective of the Study

The purpose of the study was to investigate the relationship between Japanese undergraduate students' receptive and productive vocabulary size at different frequency levels and to examine the correlation with their performance in reading and writing.

1.3. Research Questions

1. Is receptive vocabulary size larger than productive vocabulary size?
2. How does the relationship between receptive and productive vocabulary size change at different word frequency levels?
3. How does the relationship between receptive and productive vocabulary size change for learners with different receptive vocabulary size?
4. Is there any difference in vocabulary size between loanwords and non-loanwords?
5. How does receptive/productive vocabulary size correlate with reading comprehension?
6. How does receptive/productive vocabulary size correlate with writing ability?

1.4. Significance of the Study

An investigation of learners' receptive and productive vocabulary size helps teachers determine how much their learners know and what more they need to learn. The results could be used to set goals for their learners, evaluate a level of achievement, analyze a syllabus or teaching methods, and decide what more to teach to accomplish learners' goals. This study will be beneficial for teachers in providing practical tools to assess and measure learners' vocabulary size, both receptively and productively.

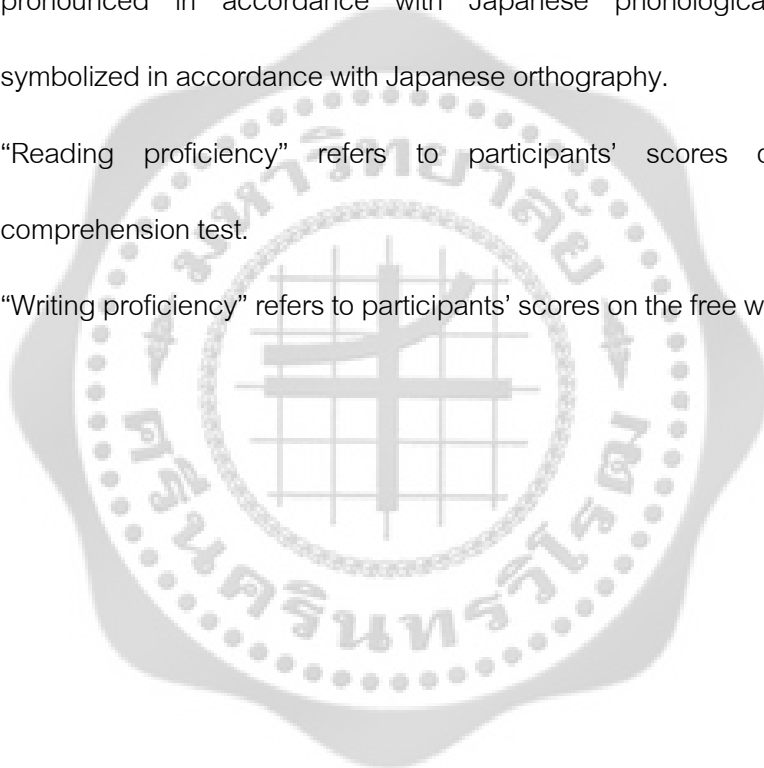
1.5. Scope of the Study

The study is limited to the following.

1. The study set out to manage a group consisting of 49 Japanese undergraduate students who were not majoring in English.
2. The study focused on receptive and productive vocabulary size and the correlation between vocabulary size with reading comprehension and writing ability, but not with listening and speaking proficiency.
3. The study was intended to investigate the breadth of vocabulary knowledge – how many words learners knew, but was not concerned with depth – how well those words were known.

1.6. Definition of Terms

1. “Receptive vocabulary size” refers to the number of words that a learner can recognize in its written form and recall its meaning when met.
2. “Productive vocabulary size” refers to the number of words that a learner can retrieve and produce appropriately in written form when needed.
3. “A loanword” refers to a Japanese word that was borrowed from English. It is pronounced in accordance with Japanese phonological structure and symbolized in accordance with Japanese orthography.
4. “Reading proficiency” refers to participants’ scores on the reading comprehension test.
5. “Writing proficiency” refers to participants’ scores on the free writing essay.



Chapter 2

Review of the Literature

There are two main purposes of the present study. The first is to investigate the relationship between receptive and productive vocabulary size at different frequency bands. The second is to examine the correlation between receptive and productive vocabulary size and scores on reading comprehension test and writing performance. The related literature will be discussed in the following order; (1) vocabulary learning, (2) testing vocabulary size and knowledge, (3), vocabulary size test (4) writing assessment, and (5) correlation studies.

2.1. Vocabulary Learning

This section discusses various aspects of vocabulary learning; word frequency, word knowledge, and form-meaning association.

2.1.1. Word frequency

It is estimated that an average native speaker of English knows 15,000-20,000 word families. However, this is far beyond the goal of an average L2 learner. Then, a question arises, "How many words are needed to do the things that a language user needs to do?" (Nation & Waring, 1997:6). Clearly, not all words are equally useful. Teachers and learners, therefore, need to decide which vocabulary should be learned. One indicator is word frequency – "how often the word occurs in normal use of the language" (Nation & Waring, 1997:8).

A study of word frequency reveals that knowledge of the 2,000 most frequent word families is required for basic communication (Schonell et al., 1956, as cited in Nation, 2001). These 2,000 word families make up about 87% of written texts and about 80% of typical

academic texts (Nation, 1990). The studies looking at the relationship between vocabulary size and reading comprehension support the threshold hypothesis, which states that there is a threshold of vocabulary necessary for comprehending written texts (Laufer & Sim, 1985, as cited in Nation, 2001). Hu and Nation (2000) suggest that there should not be more than one unfamiliar word out of every 20 words for a reader to comprehend the text adequately.

An L2 learner needs to know about 3,000-5,000 word families to comprehend most authentic texts (Nation & Waring, 1997). In addition, about 4,000-5,000 word families are necessary for comprehending more academic texts, such as an economics textbook (Sutarsyah, et al., 1994). Furthermore, Hazenberg and Hulstijn (1996) estimate that a base of at least 10,000 word families, half of what an average adult native speaker knows, is needed to study at the university level.

Investigating the progress of learners' vocabulary size can be considerably valuable to language research and pedagogy. For example, teachers might want to know how much instruction is needed before learners have reached the vocabulary threshold level necessary for the comprehension of an authentic text. Information on learners' vocabulary size at different stages of instruction can help teachers assess whether or not a syllabus of vocabulary learning fulfills learners' goals.

2.1.2. Word knowledge

The first step in understanding vocabulary learning is to specify what it means to know a word. There are many dimensions to “knowing a word” and many degrees of knowledge (Nation, 2001). In general, knowing a word involves form, meaning and use. However, determining exactly what it means to know a word is a complex issue. Many researchers have suggested a great number of aspects of word knowledge. These include receptive and productive knowledge of its spoken and written form, its grammatical and collocation behavior, its frequency and structure, its stylistic register constraints, its conceptual, referential, affective and pragmatic meanings, and its lexical relations to other words (Laufer, 1997; McCarthy & Carter, 1997; Melka, 1997; Nation, 1990; Schmitt, 1998; Schmitt & Meara, 1997).

The receptive/productive distinction is one of best known aspects. For example, learners have receptive knowledge of a word if they are able to retrieve the appropriate meaning when reading or hearing the word (receptive skills). Learners have productive knowledge of a word if they are able to produce the appropriate word when speaking or writing (productive skills). Nation (2001) describes various degree of word knowledge (see Table 1).

Table 1 What is involved in knowing a word (Nation, 2001: 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 recognizable in this word?
		P	What word parts are needed to express this meaning?
Meaning	form and meaning	R	What meaning does this word form signal?
		P	What word form can be used to express this meaning?
	concept and referents	R	What is included in the concept?
		P	What items can the concept refer to?
	associations	R	What other words does this make us think of?
		P	What other words could we use instead of this one?
Use	Grammatical functions	R	In what patterns does the word occur?
		P	In what patterns must we use this word?
	collocations	R	What words or types of words occur with this one?
		P	What words or types of words must we use with this one?
	Constraints on use (register, frequency)	R	Where, when, and how often would we expect to meet this word?
		P	Where, when, and how often can we use this word?

Note: R = receptive knowledge
P = productive knowledge

Laufer (1998), Laufer and Nation (1999), and Laufer and Paribakht (1998) examined three components of word knowledge; (1) passive, (2) controlled active, and (3) free active. Passive knowledge is basic receptive knowledge such as understanding the most frequent and core meaning of a word. Productive vocabulary knowledge was divided into *controlled* production and *free* production. This contrast distinguished learners' ability to produce a word when required to do so by a teacher or test (*controlled*) and their actual production when writing or speaking on their own (*free*). Controlled active knowledge is the ability to use a word when instructed to do so by a teacher or researcher. It involves producing words when prompted by a task (cue recalled). Free active or free productive knowledge, on the other

hand, is the ability to use a word through a learner's free will without any specific prompts for particular words, such as in a free essay, letters, reports, and oral presentations. The distinction between controlled and free active vocabulary has been made because those who can use particular words when required by others are not necessarily able to do so through the use of their own free will.

Vocabulary knowledge is not an "all-or-nothing" (Laufer, 1998: 256) or a "yes/no" (Laufer & Nation, 1999: 26) phenomenon, but involves degree of knowledge. For example, a learner might know the core meaning sense of a word but might not necessarily know all of the other possible meanings (Schmitt, 1998). In another case, a learner may be able to recognize a word and retrieve its meaning when encountered in a text but may not be able to recall how to spell it when asked in a test. Schmitt (2007) conceptualizes a mastery of basic spelling along a continuum for each aspect of word knowledge (see Figure 1). He claims that learning a word must be an incremental and gradual process, as various types of word knowledge are acquired at different times and rates.

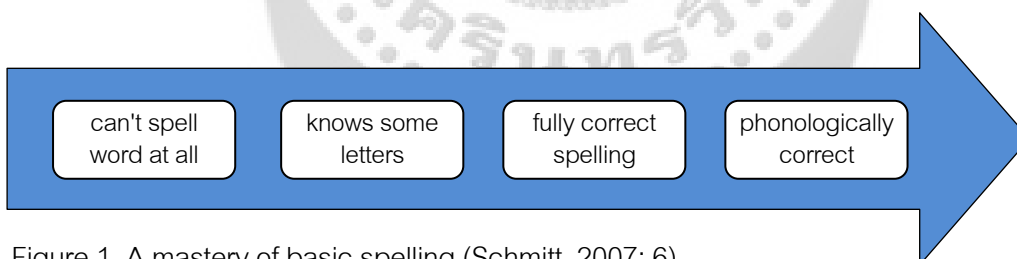


Figure 1 A mastery of basic spelling (Schmitt, 2007: 6)

2.1.3. Form-meaning association

Learners often think of knowing a word as knowing its spoken form or written form and its meaning. However, learners need to be able to connect its form and meaning instead of just knowing the two (Nation, 2001). The stronger the connection that exists between form and meaning, the easier it is for learners to retrieve the meaning of the word when seeing the word

form.

For some languages, loanwords help reduce the learning burden of making the form-meaning connection. Daulton (1998) notes the enormous number of English loanwords in Japanese: up to 38% of the 2,000 most frequent words of English, and 26% within the University Word List (Xue & Nation, 1984). Daulton's study indicates that loanwords help in the process of learning English, even though learners need to extend the limited meaning that the loanword has in Japanese.

In addition, the move to a syllabic spelling system in Japanese has brought about striking formal changes (Daulton, 1998). There are significant changes to the form of words after they have been borrowed from English, which make learners confused when spelling the original English words. The following are some examples of loanwords in Japanese: *booru* (ball), *Waarudo Shiriizu* (World Series), *depaato* (department store), *pantsu* (pants, meaning underwear), *shiiru* (to seal, meaning a sticker)

In another study regarding loanwords, Tadamitsu (2001) examined the role of "Japanized loanword" in performance on the Vocabulary Levels Test (Nation, 1983). Tadamitsu administered versions A and B of the 2,000, 3,000, and 5,000 word levels to 196 Japanese EFL learners and found that over 40% of participants scored higher on the 3,000 word level than 2,000 word level on Version A, and about 28% of participants also scored higher on Version B. He suggested that the percentage of loanwords in learners' L1 be considered when revising the Vocabulary Levels Test because there was a significant difference in some loanwords.

2.2. Testing Vocabulary Size and Knowledge

In this section, the major distinctions of vocabulary knowledge (breadth vs. depth and receptive vs. productive) will be discussed in relation to vocabulary testing.

2.2.1. Breadth vs. depth of vocabulary knowledge

One of the general issues regarding vocabulary testing is the distinction between vocabulary *breadth*, or size, and vocabulary *depth*. Vocabulary breadth tests measure how many word families a person knows and is usually concerned only with the primary meaning of the words. Two widely used breadth tests are the Vocabulary Levels Test (Nation, 1983) and the Productive Vocabulary Levels Test (Laufer & Nation, 1999). Vocabulary depth tests, on the other hand, measure how well a person knows the words, including synonyms, polysemous words, and collocations. One widely used depth test is the Vocabulary Knowledge Scale (Paribakht & Wesche, 1993).

The distinction between vocabulary breadth and depth, however, is not as clear as it seems on the surface. Qian (1999) concludes from his study with 77 Korean and Chinese speakers that knowledge of primary word meaning can often include knowledge of synonymy, polysemy, and collocation, especially if the words share these elements in the learners' L1. Thus, the notions of depth and breadth of vocabulary overlap to some extent.

As the distinction between vocabulary breadth and depth has become less obvious, vocabulary breadth tests have attracted more attention of L2 researchers and teachers. It is likely that many language teachers have observed that a large L2 vocabulary does not necessarily imply depth of knowledge about those words. Read (2000) claims that a size test is not as superficial as it seems, since it can represent the learner's vocabulary knowledge

more adequately than a vocabulary depth test in which a limited number of words are probed. Vermeer (2001) agrees that there is no good reason to suppose that vocabulary size tests are any less valuable than vocabulary depth tests since growth in vocabulary breadth and depth are dependent on the same factor – frequency of input. She administered a breadth test and a depth test to 1,600 Dutch monolingual and bilingual four-year and seven-year old children. She concludes that there is no real distinction between vocabulary breadth and depth as learning more words results in deeper knowledge of vocabulary.

An advantage of using a breadth test is that it can assess a wide range of words in a relatively short amount of time, making it especially useful for assessing learners with various levels of proficiency.

2.2.2. Receptive vs. productive vocabulary knowledge

A second issue regarding vocabulary testing is the distinction between *receptive* vocabulary and *productive* vocabulary. There has been a common perception among L2 teachers and researchers that one's receptive vocabulary is usually larger than one's productive vocabulary. For example, a learner may recognize the word *impede* and even be able to provide a synonym when encountering it within a context, but he/she may not be able to recall the word while writing or speaking (Laufer et al., 2004).

One of the initial studies that examined the gap between receptive and productive vocabulary was conducted by Morgan and Oberdeck (1930, as cited in Webb, 2008). They were among the first to find that the size of receptive vocabulary exceeded that of productive vocabulary at five levels of word frequency. The results suggest that at lower levels, receptive knowledge increases more quickly than productive knowledge, and at higher levels,

production develops faster than reception but never to the point that it equals reception.

There have been a large number of studies in which receptive and productive vocabulary size was compared using the Vocabulary Levels Test (Nation, 1983) and the Productive Vocabulary Levels Test (Laufer & Nation, 1999).

For example, Laufer (1998) administered the Productive Vocabulary Levels Test (Nation, 1983) to assess the productive vocabulary of twenty-six 16-year olds and twenty-two 17-year olds. She then compared the scores to their scores on the Vocabulary Levels Test (Laufer & Nation, 1999), which assesses receptive vocabulary knowledge. She found that receptive vocabulary size was greater than productive vocabulary size by 200 word families (or 10%) for the 16-year olds and by 950 word families (or 32%) for the 17-year olds. The ratio of productive to receptive vocabulary size was lower for 17-year olds than 16-year olds. Laufer suggests that when learner's receptive vocabulary size is small, it consists of the most frequent and useful words. She/he has no choice but to use those words for expression, thus, their productive knowledge is reinforced. As the receptive vocabulary size expands, less frequent words are learned. If a learner is not encouraged to use them, they may never be activated and, therefore, remain in receptive vocabulary. It may be possible that as receptive vocabulary size grows, the number of words that remain in receptive vocabulary increases, and the gap between receptive and productive vocabulary size widens.

A similar study was conducted by Laufer and Paribakht (1998) in which the Vocabulary Levels Test (Nation, 1983) and the Productive Vocabulary Levels Test (Laufer & Nation, 1999) were administered to assess the receptive and productive vocabulary of 79 Israeli ELF learners and 103 Canadian ESL learners. Both ESL and EFL learners had larger receptive vocabulary size than productive vocabulary size. The ratio of productive to receptive

vocabulary was higher in the EFL group, indicating that a gap between receptive and productive vocabulary size was smaller compared to the ESL group. They also compared vocabulary size of ESL learners based on length of stay in the ELS context (six months and less, seven to 24 months, and 24 months and more) and found that receptive vocabulary size did not change significantly, while productive vocabulary size expanded and the gap between receptive and productive vocabulary size decreased after two years of stay. They concluded that the development of productive vocabulary was slower and less predictable than development of receptive vocabulary. Their study indicates that it is not easy to bring receptive vocabulary into productive use, particularly for low-frequency words.

A slightly different research was conducted by Waring (1997) with the Vocabulary Levels Test (Nation, 1983) and the Productive Vocabulary Levels Test (Laufer & Nation, 1995) but the University Word List (Xue & Nation, 1984) section and 10,000 word level were deleted for both receptive and productive tests. The participants were 76 Japanese freshmen and sophomore students majoring in English. The results revealed that receptive vocabulary size was larger than productive size at all frequency levels. If a word from the highest frequency band was known receptively, it had a 64.5% chance of being known productively. However, it had only a 15.5% chance of being known productively if a word was from a less frequent level.

A different type of vocabulary test was used in Schmitt's and Meara's (1997) study in which they investigated how receptive and productive vocabulary knowledge changed over time in relation to word association knowledge and grammatical suffix knowledge. Thirty-six freshmen and 31 sophomore students in Japanese university as well as 28 Japanese high school students were tested on three word associations, inflectional and derivational suffixes

for each of 20 verbs in the beginning and the end of their academic year. They found that word association knowledge and suffix knowledge correlated with total vocabulary size. The participants gained an average of 330 words over time, and their overall receptive knowledge was from 19 to 25 percentage points higher than productive knowledge. They claim that the difference between receptive and productive vocabulary size and knowledge may vary considerably according to a learner's level of language proficiency. Unlike Laufer (1998), Schmitt and Meara argue that beginner learners are more likely to have a wider range between receptive and productive vocabulary size.

In Webb's most recent study (2008), the relationship between receptive and productive vocabulary size was examined using translation tests. The Receptive and Productive Translation tests were administered to measure receptive and productive knowledge of meaning and form, and were scored at two levels of sensitivity (sensitive and strict). Similar results were found for the relationship between receptive and productive vocabulary size. When responses were scored strictly to measure fuller knowledge of forms, receptive vocabulary size was also found to be larger than productive vocabulary size at each of three word frequency levels (701-1,900, 1,901-3,400, 3,401-6,600 most frequent words), and the difference between receptive and productive knowledge increased as the frequency of the words decreased. However, when responses were scored for partial knowledge of forms, there was little difference in vocabulary size at each frequency band. The findings also indicate that learners who have a larger receptive vocabulary are likely to know more of those words productively than learners who have a smaller receptive vocabulary.

2.3. Vocabulary Size Test

Various kinds of vocabulary tests have been used to measure learners' vocabulary

size and knowledge. In this section, two widely used vocabulary size tests and a translation test will be discussed.

2.3.1. Vocabulary Levels Test

The Vocabulary Levels Test (Nation, 1983) is extensively used to assess the breadth of vocabulary knowledge of ESL and EFL learners. It is a matching test in which participants must select the correct definition or synonym for three words from one of six options as follows.

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

Word families are tested from each of five different frequency groups: 1,000-2,000, 2,000-3,000, 3,000-5,000, and words from the University Word List (Xue & Nation, 1984).

2.3.2. Productive Vocabulary Levels Test

Another vocabulary size test is called the Productive Vocabulary Levels Test (Laufer & Nation, 1999). It is “a productive parallel of the receptive Vocabulary Levels Test” (Nation, 2001: 178) in terms of word frequency levels (2,000, 3,000, University Word List, 5,000, and 10,000). It is a cued recall test that involves participants completing a word in a sentence. To limit the answers to the target vocabulary, the first few letters of the words are provided. An example is shown below.

1. I'm glad we had this opp_____ to talk.
2. Every working person must pay income t_____.
3. The pirates buried the trea_____ on a desert island.

When comparing receptive and productive vocabulary knowledge, the two test items should be either both recognition items, or both recall items. A mixture of more than one format (e.g. recognition format for receptive knowledge and recall format for productive knowledge, see Table 2) causes an inaccurate comparison between receptive and productive vocabulary size (Nation, 2001).

Table 2 Two types of format for vocabulary knowledge test

Recognition format		Recall format	
(Recognizing a correct choice)		(Recalling an answer)	
Choose the meaning of the following word.		Translate the following word into L2.	
(L2)	(L1)	(L1)	(L2)
kaki	a. book b. leg c. face d. fruit	leg	_____

Similarly, Webb (2008) questions test formats that have been used to measure receptive and productive vocabulary size. In three studies of Laufer (1998), Laufer and Paribakht (1998), and Waring (1997), receptive and productive vocabulary size were measured by using the same tests; the Vocabulary Levels Test (Nation, 1983) and the Productive Vocabulary Levels Test (Laufer & Nation, 1999). Webb (2008) carefully compares these two tests and argues against a direct comparison of scores on the tests for the following four reasons (see Table 3).

- (a) Chance of guessing correctly: The receptive vocabulary score is measured by multiple-choice test in which the participants have six choices, whereas the productive test is a cue recall test that requires the participants to complete a word. There is 17 % possibility that the participants choose a correct answer in the receptive test without any vocabulary knowledge. However, there is little or

no chance to guess correctly in the productive test. Therefore, it is likely that the participants score higher in receptive matching test than productive test of cue recall.

- (b) Required knowledge: The Vocabulary Levels Test calls for knowledge of form and meaning, whereas the Productive Vocabulary Levels Test might also require knowledge of grammatical functions. In this sense, the productive test is more demanding.
- (c) Test format: An accurate comparison of receptive and productive tests might be unfavorable due to a difference in test format – recalling an answer and recognizing a correct answer from provided choices.
- (d) Validity of Productive Vocabulary Levels Test: A test of cued recall is not exactly a productive test but a recognition test. Morton (1979, as cited in Melka, 1997: 87) states that “the presence of partial information is often sufficient to recognize a word.”

Table 3 Comparison of Vocabulary Levels Text and Productive Vocabulary Levels Test

	Vocabulary Levels Test (Nation, 1983)	Productive Vocabulary Levels Test (Laufer & Nation, 1999)
Test type	Matching	Cued recall
Chance of guessing correctly	16.6%	Little or none
Test format	Recognition format	Recall format
Required knowledge	Form, meaning	Form, meaning, and grammatical functions

2.3.3. Receptive and Productive Translation Tests

As an alternative, Webb (2008) designed an equivalent test format as a more accurate measurement of receptive and productive vocabulary. The Receptive and Productive Translation Tests aimed to measure receptive and productive vocabulary size especially for Japanese learners of English (see Appendix B).

Webb chose 180 target words in relation to two factors: (a) frequency and (b) overlap between first language (L1) and L2 meanings.

- (a) Frequency: The target words were chosen from the COBUILD dictionary which provides frequency information based on Bank of English Corpus at three different frequency bands: 701-1,900, 1,901-3,400, and 3,401-6,600 most frequent words.
- (b) Overlap between L1 and L2 meanings: The words that had a high degree of overlap with L1 meaning and L2 were selected in order to overcome disadvantages in using translation tests. There was no guarantee that participants would write the particular target words although they knew the forms and meanings of the target words. If target words were polysemous or synonymous, it was then very difficult to ensure whether the participants had knowledge of the target words because such words elicited various responses.

Webb chose only the words that consistently had elicited the target L2 form and words that successfully matched with particular L1 meaning in a pilot study with Japanese EFL learners. Words that elicited various responses were deleted.

Two instruments were used to measure the participants' vocabulary size at three word

frequency levels. In the receptive test, the participants were required to write L1 meanings next to L2 forms as shown below.

- | | | |
|---|--------|-------|
| 1 | island | _____ |
| 2 | animal | _____ |
| 3 | sport | _____ |

In the productive test, the participants were given L1 meanings and asked to write their L2 forms as follows.

- | | | |
|-----|-----|-------|
| (1) | 木曜日 | _____ |
| (2) | 鳥 | _____ |
| (3) | 眠る | _____ |

In order to avoid any learning effect, two versions of the test booklet were prepared. The research design allowed all of the words to be tested receptively and productively and ensured that there was not a learning effect from one test to the other. Responses were scored at two levels of sensitivity (sensitive and strict) to measure partial and fuller knowledge of written form for each test at each level of frequency for all of the participants.

In summary, the most widely used vocabulary breadth tests are the Vocabulary Levels Test (Nation, 1983) and the Productive Vocabulary Levels Test (Laufer & Nation, 1999). In addition, a new entry, the Receptive and Productive Translation Test (Webb, 2008), has been discussed. Unlike the Vocabulary Levels Test and the Productive Vocabulary Levels Test, Webb proposes an equivalent test format so that there is little or no chance of guessing correctly in either the receptive or productive tests. The parallel design of the tests allows an accurate comparison of receptive and productive vocabulary size. Thus, these tests were considered when selecting the instruments for the current study.

2.4. Writing Assessment

One of the purposes of the present study is to examine the correlation between receptive/productive vocabulary size and scores in writing ability. Every piece of participants' writing composition will need to be evaluated appropriately. In this section, some of the scoring methods for L2 writing will be discussed.

2.4.1. Holistic scoring vs. analytic scoring

EFL teachers often have to choose among different types of rating scales to assess L2 writing performance. Rating scales can be classified in terms of approach and scoring method. Cooper (1977, as cited in Barkaoui, 2007) distinguishes between holistic approaches and analytic approaches.

Holistic scales (or rubrics) respond to language performance as a whole. Each score on a holistic scale represents an overall impression, thus one integrated score is assigned to a performance.

Analytic scales are divided into separate categories representing different aspects or dimensions of performance. For example, dimensions for writing performance might include content, organization, vocabulary, grammar, and mechanics. Each dimension is scored separately, and then dimension scores are added to determine an overall score. However, analytic scoring has been criticized because the parts do not necessarily add up to the whole, as Goulden (1992, as cited in Barkaoui, 2007: 87) says, "the whole is not equal to the sum of the parts." Rather, "the whole is equal to the parts and their relationships."

Research has shown that reliable and valid information gained from both analytic and

holistic scoring instruments can tell teachers a great deal about their learners' proficiency levels.

Astica (1993) analyzed 210 writing samples using ESL composition scale (Jacobs et al., 1981.) that involved content, organization, vocabulary, language use, and mechanics. A multiple regression analysis revealed that vocabulary resulted in 84% of the variance and was the largest factor in writing quality.

Different results were revealed in Bacha's (2001) study in which the same scale, ESL Composition Profile (Jacobs et al., 1981), was used to compare analytic and holistic scoring. Thirty essays were rated first holistically and then analytically. The study showed that there were highly significant positive relations between the holistic and the analytic scores using the Spearman Correlation Coefficient Test. This indicates internal consistency or reliability of the scores. Unlike Astica (1993), Bacha (2001) concludes that the raters were focusing equally on each component (content, organization, vocabulary, language, and mechanics) when rating the essays, and did not emphasize any component over the others, as some research suggests may occur.

Another study comparing holistic and analytic scoring was carried out by Song and Caruso (1996) with the use of different scales for each scoring method. They aimed to investigate differences in which professors evaluated essays written by native English speakers (NES) and ESL learners. Two faculty groups (32 English and 30 ELS professors) rated four essays holistically and analytically. The results indicate that in holistic evaluation, English and ESL faculty raters differed significantly. In analytic evaluation, the two groups did not demonstrate significant differences in rating the specifically categorized features. However, this does not imply the superiority of the analytic scale over the holistic scale

because this study only revealed the tendency of each faculty group in evaluating NES and ESL writings.

In a recent study, Barkaoui (2007) investigated how different rating scales affect rater performance. Four EFL teachers in Tunisia rated a set of 24 EFL essays silently and two other sets of four essays while thinking aloud using a holistic scale and then an analytic scale. She identified four analytic scales including Jacobs et al. (1981) and four holistic scales including Song and Caruso (1996) from the literature that were appropriate for marking argumentative essays by EFL university learners. Contrary to the hypothesis that analytic scoring would result in higher score reliability (Goulden, 1992, as cited in Barkaoui, 2007), a higher level of score reliability was achieved when the essays were marked holistically.

A study examining validation of holistic scoring for L2 writing revealed the thinking process and criteria of raters. Sakyi (2000) finds that the rater focused either on errors in the text, essay topic and presentation of ideas, or the rater's personal reaction to the text while reading. There were two different approaches to reading style observed among the raters. Some raters attempted to obtain a total impression of the essay before deciding on the score range (overall impression approach) while others decided on a range from the first paragraph or first few sentences (first impression approach). The figure shown below is Sakyi's tentative model of the holistic scoring (see Figure 2). He also suggests that there should be at least two different raters with essays.

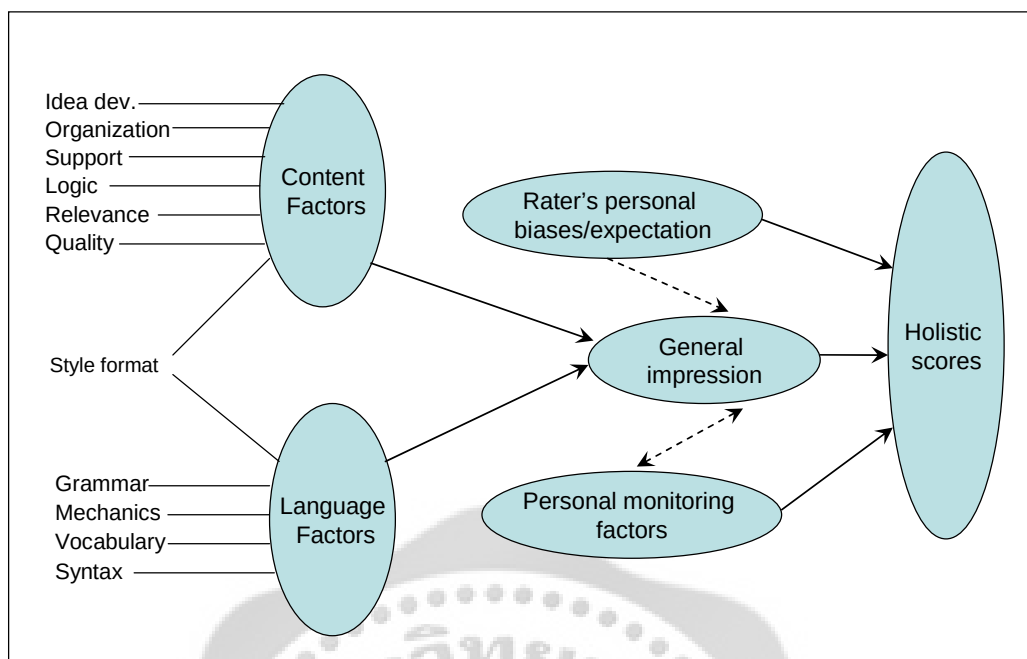


Figure 2 A tentative model showing factors affecting the holistic scores of written compositions (Sakyi, 2000: 146)

In summary, two scoring methods for L2 writing have been discussed. As a considerable number of studies have shown, holistic scoring is reliable and valid for L2 writing assessment, and thus has been chosen to evaluate the participants' essays in the present study.

2.5. Correlation Studies

Vocabulary size has been found to be correlated with general intelligence and with language proficiency. If the correlation between language skills and vocabulary size is high, a vocabulary size test may be a useful instrument to measure language proficiency (Laufer et al., 2004). As many studies have revealed, receptive vocabulary is larger than productive vocabulary. Two well-known vocabulary size tests, the Vocabulary Levels Test (Nation, 1983) and the Productive and the Vocabulary Levels Test (Laufer & Nation, 1999), have been used

to compare receptive and productive vocabulary size at various frequency levels (Schmitt, 1997; Waring, 1997; Laufer, 1998). These two vocabulary tests have been also used to investigate the correlation of vocabulary size with reading and writing proficiency (Beglar & Hunt, 1999; Laufer, 1992, 1998; Qian, 1999). The research indicates that vocabulary size is highly correlated with reading comprehension and writing ability.

2.5.1. Reading comprehension

The following studies show a statistically significant and positive correlation between receptive vocabulary size and reading comprehension.

Laufer (1992) studied the relationship between reading comprehension score and vocabulary size. An English test produced in the Netherlands and an Israeli entrance exam were used to measure reading comprehension, and for vocabulary size. Vocabulary Levels Test (Nation, 1983) and the Eurocentres Vocabulary Test 10KA (Meara & Jones, 1990), were administrated to 92 university freshmen whose native language was either Arabic or Hebrew. Those who scored higher than 55% on a reading test were regarded as successful readers. The number of successful readers exceeded that of unsuccessful readers at the level of 3,000 word families.

Qian (1999) finds stronger correlations in his study with 44 Korean speakers and 33 Chinese speakers using the Vocabulary Levels Test (Nation, 1983) and the Reading Comprehension section from the Test of English as a Foreign Language (TOEFL). The correlation between the two tests was found to be .78.

Similarly, Beglar and Hunt (1999) conducted a study with 496 Japanese high school students using four versions of the 2,000 word level from the Vocabulary Levels Test (Nation,

1983) and 464 learners with four versions of the University Word List (Xue & Nation, 1984) from the same test. They found that scores on the Reading Comprehension subsection of the TOEFL test correlated with scores on versions A and B of the 2,000 word level at .66 and .62, respectively, and with scores on versions A and B of the University Word List section at .67 and .71, respectively.

2.5.2. Writing ability

The following studies applied various methods, but found substantial correlation between vocabulary size and writing ability.

In the study mentioned above, Beglar and Hunt (1999) also compared the TOEFL Structure and Written Expression subsection with two versions of the Vocabulary Levels Test (Nation, 1983), versions A and B, measuring participants' receptive size. The results show that the correlations were .61 and .65 respectively, slightly lower than the correlations found with reading comprehension.

Two studies below compare scores on Productive Vocabulary Levels Test (Laufer & Nation, 1999) and writing performance. Laufer and Nation (1995) proposed a program to measure lexical richness, the Lexical Frequency Profile, which analyzes the proportion of word families at various frequency levels in learners' writings. Two sets of compositions written by 65 learners (22 ESL learners in New Zealand and 43 EFL learners in Israel) were examined on the basis of the number of words that the participants used from the first 1,000 most frequent words, second 1,000 most frequent words, the University Word List (Xue & Nation, 1984), and low frequency or unknown words. The aim of the study was to determine to what extent such a profile would correlate with the learners' scores on the Productive Vocabulary Levels Test. It

was found that the participants who had a larger vocabulary size used less high frequency words and, instead, used low frequency words more often than those with a smaller vocabulary size. Vocabulary scores from the University Word List section of the Productive Vocabulary Levels Test correlated with the number of academic words the participants used in the two sets of compositions at .70 and .60.

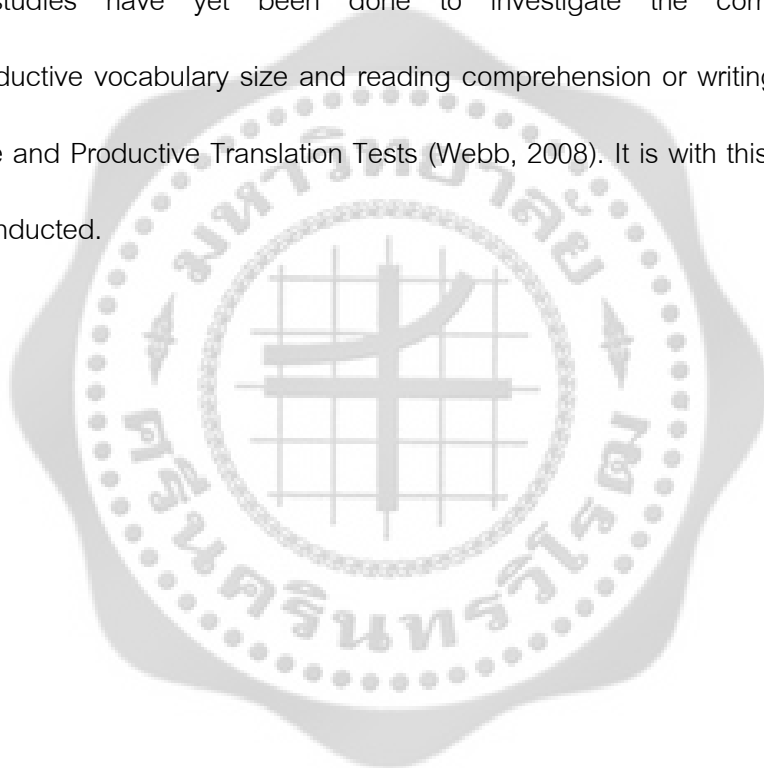
In a later study, Laufer (1998) conducted a similar study in which she correlated participants' scores on the Productive Vocabulary Levels Test (Laufer & Nation, 1999) with the type of vocabulary they used in writing a composition, as analyzed by the Lexical Frequency Profile (Laufer & Nation, 1995). Twenty 16-year olds who had studied English for six years and twenty-two 17-year olds who had studied English for seven years took part in the study. She found that the 17-year olds had larger vocabularies than the 16-year olds. Unlike her previous study, she found that the 17-year olds with larger vocabulary size did not use a greater breadth of words while writing their compositions. She claims that learners may not feel confident enough to use a word when writing or speaking, but instead use another word that is perhaps more frequent, or employ circumlocutions to express the same meaning.

The correlation studies discussed above suggest that vocabulary size should be an important contributor to proficiency in reading and writing.

To sum up, Webb's recent study (2008) proposes using an equivalent translation test to measure receptive/productive vocabulary size. He claims that receptive vocabulary and productive vocabulary are measured by different formats; recognition and recall, or a multiple-choice and a cued recall. A learner with no vocabulary knowledge can guess correctly 17% of the time in the receptive test – the Vocabulary Levels Test (Nation, 1983) –

while it is almost impossible to guess accurately in the productive test – the Productive Vocabulary Levels Test (Laufer & Nation, 1999). The receptive test requires knowledge of form and meaning, but a productive test might also involve grammatical functions. In this sense, the productive test appears to be stricter.

Webb's findings are similar to earlier studies that examined receptive/productive vocabulary size, but further study is needed to assess these recently developed translation tests. No studies have yet been done to investigate the correlation between receptive/productive vocabulary size and reading comprehension or writing ability by using the Receptive and Productive Translation Tests (Webb, 2008). It is with this in mind that this study was conducted.



Chapter 3

Methodology

The purpose of this study was to examine the relationship between receptive and productive vocabulary size and the correlation of receptive/productive vocabulary size with reading comprehension and writing ability.

This chapter is comprised of information concerning participants, instruments, scoring, procedures of the study, and data analysis.

3.1. Participants

The participants were 49 students in Daito Bunka University, Tokyo, Japan. They were native speakers of Japanese and had learned English for a minimum of six years before entering the university. The students were majoring in either economics or business administration, attending their third or fourth year of a four-year degree, and belonged to the Faculty of Business.

English was not a required subject for their degrees. Therefore, some participants had taken an elective English class since entering university and others had not. However, this factor would not affect this study as it was a within-subject design measuring an individual's own receptive and productive vocabulary. Thus, there was no need to test for exact proficiency or prior knowledge.

3.2. Instruments

3.2.1. Receptive and Productive Translation Tests (Webb, 2008)

One hundred eighty target words were selected in relation to two factors: (1) frequency (see Figure 2) and (2) overlap between first language (L1) and second language (L2) meanings.

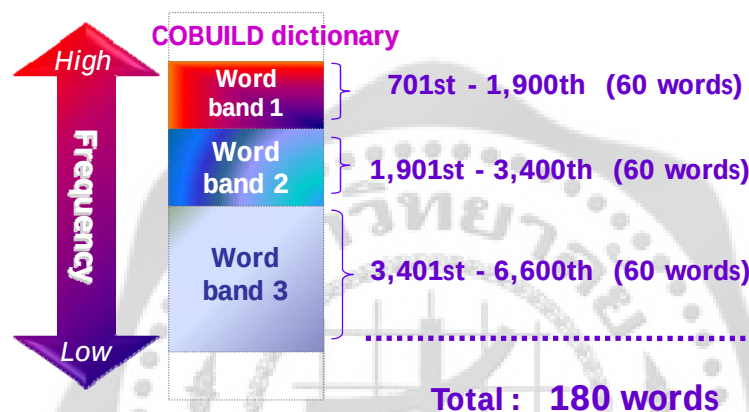


Figure 3 Word bands on the Receptive and Productive Translation Tests

The Receptive and Productive Translation Tests (Webb, 2008) were adapted to measure the participants' receptive and productive vocabulary size at three word frequency levels. In the receptive test, the participants were required to write L1 meanings next to L2 forms. In the productive test, the participants were given L1 meanings and asked to write their L2 forms. At each level, the receptive and productive tests were presented together on one page. Each page had 60 items (30 receptive and 30 productive), for a total of 180 items (90 receptive and 90 productive) on the test (see Appendix A).

If the same words appeared in both the receptive and productive tests, the participants might learn target words on one test and apply it on the other. In order to avoid any learning effect, two versions of the test booklet were prepared (Version A and Version B,

see Appendix B). The 90 L2 target words in the receptive test of Version A were used as the 90 L1 meanings in the productive test of Version B. Likewise, the 90 L1 meanings in the productive test of Version A became the 90 L2 target words in the receptive test of Version B. Half of the participants were randomly chosen to work on Version A, and the other half of the participants took Version B.

3.2.2. EIKEN Grade 2 (Society for Testing English Proficiency, 2008)

To investigate the correlation between scores on receptive vocabulary size and reading comprehension test, the proficiency test below was chosen.

EIKEN (Test in Practical English Proficiency) is Japan's most accessible and affordable English assessment tool for learners at all proficiency levels. It has been specifically designed by the Society for Testing English Proficiency (STEP) to measure the ability of Japanese learners. Backed by the Ministry of Education, Culture, Sports, Science, and Technology, the test is administered locally at 18,000 test sites in 250 cities across Japan three times each year. The number of examinees reaches 2.5 million annually (see Appendix F), for a total of 80 million examinees since 1963.

EIKEN is a pass/fail test conducted in seven bands called "grades." Each grade is a separate test with an individual set of test items and tasks designed for that ability level (see Appendix F). The test is linked with the Japanese school curriculum. Grade 2 has been designated by the Japanese education ministry as a benchmark for high school graduates.

EIKEN is administered in two stages that require four skills. The first stage is paper-based and consists of four main testing areas: (a) vocabulary, (b) reading comprehension, (c) listening comprehension, and (d) writing. The second stage is a speaking

test that is required for all examinees who pass the first stage except for Grades 4 and 5. A personal interview is conducted on-site by examiners who have been trained and certified by STEP.

EIKEN is also offered internationally in 45 countries with public sites including London, Los Angeles and New York. It has been recognized by colleges, universities and institutes in North America and Australia for international admissions.

The four-year study that involved 700 international students from 25 countries in the University of Hawaii at Manoa and Kapi'olani Community College found a strong correlation between EIKEN and Test of English as a Foreign Language (TOEFL), suggesting that EIKEN is a viable alternative. Table 4 shows the TOEFL scores as compared to the EIKEN grades.

Table 4 Comparison with the TOEFL score

EIKEN	TOEFL	
	PBT	iBT
Grade 1	600	100
Grade Pre-1	550	80
Grade 2	500	61
Grade Pre-2	400	32

Note:

- i. The PBT (paper-based test) score ranges between 310 and 677.
- ii. The iBT (new Internet-based test) is scored on a scale of 0 to 120 points.
- iii. PBT-iBT comparisons are taken from TOEFL iBT Scores, Educational Testing Service (2005, as cited in the Society for Testing English Proficiency, 2008)

To investigate the correlation between receptive vocabulary size and reading comprehension, three passages with 12 questions in Grade 2 were selected. Grade 2 was selected for the present study for two reasons.

(a) Grade 2 is considered to be a benchmark for high school graduates; thus it is suitable for undergraduate students.

(b) The Receptive and Productive Translation Tests consist of words from 701-1,900,

1,901-3,400, and 3,401-6,600 frequency levels. Grade Pre-1, comprised of the 7,500 word level, exceeds the range while Grade Pre-2 only covers up to the 3,600 word level. Grade 2 ranks in the middle of these levels, covering 5,100 words (see Table 5).

Each of three passages chosen from Grade 2 consisted of approximately 200-350 words (see Table 6) and was attached with three to five comprehension questions (see Appendix E). The participants were required to complete all the questions within 60 minutes.

Table 5 EIKEN Grades and word levels (Society for Testing English Proficiency, 2008)

EIKEN Grade	Word level	Recommended level
1	10,000-15,000	University
Pre-1	7,500	University
2	5,100	High school graduates
Pre-2	3,600	High school
3	2,100	Junior high school graduates
4	1,300	Grade 8
5	600	Grade 7

Table 6 Number of words and questions in each passage in EIKEN Grade 2

Passage number	Number of words in text	Number of questions
1	216	3
2	367	4
3	356	5
Total	939	12

3.2.3. Free writing

To analyze the correlation between receptive/productive vocabulary size and writing ability, the participants were asked to write an essay within 60 minutes. The participants were assigned to choose one of four topics in the list and express their ideas on the lined page of provided paper. All of the topics were considered to be highly relevant to undergraduate

students in Japan (see Appendices G and H).

The directions used in the study by Song and Caruso (1996) were developed and given to the participants (see Appendix I). It was likely that most participants had not been trained to write essays, and some may have felt reluctant to try their best if the instructions appeared to be too demanding. Therefore, it was not necessary to state that correct grammar and sentence structure were important as stated in the original directions.

3.3. Scoring

3.3.1. Receptive and Productive Translation Tests

The tests were scored by two raters based on the guidelines used in Webb's (2008) study. The scoring system, grammar mistakes, and semantically similar responses need to be discussed in detail as follows.

A correct response received one point and an incorrect response received zero points. All the responses were scored at two levels of sensitivity: sensitive and strict.

Strict scoring system

- Receptive test:

A learner's response was marked as correct if it demonstrated that the learner was able to recall the L1 meaning and it was spelled correctly.

- Productive test:

Responses were only marked as correct if the target words were spelled correctly.

The following two responses were considered to be acceptable because they were commonly used as spoken language in the US as well as taught in Japanese middle school; math for mathematics and gas for gasoline.

Sensitive scoring system

– Receptive test:

Responses that demonstrated L1 knowledge of meaning were marked as correct even if they were misspelled. When the L1 word was misspelled, both graders had to agree that the response demonstrated L1 knowledge of meaning for the target word.

– Productive test:

If it was not certain that the participants had intended to write the target word, it was marked as incorrect. Otherwise, a response was marked as correct. Examples of acceptable and unacceptable responses were as follows.

<Target word>	<Acceptable responses>	<Unacceptable responses>
address	*adress, *adres	*aboless
kitchen	*kitchin, *kichin, *kichen	*kicn
joke	*jork, jorc	*jook, *zhoku
castle	*cassle	*calot, *ckassle, *casr
button	*buppen, *botton, *botten	butter
bridge	*brige	*brich, breath

In Webb's (2008) study, if a target response was *bubble*, *bubbl*, *buble*, *bubel*, and *bubl* were acceptable responses while *babble* and *bobble* were not acceptable.

However, the responses mentioned above were all marked as correct with sensitive scoring in the present study. There were few differences found between sensitive and strict scoring when the responses such as *babble* or *bobble* were all rejected. It was unlikely that the participants had ever encountered *babble* or *bobble* before and that they even knew such words existed. From this perspective, it was reasonable to consider that the learners had intended to write the target word *bubble*. They had at least demonstrated partial knowledge of the target word. Otherwise, it would have not been possible to detect any partial knowledge. The following responses were marked

as correct for the same reason.

<Target word>	<Acceptable response>
cheap	cheep
hug	hag
hammer	hummer
jungle	jangle
drum	dram
cow	caw
bird	bard

However, the following responses were marked as incorrect since it was difficult to ensure that the participants had intended to write the target words.

<Target word>	<Unacceptable response>
weak	week
collar	color
bone	born
ear	year
boots	booth
steal	steel

In addition, Japanese learners are often confused between the letters *l* and *r* because no equivalent consonants exist in Japanese. Although the overall shape of a response was not a close approximation of the target word, responses showing confusion between *l* and *r* were marked as correct because it was clear that the participants had attempted to respond to the target words. Examples of responses marked as correct were;

<Target word>	<Acceptable response>
February	*Februaly
hungry	*hungly
protein	*plotein
hero	*helo
beer	*beel
mirror	*millor
spray	splay
miracle	milacle
pilot	*pirot

license	*ricense
lion	*rion
milk	*mirk
calendar	*carender
pillow	*pirrow
yellow	*yerrrow
translation	*transration

However, *glass* for *grass* was marked as incorrect because it was not certain that it was meant to be the target word.

Grammatical mistakes

– Receptive test:

Responses that demonstrated L1 knowledge of meaning but were in the wrong grammatical form were marked as correct in both strict and sensitive scoring systems since the current study focuses on meaning and form rather than grammatical knowledge of words.

– Productive test:

Responses in the wrong grammatical form were marked as correct for both scoring methods for the following two reasons.

- It was assumed that the participants who responded with *leaves* and *burning* would also know the target words *leaf* and *burn*.
- The translation tests were essentially tests of meaning and form. Although it might be possible to score responses for grammatical form, such scoring might not give an accurate assessment of grammatical knowledge.

Semantically similar responses

There were a small number of target meanings that elicited varied responses.

– Receptive test:

When the participants responded with another sense of a target word's meaning, such

responses were scored as correct. It was likely that the target words were matched with their most common meanings, but less common meanings for L2 target words should be marked as correct because such responses suggested that the participants had a larger knowledge than was being tested. (It was unlikely that a Japanese student would refer to a *spring* as a “spa” without knowing that the most common meaning of *spring* was the season between winter and summer.)

– Productive test:

Responses were marked as correct if target meanings elicited correct L2 translations. Learners were more likely to know more frequent synonyms than less frequent synonyms. If a less frequent synonym was written as a response, it was likely that the participants also knew the more frequent synonym. Examples of semantically similar responses marked as correct were:

<Target word>	<Semantically similar response>
gun	pistol
medicine	drug
sky	air
count	number
clothes	wear
hate	dislike
neighbor	next door

3.3.2. Reading comprehension test

For 12 answers in the reading comprehension test, a correct answer received one point and an incorrect answer received zero points.

3.3.3. Free writing

Participants' essays were evaluated holistically and given points ranging from 1 to 6 according to the Evaluation Scale for the Writing Skills Assessment Test (Song & Caruso, 1996,

see Appendix I). This holistic scale was chosen as it was suitable for participants at the undergraduate level. Some scales appeared to have set overly detailed criteria and were time-consuming to apply (e.g. American Council on the Teaching of Foreign Languages (ACTFL) Proficiency Guidelines, 2001). On the contrary, other scales were found to be so rough that a rater could hardly apply criteria. Therefore, the raters had to depend on their subjective impressions (e.g. Clark, 2003).

Two independent raters assessed all the essays in order to ensure fair representation and evaluation of each participant's writing ability.

3.4. Procedure of the Study

Pilot Study

A pilot study was conducted to assess the Receptive and Productive Translation Tests (Webb, 2008). The participants were four students in the Thai Language Intensive Course, Faculty of Humanities, Srinakharinwirot University, Bangkok. They were native speakers of Japanese and had learned English for at least six years before coming to Thailand. Their ages ranged from 20 to 27. Two of the participants were randomly chosen to work on Version A, and the other two took Version B. The participants were allowed to work on the test as long as they needed (see Appendix D for the results). No changes were made to the tests as a result of the pilot study because the participants adequately responded to the target meanings.

Data Collection

It would have been ideal to collect data within a short period of time to avoid any

changes in the language levels of the participants. However, data collection took place during a regular class period that was allocated only 90 minutes, once a week. For this reason, the overall duration of data collection was three weeks.

The three kinds of test were administrated in the following sequence.

- (1) Receptive and Productive Translation Tests (Webb, 2008)
- (2) Reading comprehension test (EIKEN Grade 2, Society for Testing English Proficiency, 2008)
- (3) Free writing

Generally, productive learning takes longer than receptive learning because it is not easy to bring receptive vocabulary into productive use (Laufer & Paribakht, 1998; Nation, 2001). Therefore, it was unlikely that the participants' writing ability changed within the two weeks after the first test was administered.

Procedure

The procedure included the following steps of tests and evaluation.

1. In the first week, the participants were asked to participate in the present study. To encourage the participants to try their best, they were informed that they would gain extra points in their final examinations based on the scores in the tests.
2. In the second week, the vocabulary translation tests were administered. Two versions of the test booklet were prepared in order to avoid any learning effect of the target words. A total of 180 target words (90 receptive and 90 productive) were presented in three pages; each page consisted of 60 target words (30 receptive and 30 productive). The participants might learn target words from one test and apply it on the other test if the same words appeared in both the receptive and productive tests. The 90 L2 target words

in the receptive test of Version A appeared as the 90 L1 meanings in the productive test of Version B. Similarly, the 90 L1 meanings in the productive test of Version A were used as the 90 L2 target words in the receptive test of Version B. Half of the participants were randomly chosen to work on Version A, and the other half of the participants took Version B. The participants were allowed to work on the task as long as they needed, since the tests were designed to measure receptive/productive vocabulary size and a limitation of time might have affected the participants' performance in the tests.

3. Responses on the receptive and productive tests were scored at two levels of sensitivity: sensitive and strict. First, all of the responses were scored by a Japanese native speaker. Responses were only marked as correct if the target words were spelled correctly. Next, another Japanese native speaker, the researcher, checked all of the incorrect or misspelled responses (see the scoring guideline in Chapter 3 Scoring, pp. 35-39). Both of the raters were native speakers of Japanese as the tests involved Japanese translation.
4. In the third week, reading comprehension test was administered. Once the test was distributed, all the participants were instructed to read three passages and the comprehension questions. They were instructed to select the best answer from four choices for each question within 60 minutes.
5. All of the responses on the reading comprehension tests were scored by the researcher. A correct answer received one point and an incorrect answer received zero points.
6. In the fourth week, the participants were assigned to choose one of four topics and write a well organized essay within a 60-minute time period. They were not allowed to use a dictionary because this study investigated the correlation between vocabulary knowledge and writing proficiency. Dictionary use might have affected their vocabulary knowledge

and writing performance.

7. Two raters evaluated each essay holistically according to Song and Caruso's (1996) assessment guidelines, the Evaluation Scale for Writing Skills Assessment Test (see Appendix I). The second rater, the researcher, evaluated all the essays without seeing the scores given by the first reader in order to give a reliable judgment without bias.

3.5. Data Analysis

The statistical analyses in response to each research question were based on One-way ANOVAs and Spearman Correlation Coefficient Test.

One-way ANOVA

A One-way ANOVA was performed to examine the following relationships. The analyses were performed separately for each frequency band and scoring.

1. Comparison of overall receptive and productive vocabulary size
 - Total receptive and productive vocabulary scores at each scoring (sensitive and strict)
2. Relationship between receptive and productive vocabulary size at different word frequencies

A One-way ANOVA was performed with receptive and productive vocabulary scores at

- each frequency band (701-1,900, 1,901-3,400, and 3,401-6,600 most frequent words)
 - each scoring (sensitive and strict)
3. Relationship between receptive and productive vocabulary size for learners with different receptive vocabulary scores

The participants were split into two groups according to their scores on the receptive translation test using both strict and sensitive scoring. The upper group consisted of 14 participants with the highest scores of both scoring methods, whereas the lower group was composed of 15 participants with the lowest scores of sensitive scoring and 14 participants with the lowest scores of strict scoring (see Figure 4) due to the distribution of test scores.

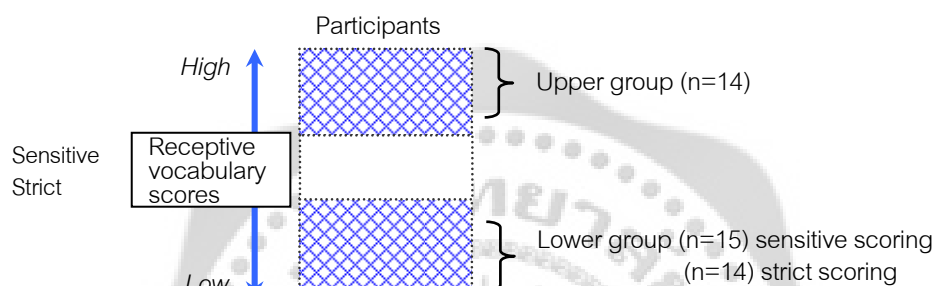


Figure 4 Upper and lower groups of the receptive translation test

A One-way ANOVA was performed with receptive and productive vocabulary scores

at

- each frequency band
- each group
- each scoring (sensitive and strict)

4. Relationship between loanwords and non-loanwords at different word frequencies

A One-way ANOVA was performed with receptive and productive vocabulary scores

of loanwords and non-loanwords at

- each frequency band
- each scoring (sensitive and strict)

Spearman Correlation Coefficient Test

Spearman Correlation Coefficient Test was performed to examine the correlation of the following values.

5. Correlation between overall receptive and productive vocabulary size
6. Correlation between receptive/productive vocabulary size and reading comprehension scores
7. Correlation between receptive/productive vocabulary size and free writing scores

In summary, participants, instruments, scoring methods, procedures and data analysis have been discussed in this chapter. Three instruments, the Receptive and Productive Translation Tests (Webb, 2008), EIKEN Grade 2 (Society for Testing English Proficiency, 2008) and the Evaluation Scale for the Writing Skills Assessment Test (Song & Caruso, 1996), were chosen to investigate the relationship between Japanese undergraduate students' receptive and productive vocabulary size at different frequency levels and to examine the correlation with their performance in reading and writing.

Chapter 4

Findings

The current study investigated the relationship between receptive and productive vocabulary size and the correlation of receptive/productive vocabulary size with reading comprehension and writing ability of 49 Japanese undergraduate students. This chapter presents quantitative analyses of data obtained from the three research instruments; the vocabulary tests, the reading comprehension test, and free writing.

The analysis of data gained from the Receptive and Productive Translation Tests (Webb, 2008) will be presented, followed by the analysis of data obtained from the reading comprehension test (EIKEN Grade 2, Society for Testing English Proficiency, 2008) and free essay.

4.1. Overall receptive and productive vocabulary size

Table 7 presents the means, standard deviations and ratios for receptive and productive vocabulary on the Receptive and Productive Translation Tests (Webb, 2008). The raw scores for receptive and productive vocabulary size are out of 90.

Table 7 Means, standard deviations and ratios for the total receptive and productive scores

Scoring	<u>Receptive</u>		<u>Productive</u>		**P/R ratio
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Sensitive	56.57*	14.65	42.16*	20.67	75%
Strict	56.33*	14.67	32.22*	20.39	57%

Note. Maximum score = 90, $n=49$, $*p < .001$

** Productive scores / receptive scores x 100

The results show that receptive scores were higher than productive scores using both the sensitive and strict scoring methods (see Table 7). The ratio of productive to receptive knowledge was much greater using sensitive scoring (75%) than strict scoring (57%).

One-way ANOVAs were performed to determine whether there were any differences between the receptive and productive vocabulary scores. The purpose of analysis of variance (ANOVA) is to test for significant differences between means. If variance components are found to be significant, the means for the receptive and productive vocabulary scores are different from each other. The independent variable was the type of knowledge: receptive and productive. One-way ANOVAs revealed significant differences between overall receptive and productive vocabulary scores. Receptive scores were found to be significantly higher than productive scores using sensitive scoring, $F(1, 48) = 79.810, p < .001$, and strict scoring, $F(1, 48) = 162.451, p < .001$ (see Table 7). The results indicate that receptive vocabulary size is larger than productive vocabulary size.

Participants had greater knowledge of meaning and form for the target words presented on Version B of both the receptive and productive tests. However, this would not affect the results, because the study investigated how receptive and productive vocabulary knowledge related to each other in the same individual using a within-subject design rather than a between-subjects design. The data acquired from the receptive test on Version A were compared with the scores for the productive test on Version A only. Similarly, the data obtained from the receptive test on Version B were solely compared with the scores for the productive test on Version B.

Additionally, Spearman correlation tests were performed to observe the correlation between receptive and productive vocabulary size. The results revealed that a strong and

significant correlation exists between receptive and productive scores as shown in Table 8. This indicates that the participants who had a higher receptive vocabulary size had a higher productive vocabulary size. The receptive scores correlated slightly higher with productive scores in sensitive scoring (.84) than in strict scoring (.81).

Table 8 Correlation between receptive and productive scores

Scoring	Productive	
	Sensitive	Strict
Receptive		
Sensitive	.84*	-
Strict	-	.81*

Note. * $p < .01$, $n=49$

4.2. Relationship between receptive and productive vocabulary size at different word frequencies

The second research question focused on the relationship between receptive and productive vocabulary size at different word frequencies. The scores on the receptive and productive translation tests at each word frequency band are presented in Table 9. The maximum score for receptive and productive vocabulary at each frequency band is 30.

Table 9 shows that receptive scores were higher than productive scores at all the frequency bands with both strict and sensitive scoring methods. One-way ANOVAs revealed that receptive scores were higher than productive scores when scoring for partial knowledge at word band 1, $F(1, 48) = 57.347$, $p < .001$, word band 2, $F(1, 48) = 33.014$, $p < .001$, and word band 3, $F(1, 48) = 47.025$, $p < .001$ (see Table 9). Receptive scores were also significantly higher at word band 1, $F(1, 48) = 132.630$, $p < .001$, word band 2, $F(1, 48) = 68.471$, $p < .001$, and word band 3, $F(1, 48) = 130.934$, $p < .001$, using strict scoring. In addition, Figure 5 illustrates the differences between receptive and productive size at each frequency band when scoring for partial knowledge (sensitive scoring) and fuller knowledge

(strict scoring). Furthermore, the ratio of productive to receptive scores decreased as the word frequency decreased using both scoring methods (see Figure 6).

Table 9 Means, standard deviations and ratios for receptive and productive scores at different word frequency bands

	<u>Word band 1</u>		<u>Word band 2</u>		<u>Word band 3</u>	
Scoring	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Sensitive						
Receptive	24.33*	4.65	19.35*	5.46	12.90*	5.79
Productive	19.35*	7.14	14.27*	8.21	8.55*	6.62
P/R ratio	80%		74%		66%	
Strict						
Receptive	24.24*	4.68	19.24*	5.48	12.84*	5.83
Productive	15.86*	7.57	11.00*	8.26	5.37*	5.70
P/R ratio	65%		57%		42%	

Note. Maximum score = 30, $n=49$, * $p < .001$

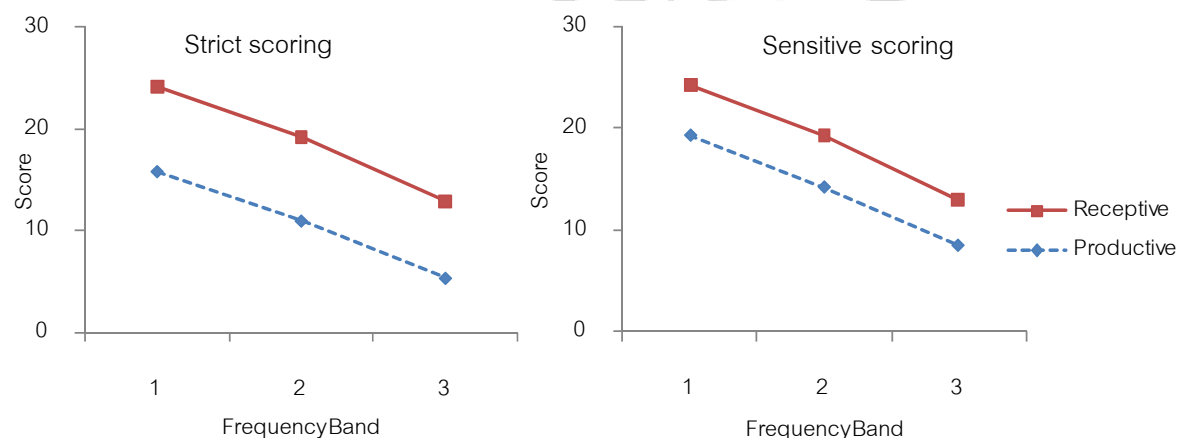


Figure 5 Receptive and productive scores at each frequency band

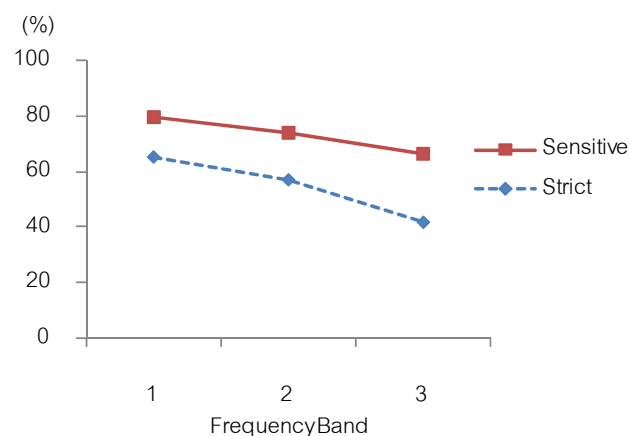


Figure 6 Ratio of productive to receptive scores at each frequency band

4.3. Relationship between receptive and productive vocabulary size for learners with different receptive vocabulary scores

The third research question focused on the relationship between receptive and productive vocabulary size for learners with different receptive vocabulary scores. Table 10 presents the data for the receptive and productive translation tests at each word frequency band for the participants with the highest and lowest scores when scoring for partial knowledge. Receptive scores were significantly higher for upper level participants at word band 1, $F(1, 13) = 10.515, p < .01$, and word band 3, $F(1, 13) = 10.957, p < .01$. There was no difference between scores at word band 2, $F(1, 13) = 3.222, p < .1$. Receptive scores were significantly higher for the lower level group at word band 1, $F(1, 13) = 28.213, p < .001$, word band 2, $F(1, 13) = 12.261, p < .005$, and word band 3, $F(1, 13) = 7.411, p < .025$.

Table 11 shows the scores on the receptive and productive tests at each word frequency band for the participants with the highest and lowest scores using strict scoring. One-way ANOVAs revealed significant differences for the upper level group at word band 1, $F(1, 13) = 17.182, p < .005$, word band 2, $F(1, 13) = 9.9, p < .05$, and word band 3, $F(1, 13) = 69.129, p < .001$. Receptive scores were significantly higher for the lower level group at word band 1, $F(1, 13) = 58.710, p < .001$, word band 2, $F(1, 13) = 14.541, p < .005$, and word band 3, $F(1, 13) = 9.522, p < .01$.

Receptive vocabulary size was larger than productive vocabulary size at all frequency levels in both upper and lower groups with both strict and sensitive scoring. Additionally, the ratios of productive to receptive knowledge as well as the scores decreased as word frequency decreased, regardless of the scoring methods. Productive-receptive ratios indicate that the gap between receptive and productive vocabulary size was wider at the lower level group.

Table 10 Means, standard deviations and ratios for the participants with the highest and lowest receptive scores at different word frequency bands using sensitive scoring

Scoring	Word band 1		Word band 2		Word band 3	
	M	SD	M	SD	M	SD
Upper group (n=14)						
Receptive	28.64*	1.45	24.93	2.20	19.21*	3.83
Productive	25.36*	3.79	22.14	6.06	15.86*	4.88
P/R ratio	89%		89%		83%	
Lower group (n=15)						
Receptive	19.20*	4.33	13.27*	3.73	7.27*	3.97
Productive	11.80*	6.52	7.07*	6.37	3.53*	4.79
P/R ratio	61%		53%		49%	

Note. Maximum score = 30, * $p < .01$

Due to the distribution of receptive scores, the number of participants in the lower group differs between scoring methods.

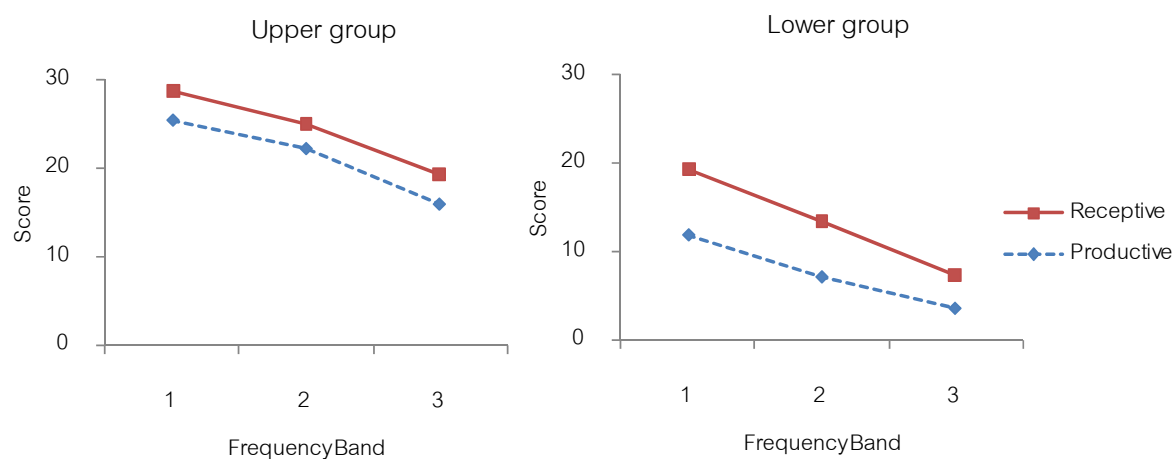


Figure 7 Receptive and productive scores for the participants with the highest and lowest receptive scores at different word frequency bands using sensitive scoring

Table 11 Means, standard deviations and ratios for the participants with the highest and lowest receptive scores at different word frequency bands using strict scoring

Scoring	Word band 1		Word band 2		Word band 3	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Upper group (<i>n</i> =14)						
Receptive	28.57*	1.50	24.79*	2.22	19.21*	3.83
Productive	22.43*	5.67	18.14*	8.24	10.64*	5.96
P/R ratio	79%		73%		55%	
Lower group (<i>n</i> =14)						
Receptive	18.71*	4.20	13.14*	3.76	6.93*	3.89
Productive	8.71*	6.28	5.43*	6.73	2.36*	4.88
P/R ratio	47%		41%		34%	

Note. Maximum score = 30, * $p < .05$

Due to the distribution of receptive scores, the number of participants in the lower group differs between scoring methods.

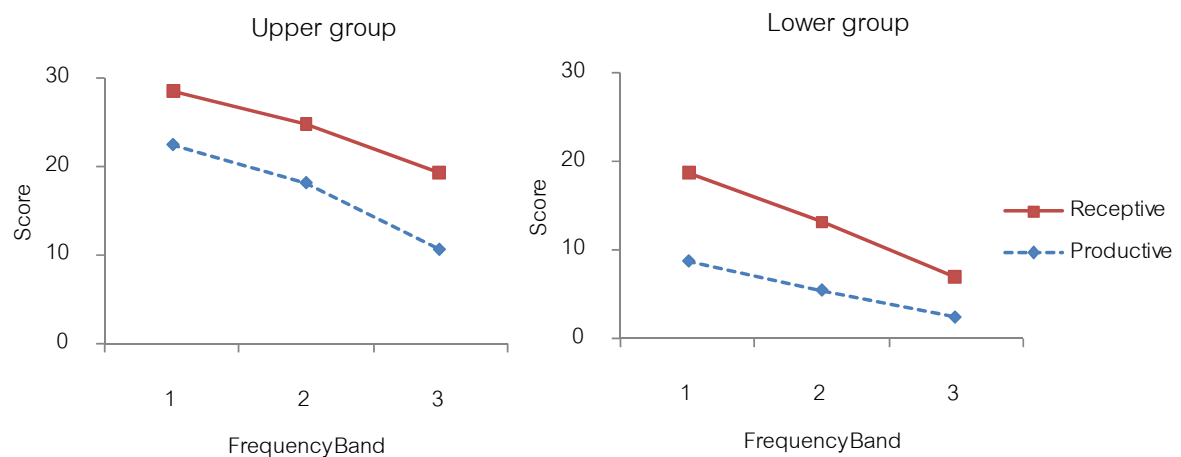


Figure 8 Receptive and productive scores for the participants with the highest and lowest receptive scores at different word frequency bands using strict scoring

4.4. Relationship between loanwords and non-loanwords at different word frequencies

The forth question aimed to determine if there was any difference in vocabulary size scores between loanwords and non-loanwords at different word frequencies.

Table 12 shows the number and percentage of loanwords in the target words (see Appendix C for loanwords on the test). Out of the 180 words on the Receptive and Productive Translation Tests, 48 (or 27%) were loanwords. All participants encountered all 48 loanwords either on the receptive or productive test, regardless of the version of the test because the target words used in the receptive test on Version A appeared as productive target words on Version B and the receptive target words on Version B were used as the target words on the productive test on Version A.

Table 12 Number and percentage of loanwords at each word frequency band

	n	%
Word band 1	10	17
Word band 2	19	32
Word band 3	19	32
Total	48	27

Note. Each word band contains 60 words.

Table 13 presents the percentage of correct responses for loanwords and non-loanwords. The participants scored higher for loanwords than non-loanwords on the receptive test with sensitive and strict scoring, while the differences between loanwords and non-loanwords were fairly small on the productive test. The percentage of correct responses for loanwords on the productive test was slightly lower than that of non-loanwords for fuller knowledge.

Table 13 Percentage of correct responses for loanwords and non-loanwords

Scoring	<u>Receptive</u>		<u>Productive</u>	
	Loanword (%)	Non-loanword (%)	Loanword (%)	Non-loanword (%)
Sensitive	74	59	51	45
Strict	74	58	34	36

Note. $n=49$

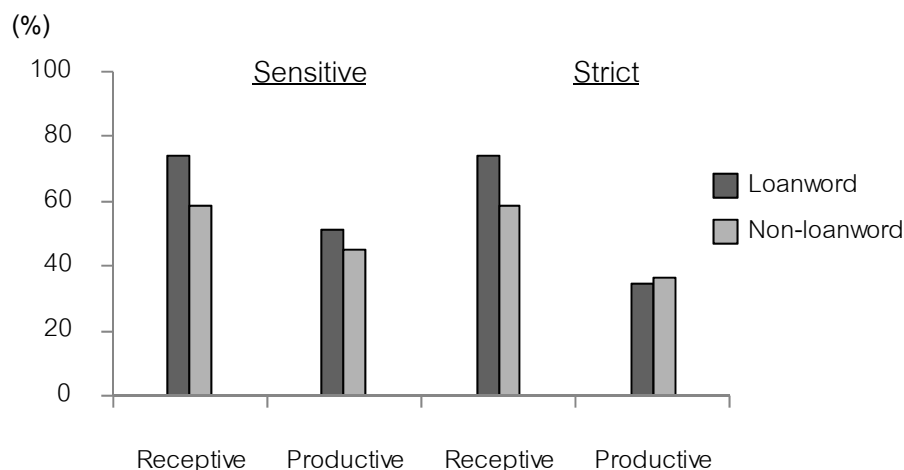


Figure 9 Percentage of correct responses for loanwords and non-loanwords

Table 14 shows the percentage of correct responses for loanwords and non-loanwords at each frequency band. On the receptive test, the participants scored higher for loanwords than non-loanwords at all the word bands with both scoring methods. The graph provides a visual representation that the participants recognized the loanwords and retrieved their meanings more successfully than on non-loanwords (see Figure 10 and 11). Additionally, the ratio of correct responses for both loanwords and non-loanwords decreased as the word frequency decreased, regardless of scoring methods.

On the productive test, the participants scored lower for loanwords than non-loanwords at word band 1, with both sensitive and strict scoring. It can be assumed that frequent loanwords did not always help the learners to produce the correct forms. As for less frequent words at word bands 2 and 3, the ratio of correct responses for loanwords was higher than non-loanwords for both partial and fuller knowledge.

Table 14 Percentage of correct responses for loanwords and non-loanwords at each word frequency band

	Word band 1		Word band 2		Word band 3	
	Loanword (%)	Non-Loanword (%)	Loanword (%)	Non-Loanword (%)	Loanword (%)	Non-Loanword (%)
Receptive						
Sensitive	93	79	78	58	60	35
Strict	93	78	78	58	60	35
Productive						
Sensitive	53	67	56	43	44	21
Strict	42	55	41	35	24	15

Note. $n=49$

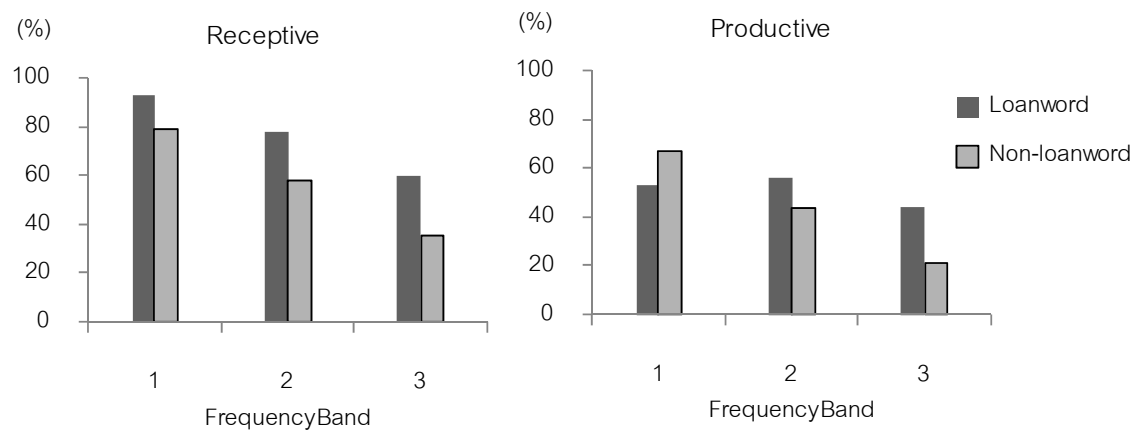


Figure 10 Percentage of correct responses for loanwords and non-loanwords at each frequency band using the sensitive scoring method

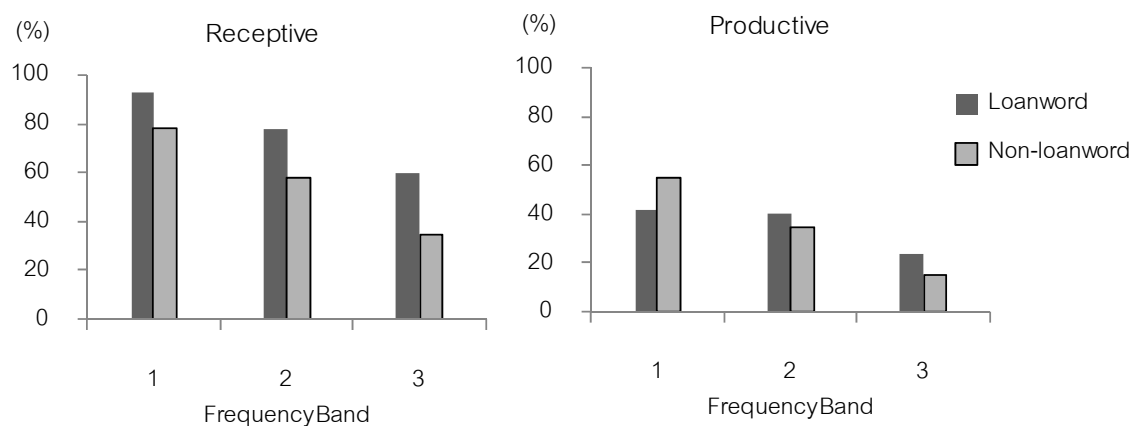


Figure 11 Percentage of correct responses for loanwords and non-loanwords at each frequency band using the strict scoring method

4.5. Correlation between vocabulary size and reading comprehension scores

The results of Spearman correlation tests revealed that both receptive and productive vocabulary size correlated moderately, but significantly, with reading comprehension scores. The correlation between receptive scores and reading comprehension scores was found to be .37 when scoring for partial knowledge and .38 with strict scoring. It is interesting to note that reading comprehension scores correlated slightly higher with productive scores than with receptive scores using both scoring methods (see Table 15).

Table 15 Correlation of receptive and productive vocabulary with reading comprehension

Vocabulary	Overall
Receptive	
Sensitive	.37*
Strict	.38*
Productive	
Sensitive	.41*
Strict	.41*

Note. * $p < .01$, $n=49$

Table 16 shows the correlation between scores on the receptive/productive tests and reading comprehension for the participants with the highest and lowest scores on the receptive vocabulary test. Spearman correlation tests revealed a significant and strong correlation of receptive and productive vocabulary size with reading for the upper level group. The correlation between the scores on the productive vocabulary test using strict scoring and reading comprehension found to be highest (.87) among others. The correlation of scores on the receptive test with reading was also found to be strong; .75 for fuller knowledge and of .73 for partial knowledge. Non-significant and negative correlations were found in the lower level group.

Table 16 Correlation of receptive and productive vocabulary with reading comprehension for the participants with the highest and lowest receptive scores

Vocabulary	Upper group (<i>n</i> =14)	Lower group (** <i>n</i> =15, *** <i>n</i> =14)
Receptive		
Sensitive	.73*	-.23
Strict	.75*	-.26
Productive		
Sensitive	.86*	-.18
Strict	.87*	-.14

Note. * $p < .01$, ** Sensitive scoring, *** Strict scoring

Due to the distribution of receptive scores, the number of participants in the lower group differs between scoring methods.

4.6. Correlation between vocabulary size and free writing scores

The results of Spearman correlation tests showed significant but weak correlation of receptive and productive scores with writing. The writing scores correlated slightly higher with receptive scores than with productive scores when using both scoring methods (see Table 17).

Table 17 Correlation of receptive and productive vocabulary with free writing

Vocabulary	Overall
Receptive	
Sensitive	.22*
Strict	.22*
Productive	
Sensitive	.11*
Strict	.12*

Note. * $p < .05$, *n*=49

Table 18 presents the correlation between receptive/productive scores and the free writing scores for the participants with the highest and lowest scores on the receptive test.

Spearman correlation tests revealed a significant and moderate correlation of

receptive and productive vocabulary size with writingt for the upper level group. The receptive scores correlated slightly higher with free writing than with productive scores using both types of scoring. A negligible correlation was found in the lower level group.

Table 18 Correlation of receptive and productive vocabulary with free writing for the participants with the highest and lowest receptive scores

Vocabulary	Upper group (n=14)	Lower group (**n=15, ***n=14)
Receptive		
Sensitive	.51*	.14
Strict	.52*	.01
Productive		
Sensitive	.47*	-.06
Strict	.43*	.03

Note. * $p < .05$ ** Sensitive scoring *** Strict scoring

Due to the distribution of receptive scores, the number of participants in the lower group differs between scoring methods.

Chapter 5

Conclusion and Discussion

In this chapter, the analyzed data will be discussed to respond to the research questions. In addition, general recommendations, limitations of the study as well as practical suggestions for further studies will be presented, respectively.

5.1. Objective of the Study

The purpose of the study was to investigate the relationship between Japanese undergraduate students' receptive and productive vocabulary size at different frequency levels and to examine the correlation with their performance in reading and writing.

5.2. Research Questions

Six research questions in the present study are as follows.

1. Is receptive vocabulary size larger than productive vocabulary size?
2. How does the relationship between receptive and productive vocabulary size change at different word frequency levels?
3. How does the relationship between receptive and productive vocabulary size change for learners with different receptive vocabulary size?
4. Is there any difference in vocabulary size between loanwords and non-loanwords?
5. How does receptive/productive vocabulary size correlate with reading comprehension?
6. How does receptive/productive vocabulary size correlate with writing ability?

5.3. Procedures

The study was conducted for three weeks in October, 2008. Forty-nine Japanese undergraduate students in Tokyo, Japan, were asked to participate in the present study.

The Receptive and Productive Translation Tests (Webb, 2008) were administered to investigate the relationship between receptive and productive vocabulary size at different frequency levels. A total of 180 target words (90 receptive and 90 productive) chosen from three different frequency bands were presented on three pages. The participants were allowed to work on the task as long as they needed, since the limitation of time might have affected the participants' performance on the tests. Responses on the tests were scored at two levels of sensitivity (sensitive and strict) to assess partial and fuller knowledge of written forms.

In the next week, the reading comprehension test, EIKEN Grade 2 (Society for Testing English Proficiency, 2008) was administered to examine the correlation of vocabulary size with reading proficiency. The participants were instructed to read three passages and 12 comprehension questions, from which they would select the best answer from among four choices for each question. They were allowed 60 minutes to complete the test.

Finally, the participants were assigned to write an essay in order to investigate the correlation between vocabulary size and writing ability. They chose one of four provided topics and wrote a well-organized essay within 60 minutes. Dictionary use was not permitted because it might have affected their writing performance. Two raters evaluated each essay holistically based on Song and Caruso's (1996) assessment guidelines: the Evaluation Scale for the Writing Skills Assessment Test.

5.4. Research Findings

The main findings of the study are summarized as follows.

1. Receptive vocabulary size was higher than productive vocabulary size using both the sensitive and strict scoring methods. The ratio of productive to receptive was much greater using sensitive scoring than strict scoring. The results of Spearman correlation tests revealed that the participants who had a higher receptive vocabulary size also had a higher productive vocabulary size.
2. Receptive vocabulary size was higher than productive vocabulary size at each frequency band. The ratio of productive to receptive vocabulary size decreased as the word frequency decreased with both scoring methods.
3. Receptive vocabulary size was larger than productive vocabulary size in both upper and lower groups of the receptive vocabulary test. The ratio of productive to receptive knowledge indicates that the gap between receptive and productive vocabulary size was wider at the lower level group.
4. The participants scored higher for loanwords than non-loanwords on the receptive test. However, the differences between loanwords and non-loanwords were fairly small on the productive test.
5. Spearman correlation tests revealed that both receptive and productive vocabulary size correlated moderately with reading comprehension. A strong correlation was found between receptive/productive vocabulary size and reading for the upper level group, while a non-significant correlation was found in the lower level group.
6. The results of Spearman correlation tests showed a weak correlation of

receptive and productive vocabulary size with writing ability. A moderate but significant correlation between receptive/productive vocabulary size and writing was found for the upper level group. No significant correlation was found for the lower level group.

5.5. Discussion of the Findings

The research findings will be discussed in detail in the following order; (1) the overall receptive and productive vocabulary size, (2) the relationship between receptive and productive vocabulary size at different word frequencies, (3) the relationship between receptive and productive vocabulary size for learners with different receptive vocabulary size, (4) the relationship between loanwords and non-loanwords at different word frequencies, (5) the correlation between vocabulary size and reading comprehension, and (6) the correlation between vocabulary size and writing ability.

5.5.1. Overall receptive and productive vocabulary size

In order to investigate the relationship between receptive and productive vocabulary knowledge, the equivalent test formats and two levels of scoring methods (sensitive and strict) were used for vocabulary size tests.

The results support findings in previous studies of Laufer (1998), Laufer and Paribakht (1998), Morgan and Oberdeck (1930) and Waring (1997), that a learner's receptive vocabulary knowledge is larger than that of his/her productive one. Remarkably, Spearman correlation tests found a significant and strong correlation between receptive and productive scores. It was discovered that the participants who had a higher receptive vocabulary size also had a higher productive vocabulary size.

When responses on the translation tests were checked with sensitive scoring, focusing on the meaning and partial knowledge of the target words, the ratio of productive to receptive knowledge was 75% (see Table 7). This indicates that three out of four words that received an appropriate response on the receptive test were also marked as correct on the productive test. When responses were checked with strict scoring, focusing on the meaning and the correct forms of the target words, the productive-receptive ratio was 57%. The gap between receptive and productive vocabulary size was larger when using strict scoring than sensitive scoring. The reason for this is that the participants made very few orthographic errors in their L1 on the receptive test while they made more errors in L2 spelling on the productive test.

In addition, the participants' receptive vocabulary size was 16 percentage points higher than their productive vocabulary size with sensitive scoring and 27 percentage points higher with strict scoring (see Table 19). The gap between receptive and productive knowledge was similar to Schmitt and Meara's (1997) findings that learner's receptive vocabulary size was 19 to 25 percentage points higher than their productive vocabulary size.

Table 19 Percentage of correct responses for total receptive and productive scores

Scoring	Receptive	Productive
Sensitive	63%	47%
Strict	63%	36%

n=49

The data presented here, as well as in previous findings, give a clear indication that receptive vocabulary knowledge is larger than productive vocabulary knowledge. A reasonable explanation for this could be that words are easier to access receptively than to

use productively (Waring, 1997). It is more demanding for a learner to use a word than to know a word. For reception, a learner needs to recognize a word in its written and spoken forms and to recall the meaning when met. For production, a lot more is required; such as knowledge of the spelling and pronunciation as well as the meaning, an ability to use it correctly with collocation, an ability to know how to mesh the word with grammar rules and the sociolinguistic aspects of language production (Waring, 1997) and so on. Production of language places a heavier demand on learners than reception of language. This would explain why a person has more receptive knowledge than productive knowledge.

5.5.2. Relationship between receptive and productive vocabulary size at different word frequencies

The second question examined the relationship between receptive and productive vocabulary size at different frequency levels. The results showed that both receptive and productive scores decreased as word frequency decreased, with both strict and sensitive scoring. The gap between productive and receptive vocabulary size widened as frequency decreased, regardless of the scoring methods.

When responses were checked with strict scoring, the ratio of productive to receptive scores was 65% at word band 1, 57% at word band 2, and 42% at word band 3 (see Table 9). When responses were scored for partial knowledge, the gap between receptive and productive scores was smaller at each frequency band. The receptive-productive ratio was 80% at word band 1, 74% at word band 2, and 66% at word band 3.

As discussed in the previous question, there is a heavier demand on a learner for production than reception even though only the forms and meaning of the target words were tested in the present study. As shown in Figure 6, the upper slope presenting partial

knowledge flattens, whereas the lower slope for fuller knowledge gently falls to the right as frequency decreased. A possible explanation is that a learner can acquire partial, incomplete, or limited knowledge about L2 forms to a certain extent regardless of word frequency. However, much more effort is needed to acquire fuller knowledge, especially at the lower frequency levels.

5.5.3. Relationship between receptive and productive vocabulary size for learners with different receptive vocabulary size

The third research question investigated the relationship between receptive and productive vocabulary size for learners with different receptive vocabulary scores. To determine whether or not the relationship between receptive and productive vocabulary size changed for learners, the 14 participants with the highest receptive vocabulary scores and the 14 participants with the lowest receptive vocabulary scores were divided into two levels: upper and lower. The lower group with sensitive scoring was comprised of 15 participants due to the distribution of receptive vocabulary scores.

The results showed that scores on the tests at each frequency band decreased as word frequency decreased with both scoring methods (see Table 10 and 11). When responses were scored for partial knowledge, the difference between receptive and productive vocabulary size was very small at all frequency bands for the upper level group, with the productive-receptive ratio ranging from 83% to 89%, whereas the gap was wider for the lower level group, with a ratio ranging from 49% to 61%. When responses were scored for fuller knowledge, the results showed that the gap between receptive and productive vocabulary size widened at all frequency bands for both groups. The upper level group had a productive-receptive ratio of 79% for word band 1, 73% for word band 2, and 55% for word

band 3 (see Table 11). The lower level group had a ratio of 47% for word band 1, 41% for word band 2, and 34% for word band 3.

The findings indicate that receptive vocabulary size might give some indication of productive vocabulary size. Learners who have a larger receptive vocabulary are likely to know more of those words productively than learners who have a smaller receptive vocabulary. Participants with higher receptive vocabulary scores had greater ability at demonstrating productive knowledge of L2 forms than those with lower receptive scores.

The ratio of productive to receptive vocabulary size was higher in the upper group, whereas Laufer (1998) found that the productive-receptive ratio was lower in more advanced groups than in less advanced groups. She claims that receptive vocabulary of beginners consists of the most frequent and useful words. The small gap between receptive and productive vocabulary size in less advanced groups resulted from the limited size of receptive vocabulary as well as the productive one. As the receptive vocabulary size expands, it is difficult to activate all the receptive knowledge of a learner to productive use, therefore, more words remain in the receptive realm. As a result, the gap between receptive and productive vocabulary size widens.

However, the participants in the lower level group in the current study demonstrated a wider gap than those in the upper group. This may be caused by the degree at which the participants had been exposed to English both inside and outside the classroom. All the participants in Laufer's (1998) study were native speakers of Hebrew and had been learning English at the advanced curriculum in the Israel school system for either six or seven years. In addition, they were reported to have been exposed to the target language by watching English-speaking TV programs, playing computer games, and communicating with

English-speaking tourists outside the classroom. The participants in the present study, on the other hand, were native speakers of Japanese and had learned English at least six years before entering the university, but some had not taken an English class for two or three years after graduating from high school since English was not a required subject for their degree. Unlike the participants in Laufer's study, most participants are less likely to have exposure to English. It is necessary to increase opportunities to reinforce productive knowledge both in side and outside the classroom. Productive learning is likely to lead to larger gains in increasing productive knowledge, just as receptive learning is more effective in increasing receptive knowledge than productive knowledge (Waring, 1997). A more balanced approach to vocabulary learning is a key for growth in both receptive and productive knowledge.

5.5.4. Relationship between loanword and non-loanword items at different word frequencies

The next question examined the relationship between loanwords and non-loanwords at different frequency levels.

On the receptive test, the participants scored higher for loanwords than non-loanwords regardless of scoring methods and frequency bands (see Table 14). The learners successfully demonstrated receptive knowledge of non-loanwords as well as loanwords in word band 1 (see Figure 10 and 11). In word band 3, however, the percentage of correct answers for non-loanword declined rapidly (35%) while it remained relatively high for loanwords (60%) with both scoring methods. The results indicate that the learners recognized the L2 form of loanwords and retrieved their meaning more successfully than with non-loanwords.

On the productive test, on the other hand, the percentage of correct answers for

loanwords was slightly lower than that of non-loanwords for fuller knowledge, but the results were reversed for partial knowledge (see Table 14). The results show that loanwords did not necessarily help the learners produce the correct L2 forms, but at least helped them demonstrate partial productive knowledge. At the most frequent word level, the participants scored lower for loanwords than non-loanwords in both scoring methods (see Figures 10 and 11). A possible explanation is that there were fewer loanwords involved in word band 1 than in other bands (see Table 12).

The scores for loanwords on the productive tests were much lower than expected, especially for fuller knowledge. If L1 forms of loanwords could provide information about their L2 forms, it might be easy to demonstrate productive knowledge of form. However, none of the participants were able to demonstrate productive knowledge of *calendar*, *trumpet*, *cereal*, and *handsome* in strict scoring. It can be assumed that L1 forms of loanwords symbolized in accordance with Japanese orthography could provide a small piece of information about L2 forms.

5.5.5. Correlation between vocabulary size and reading comprehension

The next question investigated the correlation of vocabulary size with reading comprehension. It was revealed that the relationship between vocabulary size and reading comprehension scores was different from correlation results found in previous studies.

The correlation of receptive vocabulary size with reading was found to be .37 when scoring for partial knowledge and .38 in strict scoring (see Table 15). They were much lower than the findings of Qian (1999) or Beglar and Hunt (1999) as shown in Table 20. This may be due to differences of instruments used in the present study. There have been a number of

studies using the Vocabulary Levels Test (Nation, 1983) and the Reading Comprehension section from the Test of English as a Foreign Language (TOEFL) to examine the correlation between vocabulary size and reading proficiency. The current study used the Receptive and Productive Translation Tests (Webb 2008) and EIKEN (Test in Practical English Proficiency) specifically designed to measure the vocabulary size and language proficiency of Japanese learners.

Surprisingly, the results of the participants with the highest scores on the receptive vocabulary test were similar to the previous findings. The correlation of receptive vocabulary with reading was found to be .75 for fuller knowledge and .73 for partial knowledge (see Tables 16 and 20). The larger receptive vocabulary size led to better comprehension of the texts. The correlation between productive vocabulary and reading was even higher; .87 for fuller knowledge and .86 for partial knowledge. Productive vocabulary size could be an indicator of successful reading comprehension as well as receptive vocabulary size.

However, no significant correlation was found in the lower level group. It is possible that the participants in the lower group were incapable of comprehending the passages and the questions on the test. In addition, participants had a 25% chance of guessing correct answers from among four choices without reading any passages or questions as they had been instructed. The percentage of correct answers in the lower group was 23% while it increased to 47% in the upper level group. Guessing the answers for the multiple-choice questions might have resulted in a non-significant correlation in the lower group.

Table 20 Comparison of correlations between vocabulary size and reading comprehension

	Qian (1999)	Beglar & Hunt (1999)	Current study
Reading comprehension test		TOEFL	EIKEN Grade 2
Vocabulary size test	Vocabulary Levels Test (Nation, 1983)		Receptive and Productive Translation Tests (Webb, 2008)
	.78	2,000 word level	<Upper level group>
		Version A	Receptive Sensitive
		Version B	Strict
		University Word List	Productive Sensitive
		Version A	Strict
		Version B	

5.5.6. Correlation between vocabulary size and writing ability

The final question investigated the correlation of vocabulary size with writing ability. The results of the correlation analysis between vocabulary size and writing ability were relatively different from expected. The correlation of receptive vocabulary size with writing was found to be .22 with both scoring methods (see Table 17). The correlation with productive vocabulary size was lower; .11 when scoring for partial knowledge and .12 for fuller knowledge. They were much lower than the correlation found in Laufer and Paribakht's (1998) study, but higher than in Laufer's (1998) findings, in which no significant correlation was found (see Table 21). The weak correlation between vocabulary size and scores on the free essay revealed that larger receptive and productive vocabulary size do not seem to be reflected in free writing ability.

As for the participants with the highest scores on the receptive vocabulary test, the correlation of receptive vocabulary with writing was found to be .52 for fuller knowledge

and .51 for partial knowledge (see Tables 18 and 21). The correlation between productive vocabulary and writing was slightly lower; .43 for fuller knowledge and .47 for partial knowledge. However, free essay scores did not correlate with receptive and productive knowledge in the lower level group. This indicates that learners who could recognize or produce more words than other learners, if forced to, were not necessarily those who would express their thoughts and organize their ideas more clearly than others in free expression.

Table 21 Comparison of correlation between vocabulary size and writing ability

	Laufer (1998)	Laufer & Paribakht (1998)	Current study
Writing	Free composition	Free composition	Free composition
Vocabulary size test	Vocabulary Levels Test (Nation, 1983) Productive Vocabulary Levels Test (Laufer & Nation, 1999)		Receptive and Productive Translation Tests (Webb, 2008)
			<Upper level group>
	no significant correlation	Receptive .68	Receptive Sensitive .51 Strict .52
		Productive .68	Productive Sensitive .47 Strict .43

To sum up, the results of this study demonstrate that receptive vocabulary size was larger than productive vocabulary size regardless of the word frequency or vocabulary size of an individual learner. The participants who had a higher receptive vocabulary size also had a higher productive vocabulary size. In addition, the participants scored higher for loanwords than non-loanwords on the receptive test, but the differences between loanwords and non-loanwords were fairly small on the productive test. Furthermore, both receptive and

productive vocabulary size correlated with reading comprehension and writing ability for the upper level group based on the receptive vocabulary test, while no significant correlation was found for the lower level group.

5.6. General Recommendations

Several recommendations for ESL/EFL learners and teachers arise from the findings of this study.

1. The Receptive and Productive Translation Tests (Webb, 2008) are practical instruments to measure the vocabulary size of Japanese learners. In addition, an identical format to test receptive and productive knowledge is beneficial to investigate the relationship between receptive and productive vocabulary size and knowledge. The results of this study show a substantial gap between receptive and productive as seen in the other studies (Laufer, 1998; Laufer & Nation, 1999; Laufer & Paribakht, 1998; Schmitt, 1997; Waring, 1997; Webb, 2008). Teachers are encouraged to keep a balance between receptive and productive vocabulary learning.
2. It is difficult to bring receptive vocabulary into productive use (Laufer, 1998; Laufer & Paribakht, 1998, Nation, 2001). Learners need not only acquire productive knowledge, but must also have a desire and an opportunity to use a word. It is highly recommended that teachers conduct various activities to move vocabulary into productive use. Gap-fills, questionnaires, and games are good examples of production tasks (for more classroom activities, see Thornbury, 2002). When choosing tasks, teachers need to consider characteristics of learners and their

overall learning context, including their proficiency level, the L1, culture, their motivation and purposes for learning the L2.

3. The high frequency words, especially the 2,000 most frequent words, deserve individual attention since learners are likely to have a large number of opportunities to meet and use those words (Laufer & Nation, 1999). Beyond this level, teachers should focus more on strategies for vocabulary learning, such as guessing from the context and learning word parts, instead of teaching particular words directly. Learners should continue to increase their vocabulary size while learning various strategies and techniques to cope with low-frequency words (see more vocabulary teaching and learning strategies in Moir & Nation, 2002; Nation, 2001; Nation, 2002; Schmitt, 2007).
4. English spelling patterns are often predictable based on morphemic, graphemic and phonemic features, and there are few words that are completely unpredictable in spelling (Nation, 2001). Teaching various spelling rules and an application of previously learned knowledge can be helpful to move learners' partial knowledge about form to fuller knowledge.
5. A variety of vocabulary measures is useful to see if particular aspects of vocabulary knowledge are being neglected in a course. For example, learners who receive a high score in a vocabulary test but use poor vocabulary in free writing may need well designed tasks so that they are encouraged to use the full richness of their vocabulary in writing.
6. It has been suggested in the literature that vocabulary size tests may be a useful instrument to examine language proficiency (Laufer et al., 2004). The correlation

obtained from this study supports this claim. However, it is not recommended that a vocabulary test should replace other proficiency tests for the following two reasons. First, the correlation scores with reading and writing in the present study as well as other studies are not high enough. It is also difficult to determine what correlation score would be high enough for a vocabulary size test to serve as a reliable indicator of learners' reading and writing proficiency. Second, vocabulary size and knowledge is only one aspect of learners' success in language learning.

7. The equivalent test formats used in this study were beneficial to provide a more accurate measurement of receptive and productive vocabulary size. However, the following two points should be carefully considered when developing translation tests. Firstly, the target words on translation tests should be carefully chosen and should be pilot-tested to ensure that the participants are likely to respond to the target meanings. Secondly, grading tests might present practical difficulties in an ESL situation where there is a variety of learners' L1.

5.7. Limitations of the Study

Three limitations of the current study were identified as follows.

1. This study focused on a small sample. Therefore, the results may not be generalized to other groups of learners. Although the results of the present study might confirm that learners' receptive vocabulary size is larger than their productive vocabulary size, it is likely that this relationship varies from group to group. For example, the gap between receptive and productive vocabulary size was found to be smaller for EFL students than ESL students due to learning differences (Laufer & Paribakht, 1998).

2. The current study was intended to investigate the breadth of vocabulary knowledge – how many words learners know, but is not concerned with depth – how well those words are known. The tests used here only measured knowledge of meaning and form. The gap between receptive and productive knowledge could be greater if other aspects of word knowledge, such as collocation or syntax, were tested.
3. This study focused on receptive and productive vocabulary size and correlation of vocabulary size with reading and writing proficiency. However, reading requires not only recognition of written words, but also such skills as skimming, scanning, and inferring. Similarly, writing requires skills such as prewriting, producing a clear statement and topic sentences, using transitions, and control of punctuation. Clearly, the Receptive and Productive Translation Tests (Webb, 2008) is not designed to evaluate such skills. A different correlation of vocabulary size with reading and writing proficiency may be found when tested with other instruments.

5.8. Suggestions for Further Research

The proficiency level of learners and vocabulary instruction are contributory factors to vocabulary size. It would be more beneficial for teachers to examine how different methods of instruction affect vocabulary size in addition to further investigation of the relationship between receptive and productive vocabulary for different types of learners. Furthermore, the vocabulary size tests may need to be developed to serve as reliable instruments for more accurate comparisons of receptive and productive vocabulary.



References

- American Council for the Teaching of Foreign Languages. (2001). *ACTFL Proficiency Guidelines*. Retrieved February 11, 2009, from URL: <http://www.actfl.org/files/public/writingguidelines.pdf>
- Alderson, J. C. (1984). Reading in a foreign language: A reading or a language problem? In Alderson, J. C. & Urquhart, A. H. (Eds.), *Reading in a Foreign Language*, pp. 1-24. London: Longman.
- Astika, G. G. (1993). Analytical assessment of foreign students' writing. *RELC Journal*, 24, 61-72.
- Bacha, N. (2001). Writing evaluation: What can analytic versus holistic essay scoring tell us? *System*, 29, 371-383.
- Barkaoui, K. (2007). Rating scale impact on EFL essay marking: A mixed-method study. *Assessing Writing*, 12, 86-107.
- Beglar, D. & Hunt, A. (1999). Revising and validating the 2,000 word level and university word level vocabulary tests. *Language Testing*, 16, 131-162.
- Clark, I. (2003). *Concepts in Composition: Theory and Practice in the Teaching of Writing*. Mahwah, NJ: Lawrence Erlbaum.
- Daulton, F.E. (1998). Japanese loan word cognates and the acquisition of English vocabulary, *The Language Teacher Online*, 22 (1), Retrieved February 11, 2009, from URL: <http://www.jalt-publications.org/tlt/articles/1998/01/daulton>
- Goulden, R., Nation, I.S.P. & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics*, 11, 341-363.
- Hazenberg, S. & Hulstun, J. H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation, *Applied Linguistics*, 17, 145-163.
- Hu, H. M. & Nation, I.S.P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13, 403-430.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University Press.

- Jacobs, H., Zinkgraf, S., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL Composition: A Practical Approach*. Rowley, MA: Newbury House.
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In Bejoint, H. & Arnaud, P. (Eds.), *Vocabulary and Applied Linguistics*, pp. 126-32. London: Macmillan.
- (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In Coady, J. & Huckin, T. (Eds.), *Second Language Vocabulary Acquisition*, pp. 20-34. Cambridge: Cambridge University Press.
- (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics*, 19, 255-271.
- Laufer, B., Elder, C., Hill, K., & Congdon, P. (2004). Size and strength: Do we need both to measure vocabulary knowledge? *Language Testing*, 21, 202-226.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16, 307-322.
- (1999). A productive-size test of controlled productive ability. *Language Testing*, 16, 33-51.
- Laufer, B., & Paribakht, T. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning*, 48, 365-391.
- Meara, P. (1996). The dimensions of lexical competence. In Brown, G., Malmkjaer, K. & Williams, J. (Eds.), *Performance and Competence in Second Language Acquisition*, pp. 35-53. Cambridge: Cambridge University Press.
- Meara, P. & Jones, G. (1990). *The Eurocentres Vocabulary Size Tests 10KA*. Zurich: Eurocentres.
- Melka, F. (1997). Receptive versus productive aspects of vocabulary. In Schmitt, N. & McCarthy, M. (Eds.), *Vocabulary: Description, Acquisition, and Pedagogy*, pp. 84-102. New York: Cambridge University Press.
- Moir, J. & Nation, I.S.P. (2002). Learners' use of strategies for effective vocabulary learning. *Prospect*, 17 (1): 15-35.
- Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines*, 5, 12-25.
- (1990). *Teaching and Learning Vocabulary*. Rowley, MA: Newbury House.

- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nation, I.S.P. & Meara, P. (2002). Vocabulary. In Schmitt, N. (Ed.), *An Introduction to Applied Linguistics*, pp. 35-54. London: Edward Arnold.
- Nation, I.S.P. & Waring, R. (1997). Vocabulary size, text coverage and word lists. In Schmitt, N. & McCarthy, M. (Eds.), *Vocabulary: Description, Acquisition, and Pedagogy*, pp. 6–19. New York: Cambridge University Press.
- National Capital Language Resource Center. (2004). *Assessing learning: Alternative assessment*. Retrieved February 11, 2009, from URL: <http://www.nclrc.org/essentials/assessing/alternative.htm>
- Nonaka, T. (2004). English vocabulary size of Japanese university students, *Niigata Seiryo Junior College Research Report*, 34, 25-34.
- Paribakht, T.S. & Wesche, M.B. (1993). Reading comprehension and second language development in comprehension-based ESL programme. *TESL Canada Journal*, 11, 9-27.
- Qian, D. D. (1999). Assessing the roles of depth and breath of vocabulary knowledge in reading comprehension. *The Canadian Modern Language Review*, 56, 283-307.
- Read, J. (2000). *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Sakyi, A.A. (2000). Validation of holistic scoring for ESL writing assessment: How raters evaluate compositions. In Milanovic, M (Ed.), *Studies in Language Testing 9: Fairness and Validation in Language Assessment*, pp. 129-152. Cambridge: Cambridge University Press.
- Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, 48, 281-317.
- (1999). The relationship between TOEFL vocabulary items and meaning, association, collocation and word-class knowledge. *Language Testing*, 16, 189-216.

- Schmitt, N. (2007). Current trends in vocabulary learning and teaching. Cummins, J. & Davison, C. (Eds.), *The International Handbook of English Language Teaching, Vol. 2*. Springer. Retrieved February 11, 2009, from URL: https://www.nottingham.ac.uk/english/research/cral/lib/exe/fetch.php?id=people%3Aschmitt&cache=cache&media=people:schmitt:cummins_book_-_schmitt.doc
- Schmitt, N. & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition*, 19, 17-35.
- Society for Testing English Proficiency. (2008). *EIKEN Test in Practical English Proficiency*. Retrieved February 11, 2009, from URL: <http://stepeiken.org/>
- Song, C. B., & Caruso, I. (1996). Do English and ESL faculty differ in evaluating the essays of native English-speaking and ESL students? *Journal of Second Language Writing*, 5, 163-182.
- Sutarsyah, C., Nation, P. & Kennedy, G. (1994). How useful is EAP vocabulary for ESP? A corpus based case study. *RELC Journal*, 25, 34-50.
- Tadamitsu, K. (2001). *An examination of Nation's (1990) vocabulary levels test*. Retrieved February 11, 2009, from URL: <http://www1.harenet.ne.jp/~waring/vocab/colloquium/tad2001.htm>
- Thornbury, S. (2002). *How to Teach Vocabulary*. London: Longman
- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1-L2 acquisition and frequency of input. *Applied Psycholinguistics*, 22, 217-234.
- Waring, R. (1997). A comparison of the receptive and productive vocabulary sizes of some second language learners. *Immaculata (Notre Dame Seishin University, Okayama)*, 1, 53-68.
- Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, 30, 79-95.
- Xue, G. & Nation, I.S.P. (1984). A university word list. *Language Learning and Communication*, 3, 215-229.





APPENDIX A

Target Words

Word band 1 701st - 1,900th most frequent words		Word band 2 1,901st - 3,400th most frequent words		Word band 3 3,401st - 6,600th most frequent words	
island	島	medicine	薬	fog	霧
animal	動物	pocket	ポケット	glue	糊 (のり)
sport	スポーツ	palace	宮殿	socks	靴下
address	住所	hero	英雄	bruise	青あざ
earth	地球	beer	ビール	gym	ジム
airport	空港	winter	冬	hobby	趣味
doctor	医者	ear	耳	swear	誓う
cheap	安い	kiss	キス	cheek	頬 (ほお)
Monday	月曜日	hate	嫌う	owl	ふくろう
weak	弱い	sky	空	hammer	ハンマー
safe	安全	button	ボタン	dive	潜る
finger	指	belt	ベルト	funeral	葬式
wine	ワイン	hide	隠れる	bubble	泡
farm	農場	singer	歌手	thunder	雷
deep	深い	roof	屋根	spoon	スプーン
gun	銃	cow	牛	hug	抱き合う
teach	教える	wash	洗う	olive	オリーブ
scientist	科学者	sand	砂	tent	テント
rain	雨	vegetable	野菜	linen	麻
sing	歌う	sweet	甘い	ladder	はしご
soft	柔らかい	grass	草	mathematics	数学
chair	椅子	pilot	パイロット	ceiling	天井
newspaper	新聞	snow	雪	gasoline	ガソリン
queen	女王	bone	骨	ripe	熟した
secret	秘密	moon	月	pillow	枕
cold	寒い	relative	親戚	disco	ディスコ
machine	機械	salt	塩	recipe	レシピ
ticket	切符	leaf	葉	pork	豚肉
magazine	雑誌	pink	ピンク	translate	訳す
clothes	服	protein	たんぱく質	butterfly	蝶
Thursday	木曜日	pregnant	妊娠	stranger	見知らぬ人
bird	鳥	ambassador	大使	calendar	カレンダー
sleep	眠る	pollution	公害	stairs	階段
hole	穴	joke	冗談	carpet	じゅうたん
bridge	橋	jockey	騎手	collar	襟 (えり)
fruit	果物	fence	フェンス	cereal	シリアル
bottle	瓶	mask	お面	hungry	空腹な
dream	夢	ceremony	式	kidnap	誘拐
key	鍵	castle	城	spice	香辛料
warm	暖かい	kilometer	キロメートル	jungle	ジャングル
hotel	ホテル	bowl	ボウル	strawberry	イチゴ
burn	燃える	milk	牛乳	tomato	トマト
population	人口	boil	ゆでる	sour	酸っぱい
count	数える	birthday	誕生日	rainbow	虹
science	科学	mirror	鏡	wig	かつら
promise	約束	stomach	胃	handsome	ハンサム
weather	天気	boots	ブーツ	shallow	浅い
spring	春	yellow	黄色	umbrella	傘
camera	カメラ	license	免許	vertical	垂直な
box	箱	spray	スプレー	nest	巣
February	二月	jewel	宝石	trumpet	トランペット
wear	着る	actor	俳優	lily	ユリ
wind	風	toy	おもちゃ	lazy	怠惰な
corner	角 (かど)	tower	塔	zoo	動物園
weekend	週末	twins	双子	slave	奴隷
kitchen	台所	steal	盗む	massage	マッサージ
video	ビデオ	whisper	ささやく	miracle	奇跡
king	王	neighbor	隣人	candle	ろうそく
tomorrow	明日	jazz	ジャズ	lion	ライオン
train	電車	drum	ドラム	liquid	液体



APPENDIX B

Receptive and Productive Translation Test

<Version A>

Write a meaning of each word in Japanese.

Translate following words into English.

- | | | | |
|--------------|-------|--------------------------|-------|
| 1 island | _____ | (1) 木曜日 | _____ |
| 2 animal | _____ | (2) 鳥 | _____ |
| 3 sport | _____ | (3) 眠る | _____ |
| 4 address | _____ | (4) 穴 | _____ |
| 5 earth | _____ | (5) 橋 | _____ |
| 6 airport | _____ | (6) 果物 | _____ |
| 7 doctor | _____ | (7) 瓶 | _____ |
| 8 cheap | _____ | (8) 夢 | _____ |
| 9 Monday | _____ | (9) 鍵 | _____ |
| 10 weak | _____ | (10) 暖かい | _____ |
| 11 safe | _____ | (11) ホテル | _____ |
| 12 finger | _____ | (12) 燃える | _____ |
| 13 wine | _____ | (13) 人口 | _____ |
| 14 farm | _____ | (14) 数える | _____ |
| 15 deep | _____ | (15) 科学 | _____ |
| 16 gun | _____ | (16) 約束 | _____ |
| 17 teach | _____ | (17) 天気 | _____ |
| 18 scientist | _____ | (18) 春 | _____ |
| 19 rain | _____ | (19) カメラ | _____ |
| 20 sing | _____ | (20) 箱 | _____ |
| 21 soft | _____ | (21) 二月 | _____ |
| 22 chair | _____ | (22) 着る | _____ |
| 23 newspaper | _____ | (23) 風 | _____ |
| 24 queen | _____ | (24) <small>かど</small> 角 | _____ |
| 25 secret | _____ | (25) 週末 | _____ |
| 26 cold | _____ | (26) 台所 | _____ |
| 27 machine | _____ | (27) ビデオ | _____ |
| 28 ticket | _____ | (28) 王 | _____ |
| 29 magazine | _____ | (29) 明日 | _____ |
| 30 clothes | _____ | (30) 電車 | _____ |

31	medicine	(31)	妊娠
32	pocket	(32)	大使
33	palace	(33)	公害
34	hero	(34)	冗談
35	beer	(35)	騎手
36	winter	(36)	フェンス
37	ear	(37)	お面
38	kiss	(38)	式
39	hate	(39)	城
40	sky	(40)	キロメートル
41	button	(41)	ボウル
42	belt	(42)	牛乳
43	hide	(43)	ゆでる
44	singer	(44)	誕生日
45	roof	(45)	鏡
46	cow	(46)	胃
47	wash	(47)	ブーツ
48	sand	(48)	黄色
49	vegetable	(49)	免許
50	sweet	(50)	スプレー
51	grass	(51)	宝石
52	pilot	(52)	俳優
53	snow	(53)	おもちゃ
54	bone	(54)	塔
55	moon	(55)	双子
56	relative	(56)	盗む
57	salt	(57)	ささやく
58	leaf	(58)	隣人
59	pink	(59)	ジャズ
60	protein	(60)	ドラム

61	fog	(61)	見知らぬ人
62	glue	(62)	カレンダー
63	socks	(63)	階段
64	bruise	(64)	じゅうたん
65	gym	(65)	^{えり} 襟
66	hobby	(66)	シリアル
67	swear	(67)	空腹な
68	cheek	(68)	誘拐
69	owl	(69)	香辛料
70	hammer	(70)	ジャングル
71	dive	(71)	イチゴ
72	funeral	(72)	トマト
73	bubble	(73)	酸っぱい
74	thunder	(74)	虹
75	spoon	(75)	かつら
76	hug	(76)	ハンサム
77	olive	(77)	浅い
78	tent	(78)	傘
79	linen	(79)	垂直な
80	ladder	(80)	巢
81	mathematics	(81)	トランペット
82	ceiling	(82)	ユリ
83	gasoline	(83)	怠惰な
84	ripe	(84)	動物園
85	pillow	(85)	奴隷
86	disco	(86)	マッサージ
87	recipe	(87)	奇跡
88	pork	(88)	ろうそく
89	translate	(89)	ライオン
90	butterfly	(90)	液体

<Version B>

Write a meaning of each word in Japanese.

Translate following words into English.

1 Thursday

2 bird

3 sleep

4 hole

5 bridge

6 fruit

7 bottle

8 dream

9 key

10 warm

11 hotel

12 burn

13 population

14 count

15 science

16 promise

17 weather

18 spring

19 camera

20 box

21 February

22 wear

23 wind

24 corner

25 weekend

26 kitchen

27 video

28 king

29 tomorrow

30 train

(1) 島

(2) 動物

(3) スポーツ

(4) 住所

(5) 地球

(6) 空港

(7) 医者

(8) 安い

(9) 月曜日

(10) 弱い

(11) 安全

(12) 指

(13) ワイン

(14) 農場

(15) 深い

(16) 銃

(17) 教える

(18) 科学者

(19) 雨

(20) 歌う

(21) 柔らかい

(22) 椅子

(23) 新聞

(24) 女王

(25) 秘密

(26) 寒い

(27) 機械

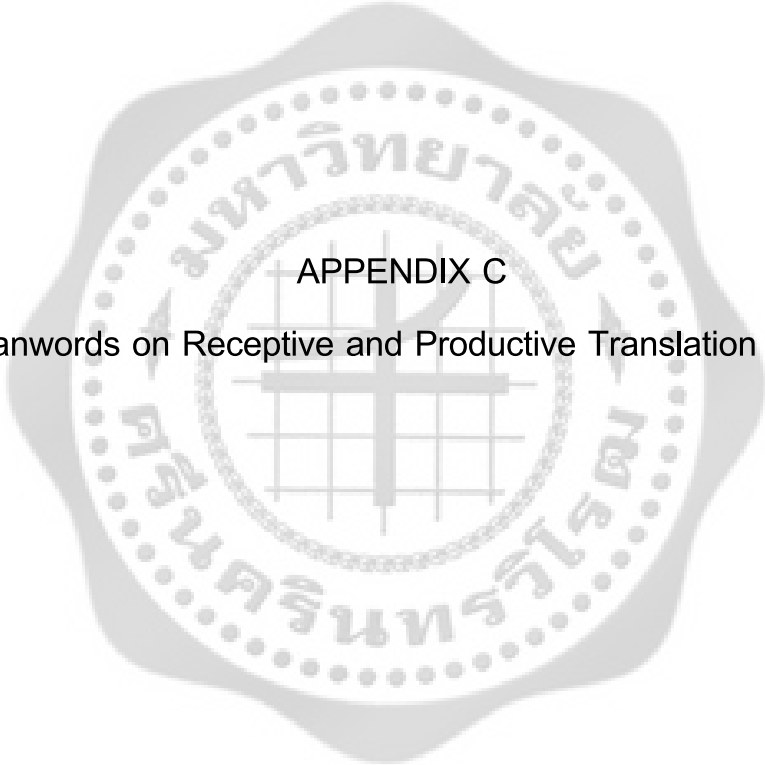
(28) 切符

(29) 雑誌

(30) 服

31	pregnant	_____	(31)	薬	_____
32	ambassador	_____	(32)	ポケット	_____
33	pollution	_____	(33)	宮殿	_____
34	joke	_____	(34)	英雄	_____
35	jockey	_____	(35)	ビール	_____
36	fence	_____	(36)	冬	_____
37	mask	_____	(37)	耳	_____
38	ceremony	_____	(38)	キス	_____
39	castle	_____	(39)	嫌う	_____
40	kilometer	_____	(40)	空	_____
41	bowl	_____	(41)	ボタン	_____
42	milk	_____	(42)	ベルト	_____
43	boil	_____	(43)	隠れる	_____
44	birthday	_____	(44)	歌手	_____
45	mirror	_____	(45)	屋根	_____
46	stomach	_____	(46)	牛	_____
47	boots	_____	(47)	洗う	_____
48	yellow	_____	(48)	砂	_____
49	license	_____	(49)	野菜	_____
50	spray	_____	(50)	甘い	_____
51	jewel	_____	(51)	草	_____
52	actor	_____	(52)	パイロット	_____
53	toy	_____	(53)	雪	_____
54	tower	_____	(54)	骨	_____
55	twins	_____	(55)	月	_____
56	steal	_____	(56)	親戚	_____
57	whisper	_____	(57)	塩	_____
58	neighbor	_____	(58)	葉	_____
59	jazz	_____	(59)	ピンク	_____
60	drum	_____	(60)	たんぱく質	_____

61	stranger		(61)	霧	
62	calendar		(62)	のり 糊	
63	stairs		(63)	靴下	
64	carpet		(64)	青あざ	
65	collar		(65)	ジム	
66	cereal		(66)	趣味	
67	hungry		(67)	誓う	
68	kidnap		(68)	ほお 頬	
69	spice		(69)	ふくろう	
70	jungle		(70)	ハンマー	
71	strawberry		(71)	潜る	
72	tomato		(72)	葬式	
73	sour		(73)	泡	
74	rainbow		(74)	雷	
75	wig		(75)	スプーン	
76	handsome		(76)	抱き合う	
77	shallow		(77)	オリーブ	
78	umbrella		(78)	テント	
79	vertical		(79)	麻	
80	nest		(80)	はしご	
81	trumpet		(81)	数学	
82	lily		(82)	天井	
83	lazy		(83)	ガソリン	
84	zoo		(84)	熟した	
85	slave		(85)	枕	
86	massage		(86)	ディスコ	
87	miracle		(87)	レシピ	
88	candle		(88)	豚肉	
89	lion		(89)	訳す	
90	liquid		(90)	ちょう 蝶	

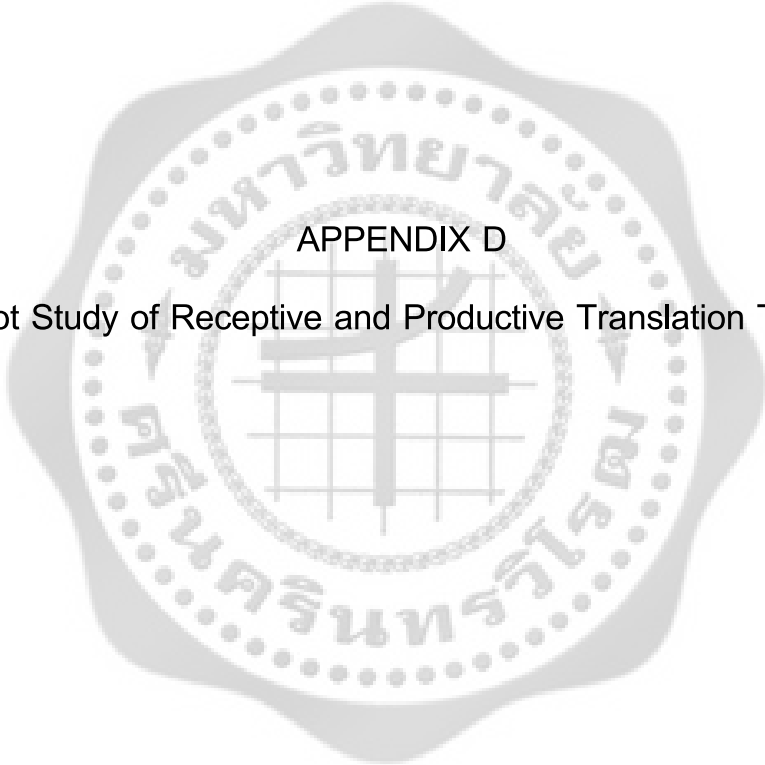


APPENDIX C

Loanwords on Receptive and Productive Translation Tests

Receptive (Version A) Productive (Version B)		Receptive (Version B) Productive (Version A)		
Word band 1				
	sport	3	* bottle	7
*	ticket	28	camera	19
	wine	13	* fruit	6
			hotel	11
			* kitchen	26
			video	27
			* wear	22
Word band 2				
	beer	35	boots	47
	belt	42	bowl	41
	button	41	drum	60
*	hero	34	fence	36
	kiss	38	jazz	59
	pilot	52	kilometer	40
	pink	59	* mask	37
	pocket	32	* milk	42
*	protein	60	spray	50
			* tower	54
Word band 3				
	disco	86	calendar	62
	gasoline	83	* carpet	64
	gym	65	cereal	66
	hammer	70	handsome	76
	olive	77	jungle	70
	recipe	87	lion	89
*	socks	63	massage	86
	spoon	75	* spice	69
	tent	78	tomato	72
			trumpet	81

Note. * A word that has an equivalent Japanese translation but is commonly accepted as a loanword. Loanwords are listed in alphabetical order. The number following each word indicates the item number on the tests.



APPENDIX D

Pilot Study of Receptive and Productive Translation Tests

Means, standard deviations, and ratios for the total receptive and productive scores

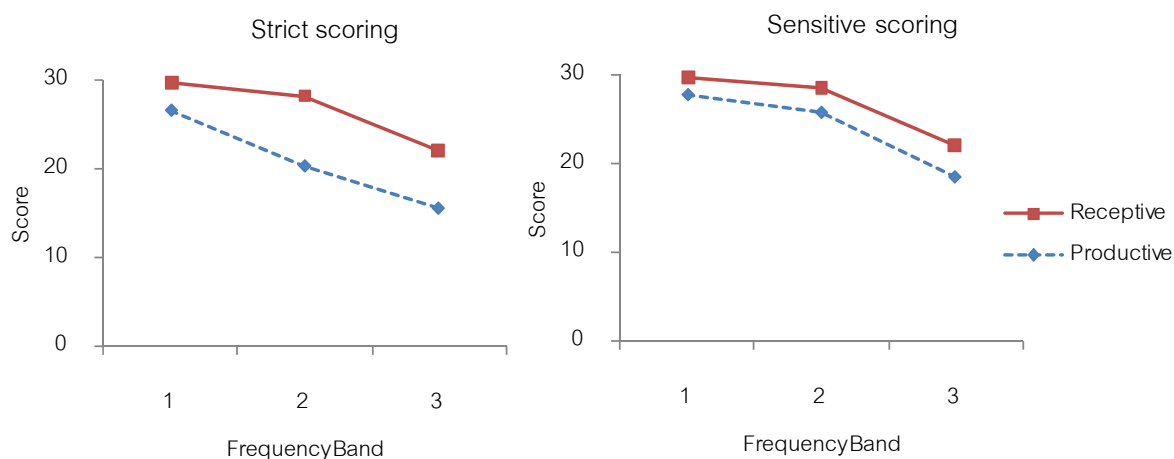
Scoring	Receptive	Productive	P/R ratio
	M	M	
Sensitive	80.25	72.00	90%
Strict	80.00	62.25	78%

Maximum score = 90, n=4

Means, standard deviations, and ratios for the receptive and productive scores at different word frequency bands

Scoring	Word band 1	Word band 2	Word band 3
	M	M	M
Sensitive			
Receptive	29.75	28.50	22.00
Productive	27.75	25.75	18.50
Ratio	93%	90%	84%
Strict			
Receptive	29.75	28.25	22.00
Productive	26.50	20.25	15.50
Ratio	89%	72%	70%

Maximum score = 30, n=4



Receptive and productive scores at each frequency band



APPENDIX E

Reading Comprehension Test (EIKEN Grade2)

<Directions>

Read each passage and choose the best answer from the four choices for each question. Then, on your answer sheet, find the number of the question and circle your answer.

PASSAGE 1

From: Tom North <tnorth@interlink-s.com>
To: Stephanie Brenner <sbrenner@steptwopublishing.com>
Date: June 13, 2008
Subject: Next week

Dear Ms. Brenner,

I have made the arrangements for your visit to my company, Interlink Services, next week. My co-workers and I are excited about having the chance to meet the author of *The Secrets of Successful Marketing* in person. You said in your e-mail that your presentation would be based on the marketing issues you wrote about in your book. I think that is an excellent idea. I put your plane tickets in the mail yesterday, so they should arrive soon. You will be taking Southeastern Airways Flight 258, departing from Atlanta at 9:30 a.m. on June 18 and arriving in Boston at 12:10 p.m. I will meet you at the airport and take you to your hotel. I have booked a room for you at the Majestic Hotel, which is not far from our office.

We will use a conference room in our office building for the presentation. It holds about 50 people. We will set up a whiteboard for you to use, but please let us know if you will need any other equipment, such as a computer or a projector. Also, are you available after the presentation? If so, perhaps you would like to join me and the president of Interlink Services for dinner.

Sincerely yours,

Tom North

Human Resources Manager, Interlink Services

- (1) What will Ms. Brenner do in Boston?
- A. Meet a famous author in person.
 - B. Discuss her ideas for a new book.
 - C. Give a presentation on marketing.
 - D. Make arrangements for a trip.
- (2) What does Tom North say he will do on June 18?
- A. Fly to Boston from Atlanta.
 - B. Put some flight tickets in the mail.
 - C. Book a room at the Majestic Hotel.
 - D. Pick Ms. Brenner up from the airport.
- (3) Tom North asks Ms. Brenner if she
- A. will bring a whiteboard with her.
 - B. is free for a meal in the evening.
 - C. wants to use a larger conference room.
 - D. has met the president of Interlink Services.

PASSAGE 2

Make Yourself Comfortable

Parents often tell their children to sit with their backs straight because they believe it is better for the backbone. Until recently, however, scientists had not been able to test the truth of this idea. This was because the scanners used by doctors to look at the backbone only worked when patients were lying down. Now, a new machine has been developed that can scan the bones of a person in a sitting position. A team of scientists in Aberdeen, Scotland, has used the new machine to test various theories about sitting and has come up with some surprising results.

The research team studied 22 healthy volunteers who had no problems with their backs. The new machine was used to scan the volunteers while they were sitting in a variety of positions. It was found that sitting with the back at a 90-degree angle actually puts a lot of stress on the lower part of the backbone. Leaning forward over a desk or a computer is even more harmful. Surprisingly, the healthiest position was found to be leaning backward at an angle of 135 degrees, with the feet flat on the floor.

Dr. Waseem Bashir, the leader of the research team, believes that these results are important because back pain is such a common problem. According to research carried out in the United States, 80 percent of people suffer from back pain at some time in their lives. Moreover, it is one of the most common reasons for people to visit a doctor. Dr. Bashir believes that companies should provide employees with chairs that allow them to avoid damaging their backs. If they do this, he says, they will be able to significantly reduce the number of days that people miss work because of back pain.

Dr. Bashir points out that people need to be practical as well. There are many tasks, such as using a computer, that are difficult to do if one is leaning backward. Also, it is important to move around and change position from time to time. Above all, he says, people should avoid spending a long time leaning over computers or video games. Parents will certainly agree with this piece of advice.

- (4) Recently, scientists have
- A. been testing a special kind of chair made for children.
 - B. been trying to develop machines that can repair damaged bones.
 - C. found a way to scan the backbones of patients who are not lying down.
 - D. discovered that scanners can damage people's backs if used incorrectly.
- (5) What is one discovery the researchers made?
- A. The most common sitting position is at a 90-degree angle.
 - B. Sitting up straight in a chair can be harmful to the backbone.
 - C. The key to sitting comfortably is to put one's feet on a chair.
 - D. People tend to sit in a variety of positions during the day.
- (6) What is one thing Dr. Waseem Bashir believes?
- A. People should take more time off from work to avoid getting back pain.
 - B. His research may result in a cure for 80 percent of people with back pain.
 - C. Companies should use chairs with an appropriate design for their employees.
 - D. More research is needed to find the healthiest way of sitting when at work.
- (7) Which of the following statements about Dr. Bashir is true?
- A. He understands that it is not always possible for people to sit back at an angle of 135 degrees.
 - B. He says that it is not practical for parents to tell their children to stop playing video games.
 - C. He thinks that a major cause of back pain is changing sitting positions too often.
 - D. He found that the best treatment for back pain is different for every individual.

PASSAGE 3

A SMART Idea

Kuala Lumpur, Malaysia, is a rapidly growing city built at the point where the Klang River and the Ampang River meet. The city is known for its traffic jams, as well as for the sudden floods that can cover the streets with water during heavy rainstorms. During the morning and evening rush hours, there are often traffic jams several kilometers long, and when heavy rains flood the roads, driving becomes extremely dangerous and traffic problems get worse. In the past, it was too costly to construct effective defenses against floods, but now engineers believe they have found a way to stop floods and reduce traffic jams in Kuala Lumpur at the same time.

These engineers have designed a three-level tunnel known as the Stormwater Management and Road Tunnel (SMART). In this tunnel, the upper and middle levels function as highways for traffic to flow smoothly, providing cars with direct access to the city center. When necessary, the bottom level of the tunnel carries rainwater to a storage lake.

The SMART tunnel has three different modes of operation. In the normal mode, when there is little rainfall, it is used only for traffic. In the second mode, during moderate storms, the bottom level carries away rainwater while the upper two levels remain open for traffic. During the severe storms that hit Kuala Lumpur two or three times a year and flood the area with an enormous amount of water, the tunnel goes into the third mode. In this mode, the tunnel is closed to traffic entirely. Special gates are then opened so that water can fill the whole tunnel and flow away quickly.

The tunnel has been carefully designed to be as safe as possible. When it goes into the third mode, monitoring stations make sure that every vehicle has left before the gates are opened to the floodwaters. The tunnel carries these floodwaters away so efficiently that vehicles are allowed to enter the tunnel again within 48 hours, bringing traffic back to normal. The Malaysian government is confident that this new tunnel will be able to serve both of the functions for which it was designed.

- (8) In Kuala Lumpur,
- A. engineers are building flood walls that are several kilometers long.
 - B. bad weather often makes the construction of new roads impossible.
 - C. the heaviest rainstorms are usually during the morning.
 - D. floods sometimes cause traffic jams to get worse.
- (9) The Stormwater Management and Road Tunnel provides a way for
- A. water to pass directly to the city center for use in homes.
 - B. traffic to reach a lake without crowding the highways.
 - C. rainwater to flow away safely below traffic lanes.
 - D. engineers to change the direction of traffic when there are problems.
- (10) What is one thing that happens when the SMART tunnel is in its second mode?
- A. Special gates open to allow traffic into the bottom level.
 - B. Water flows into the upper level from the storage lake.
 - C. Moderate storms force people to use the tunnel.
 - D. The tunnel is used both for rainwater and traffic.
- (11) Floodwater can enter the tunnel's three levels when
- A. all the traffic inside has been cleared out.
 - B. the storm has continued for more than 48 hours.
 - C. monitoring stations have made sure the water level is not too high.
 - D. its designers have checked to see if the gates are working efficiently.
- (12) Which of the following statements is true?
- A. The gates are closed to stop too much traffic from entering the tunnel.
 - B. The entire tunnel will be used for floodwater only a few times a year.
 - C. The routes of two rivers in Kuala Lumpur were changed by engineers.
 - D. The Malaysian government hopes to build a second SMART tunnel in Kuala Lumpur.



APPENDIX F
Comprehensive Data on EIKEN

Number of applicants and those who passed EIKEN

Year	1	Pre-1	2	Pre-2	3	4	5	Total
	Applicant	Applicant	Applicant	Applicant	Applicant	Applicant	Applicant	Applicant
	Passed	Passed	Passed	Passed	Passed	Passed	Passed	Passed
2005	22,033	68,614	316,323	514,158	714,901	524,358	329,027	2,489,414
	1,949	8,462	74,608	187,407	372,605	370,310	277,980	1,293,321
	8.8%	12.3%	23.6%	36.4%	52.1%	70.6%	84.5%	52.0%
2006	24,076	71,211	318,099	509,872	698,490	505,108	322,073	2,448,929
	2,456	10,198	69,987	183,655	375,472	359,897	272,114	1,273,779
	10.2%	14.3%	22.0%	36.0%	53.8%	71.3%	84.5%	52.0%
2007	24,967	69,469	310,726	511,742	681,785	483,247	315,151	2,397,087
	2,548	10,807	68,834	183,323	374,860	349,016	264,509	1,253,897
	10.2%	15.6%	22.2%	35.8%	55.0%	72.2%	83.9%	52.3%
University Students (2007)	2,861	12,963	19,993	3,785	840	179	59	40,680
	344	2,170	6,950	1,661	410	154	51	11,740
	12.0%	16.7%	34.8%	43.9%	48.8%	86.0%	86.4%	28.9%

The Society for Testing English Proficiency (2008)

Motivational goals of EIKEN Grade and benchmarks designated by Ministry of Education, Culture, Sports, Science, and Technology

EIKEN Grade	Ability Level	
1	Can understand the use of language necessary to practice effectively in a wide range of social, professional, and educational situations.	
Pre-1	Can understand the use of language necessary to practice effectively in social, professional, and educational situations.	Benchmark
2	Can understand the use of language at a level sufficient to allow him/her to take part in social, professional, and educational situations.	
Pre-2	Can understand the use of language at a level sufficient to allow him/her to take part in general aspects of daily life.	Recommended level to be achieved upon graduation from junior high school.
3	Can understand the use of language concerning familiar everyday topics such as likes and dislikes and basic personal family information.	
4	Can understand and use simple words, phrases, and short sentences.	
5	Can understand and use basic words, phrases, and simple greetings.	

The Society for Testing English Proficiency (2008)



APPENDIX G

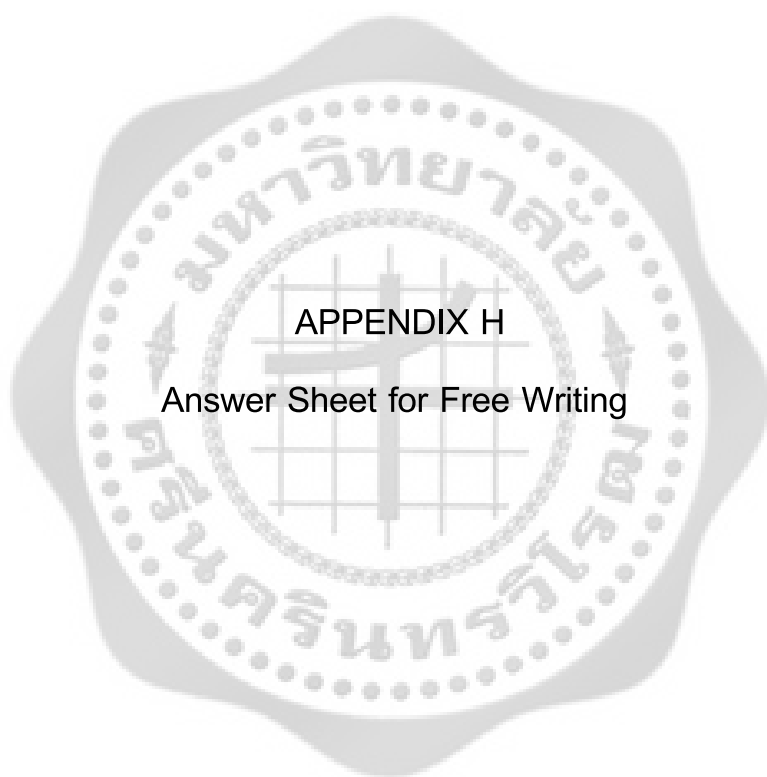
Directions for Free Writing

<Directions>

You will have 60 minutes to plan and write the essay assigned below.

1. You may wish to use your 60 minutes in the following way:
 - (1) 10 minutes planning what you are going to write
 - (2) 40 minutes writing
 - (3) 10 minutes reading and correcting what you have written.
2. Write your essay on the lined page of distributed paper.
3. You must write your essay on one of the following topics.
 - A)Family / Hometown
 - B)Club activity / Special interest
 - C)Part-time job / Volunteer work
 - D)Career / Future plan
4. Try to express your thoughts clearly and organize your ideas so that they will make sense to a reader.

Developed from Song and Caruso (1996)



APPENDIX H

Answer Sheet for Free Writing

APPENDIX I

Evaluation Scale for the Writing Skills Assessment Test

6	The essay provides a well-organized response to the topic and maintains a central focus. The ideas are expressed in appropriate language. A sense of pattern of development is present from beginning to end. The writer supports assertions with explanation or illustration, and the vocabulary is well suited to the context. Sentences reflect a command of syntax within the ordinary range of standard written English. Grammar, punctuation, and spelling are almost always correct.
5	The essay provides an organized response to the topic. The ideas are expressed in clear language most of the time. The writer develops ideas and generally signals relationships within and between paragraphs. The writer uses vocabulary that is appropriate for the essay topic and avoids oversimplifications or distortions. Sentences generally are correct grammatically, although some errors may be present when sentence structure is particularly complex. With few exceptions, grammar, punctuation, and spelling are correct.
4	The essay shows a basic understanding of the demands of essay organization, although there might be occasional digression. The development of ideas is sometimes incomplete or rudimentary, but a basic logical structure can be discerned. Vocabulary generally is appropriate for the essay topic but at times is oversimplified. Sentences reflect a sufficient command of standard written English to ensure reasonable clarity of expression. Common forms of agreement and grammatical inflection are usually, although not always, correct. The writer generally demonstrates through punctuation an understanding of the boundaries of the sentence. The writer spells common words, except perhaps so-called "demons," with a reasonable degree of accuracy.
3	The essay provides a response to the topic but generally has no overall pattern of organization. Ideas are often repeated or undeveloped, though occasionally a paragraph within the essay does have some structure. The writer uses informal language occasionally and records conversational speech when appropriate written prose is needed. Vocabulary often is limited. The writer generally does not signal relationships within and between paragraphs. Syntax is often rudimentary and lacking in variety. The essay has recurrent grammatical problems, or because of an extremely narrow range of syntactical choices, only occasional grammatical problems appear. The writer does not demonstrate a firm understanding of the boundaries of the sentence. The writer occasionally misspells common words of the language.
2	The essay begins with a response to the topic but does not develop that response. Ideas are repeated frequently or presented randomly, or both. The writer uses informal language frequently and does little more than record conversational speech. Words are often misused, and vocabulary is limited. Syntax is often tangled and is not sufficiently stable to ensure reasonable clarity of expression. Errors in grammar, punctuation, and spelling occur often.
1	The essay suffers from general incoherence and has no discernible pattern of organization. It displays a high frequency of error in the regular features of standard written English. Lapses in punctuation, spelling, and grammar often frustrate the reader. Or, the essay is so brief that any reasonably accurate judgment of the writer's competence is impossible.



VITAE

VITAE

Name : Rie Sakai

Date of Birth: February 26, 1974

Place of Birth : Japan

Address : 20, Sukhumvit 31, Klongtoey, Wattana, Bangkok, 10110

Educational Background:

1998 Bachelor of Arts (Social Science), Waseda University, Tokyo, Japan

2009 Master of Arts (Teaching English as a Foreign Language) ,
Srinakharinwirot University, Bangkok

