Fall 2014

Relationships between lexical proficiency and L2 oral proficiency

Yunjung You
Purdue University

Graduate School Form 30
(Updated 11/20/2014)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Yunjung You

Entitled

RELATIONSHIPS BETWEEN LEXICAL PROFICIENCY AND L2 ORAL PROFICIENCY

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

April Ginther

Tony Silva

Margie Berns

Dwight Atkinson

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification/Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

April Ginther

Approved by Major Professor(s): _____

_____

Approved by: Nancy Peterson 12/01/2014

Head of the Department Graduate Program Date

RELATIONSHIPS BETWEEN LEXICAL PROFICIENCY AND L2 ORAL PROFICIENCY

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Yunjung You

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2014

Purdue University

West Lafayette, Indiana

Dedicated to my kids *Soojung* and *Minjae*, my husband *Dr. Sungwon Kim*,

and my mother and father in Korea *Sookhee Hong* and *Hogu You*.

Without their encouragement and constant love, none of this would have happened.

이 논문을 사랑하는 아이들 수정이와 민재, 남편 김성원 박사,

그리고 한국에 계신 어머니 아버지께 헌정합니다.

이 분들의 격려와 사랑이 아니었다면 이 논문은 시작되지도 끝나지도 않았을 것입니다.

ACKNOWLEDGEMENTS

The last four years to study and work in the Department of English at Purdue University have been such blessing moments in my life. However, it didn't always go smooth. Learning something new was fun but painful many times. While I went through these harsh moments, it was four people who helped me stand up again: Dr. April Ginther, Dr. Tony Silva, Dr. Margie Berns, and Dr. Dwight Atkinson.

I would like to express my deepest appreciation to Dr. April Ginther for her endless support and advice during my whole doctoral period. While passing through this long tunnel, Dr. Ginther has been an anchor for me not only academically but also mentally. I know that all the praises and criticism from her have nourished me. I am and was truly happy I could be an advisee of hers.

Also, I would like to express my sincere gratitude to Dr. Tony Silva, Dr. Margie Berns, and Dr. Dwight Atkinson. Dr. Silva is the one who first introduced me the field of Second Language Studies and gave me a lot of advices throughout. Without Dr. Bern's encouragement and advice, I could've almost given up all this work when I seriously considered stopping due to the hard family situation. Also, Dr. Atkinson has inspired me academically, and from his advice, I was motivated to keep working in the way of scholar.

I truly appreciate you, Dr. Ginther, Dr. Silva, Dr. Berns, and Dr. Atkinson. It was such an honor for me to have the opportunity to be with you in my life.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

You, Yunjung. Ph.D., Purdue University, December 2014. Relationships Between Lexical Proficiency and L2 Oral Proficiency. Major Professor: April Ginther.

Lexical proficiency is strongly correlated with L2 productive language ability (Crossley, Salsbury, & McNamara, 2011; Laufer & Nation, 1995). While many studies have investigated the relationship between lexical proficiency and L2 writing ability, not many studies have been conducted in terms of spoken language. Also, few studies have focused on how non-native English speakers with different L1 backgrounds differently or similarly develop lexical proficiency in L2 speaking. Based on this background, the present study is conducted with three purposes: 1) to compare the effectiveness of measures of lexical proficiency in terms of their ability to predict the quality of L2 spoken production, as determined by scores on the Oral English Proficiency Test (OEPT) at Purdue University; 2) to compare the different lexical features among different levels of L2 oral proficiency; 3) to see whether different L1 backgrounds present different levels of development of lexical proficiency.

A quantitative research approach was selected in this study. Three hundred and three speech samples from the Oral English Proficiency Test (OEPT), representing four different L1 groups of Korean, Mandarin, Hindi, and English, were collected, and in order to measure lexical proficiency, the Lexical Frequency Profile (LFP) was employed.

Statistical inferences were based on Spearman rank order correlation coefficients and descriptive statistics.

Results showed that OEPT scores have strong or moderately strong correlations with some indices of lexical proficiency. Interestingly, however, different results were obtained when the Spearman rank-order correlation coefficients were rerun without Mandarin groups: the correlation coefficients increased for the variables with Types, TTR, D, K1 Type, K2 Type, AWL Type, and Off Type. Descriptive statistics suggested some reasons for the different results with and without the Mandarin group: basically, the Mandarin group created more lexical diversity and produced more words as compared to other sub groups than the Korean or Hindi groups did. Also, the Mandarin group created dissimilar trends in each LFP variable, unlike the Korean and Hindi groups. Also, descriptive statistics on the LFP indicate that L2 advanced speakers express their ideas employing frequent words as a rule rather than infrequent ones. This result indicates that when it comes to vocabulary learning for oral proficiency, frequent words should be a more important focus in L2 vocabulary learning than infrequent words. Each of these results is explored in detail in the Discussion section. Pedagogical implications and suggestions for future studies are discussed as well.

CHAPTER 1. INTRODUCTION

1.1 Introduction

"Close the book, get your desk empty, and get ready to have a vocabulary quiz" the English teacher said. So nervously, I went through the items on the quiz that the teacher passed out. Ten items. They were all about translating English words into Korean. I memorized 30 difficult words last night for this quiz. *I have to write the answers all before forgetting them! Recognize,인식하다. Notion, 개념. Joyful, 즐거운..*

Fortunately, I still held them in my head. It seems I have all the 10 items right. Oh, yeah!

The above scene is what happened in one of the English as a foreign language classes that I had when I was in high school in Korea. I remember that I recited 30 words of English almost every day. I believed the more words I memorized, the better English proficiency I would have. That was also what teachers recommended. So I bought a vocabulary book and memorized the spelling of English words as well as their meanings in Korean, as a one to one correspondence. Memorizing vocabulary words was not easy, but actually painful. For more effective memorization, I wrote them down on empty sheets hundreds of thousands of times and sometimes guessed word meanings by covering only part of the meaning in the vocabulary book. However, even though I eventually learned the word meanings and their spellings, I didn't know how to pronounce nor how to use them. Even worse, I didn't know what I didn't know.

Anyway, the scores of English exams and quizzes were not too bad even though I didn't enjoy learning English very much. When I first came to the U.S. at the age of 23; however, I could not even use one word in conversation, even though I had memorized them like crazy.

I was puzzled. What was the problem? I worked so hard on all of these difficult words, but still couldn't speak well. Did I choose the wrong selection of words? Or did I have to study them in different ways? I definitely misunderstood how to learn and what words I need to know for speaking in a casual conversation. Furthermore, my misunderstanding led me to the wrong method of vocabulary learning.

Many people who study a second language recognize the importance of lexical proficiency and make an effort to increase their lexical proficiency. However, they tend to be ignorant of how to learn the words. Why? This is because they are not aware of what characteristics and elements a vocabulary word contains as well as what lexical proficiency means. This is the same not only for learners but also for teachers. That is, in many cases, both learners and teachers do not recognize what they need to learn or teach to increase lexical proficiency because they do not have knowledge of the characteristics of vocabulary nor the meaning of lexical proficiency.

My personal experience has brought me to some fundamental questions: To speak well in a foreign or second language, how many vocabulary words should one master? Are there words more important and more frequently used in speech than others? Is there an effective way of learning vocabulary for improving L2 speaking proficiency? These questions from my own experience on second language learning have remained on my mind for a long time and served as a starting point for this study.

1.2 Background of the Problem

Lexical proficiency has been strongly correlated with L2 productive language ability (Crossley, Salsbury, & McNamara, 2011; Engber, 1995; Ferris, 1994; Goodfellow, Lamy, & Jones, 2002; Laufer & Nation, 1995; Mellor, 2011), and the development of studies on lexical proficiency and L2 proficiency in production parallels the development of measures of lexical proficiency. In the initial stage, the main methods of measuring lexical proficiency simply involve 1) counting the total number of words or the number of different words, 2) calculating the proportion of content words (lexical density), or 3) calculating the type-token ratio (TTR). For example, Engber (1995) employed lexical density and TTR for measuring lexical proficiency in L2 writing. However, a limitation of TTR is that its results are inconsistent with texts longer than 150 words; therefore, many transformations of TTR have been developed to overcome the weakness of TTR (e. g., Guiraud's Index, Yule's K, the measurement D, Hapax, an estimate of Advanced Guiraud). Mellor (2011) is a good example of a study that employs transformations of TTR to more accurately measure lexical proficiency. These transformations are considered more stable than the traditional TTR, although they still contain the fundamental problems of being affected by the number of tokens (Malver, Richards, Ciper, & Duran, 2004). To address this issue, the measurement D was developed as a statistical model of TTR that could discriminate L2 proficiency better than traditional TTR and its transformations (Malvern et al., 2004). In addition, the Lexical Frequency Profile was developed by Laufer and Nation (1995), different from the category of lexical diversity and based on the assumption that each word is used with different frequency in reality, which should therefore reflect how easily the word can be learned. The studies of

Laufer and Nation (1995) and Goodfellow et al. (2002) are examples of studies that employ a Lexical Frequency Profile. In short, the above studies on lexical proficiency and L2 productive language proficiency can be used to trace the development of even increasingly more sensitive measures of lexical proficiency.

1.3 Purposes of the Current Research

The present study has three purposes: 1) to compare the effectiveness of measures of lexical proficiency in terms of their ability to predict the quality of L2 spoken production, as determined by scores on the Oral English Proficiency Test (OEPT) at Purdue University; 2) to compare the different lexical features among different levels of L2 oral proficiency; 3) to see whether different L1 backgrounds present different levels of lexical proficiency.

1.4 Research Design

A quantitative research approach was selected for this study. Three hundred and three item responses from the Oral English Proficiency Test (OEPT), representing four different L1 groups of Korean, Mandarin, Hindi, and English, were selected and analyzed. For data analysis, two software programs, VocabProfile 3.0 and Coh-Metrix, were adopted. In order to measure lexical proficiency, lexical diversity (tokens, types, type/token ratio, and *D*) and the Lexical Frequency Profile were employed. Statistical inference was made using Spearman rank-order correlation coefficients and descriptive statistics.

CHAPTER 2. THEORETICAL FRAMEWORK

2.1 Lexical Proficiency

The nature and operationalization of lexical proficiency is a topic of great interest and debate in the literature on second language acquisition. For example, Anderson and Freebody (1981) define lexical proficiency by its breadth and depth, while Henriksen (1999) sees lexical proficiency as having three dimensions: "partial to precise knowledge," "depth of knowledge," and "receptive to productive" use ability. Nation (2001) and Milton (2009) agree that a lexical item has two dimensions in terms of its use: receptive and productive. Nation (2001), however, argues that knowing a word involves three different aspects: form, meaning, and use of vocabulary. Daller, Milton, and Treffers-Daller (2007) also explain lexical proficiency within three dimensions: breadth, depth, and fluency. Each of these studies emphasizes and operationalizes lexical proficiency in compelling and useful ways and shall receive more detailed treatment below.

2.1.1 The Theory of Breadth and Depth Dimensions

Anderson and Freebody (1981) suggested one way to view lexical proficiency: breadth and depth of word knowledge. Breadth of lexical proficiency simply concerns how many words one knows, while depth of lexical proficiency is related to what one knows about that vocabulary (e.g., collocations, associations, grammatical rules, and

specific usage). The dimensions of breadth and depth are concretized by two models: Henriksen's three dimensions of lexical competence (1999) and the theory of lexical space (Daller et al., 2007).

2.1.1.1 Henriksen (1999)'s Three Dimensions of Lexical Competence

Henriksen (1999) argues that there are three dimensions of lexical competence: (1) "partial to precise knowledge," (2) "depth of knowledge," and (3) "receptive to productive" use ability. That is, the first dimension, partial to precise knowledge, defines how precisely one recognizes and comprehends a word. This dimension is also related to the size and breadth of one's vocabulary pool (i.e., how many words one knows). Second, the depth of knowledge dimension is related to "the quality of the learner's vocabulary knowledge" (Read, 1993, p. 357). The third dimension, the receptive to productive dimension, is based on the idea that one's receptive ability is different from one's productive ability in terms of L2 vocabulary acquisition.

2.1.1.2 Lexical Space: Breadth, Depth, and Fluency (Daller et al., 2007)

Daller et al. (2007) explain lexical proficiency as the concept of lexical space, which is comprised of three dimensions: breadth, depth, and fluency (Figure 2-1). The first dimension, breadth, can be represented as the vertical axis. The dimension of breadth reflects how many vocabulary words a learner knows. According to Daller et al. (2007), the whole part of *form* and the *form and meaning* under *meaning* in Nation's (2001) definition of lexical proficiency can be included in the dimension of breadth. The next dimension, depth, is represented as the vertical axis. Depth is used to explain what one knows about a word. Regarding Nation's model of lexical proficiency, the dimension of

depth reflects part of *meaning* and the whole sub-divisions of *use*: *concepts and referents*, *associations*, *grammatical functions*, *collocations*, and *constraints on use*. The last dimension is fluency, which explains how one productively uses words and the level of automaticity in production. Fluency is also considered an essential dimension in the model of lexical space.

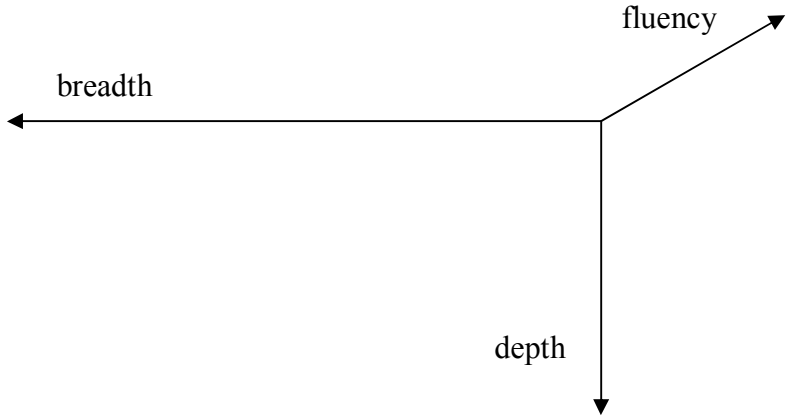

Figure 2-1 The Lexical Space: Dimensions of Word Knowledge and Ability

(Daller et al., 2007, p. 8)

The model of lexical space by Daller et al. (2007) is useful from three perspectives. First, it explains lexical proficiency with the concept of "space": in the model, lexical proficiency is expressed with three axes, and this representation is helpful to understand lexical proficiency. Second, the model of lexical space includes the dimension of *fluency*. With the dimension of fluency, one can better understand how the productivity of language can be defined in lexical proficiency. Lastly, the model of lexical space shows the relationship between Nation's (2001) definition of lexical

proficiency and the traditional categories of breadth and depth so that one better grasps the meaning and boundary of lexical proficiency.

Even though the model, lexical space, seems a reasonable representation of lexical proficiency using the three axes of breadth, depth, and fluency, it is still ambiguous in some aspects. First, the model of lexical space does tell us about what each of the three axes is; however, it does not give clear explanation about the relationship among those axes, especially between breadth and fluency and between depth and fluency. Second, fluency is not clearly explained in this model. Daller et al. (2007) argue that fluency is an attempt to define how one easily can access each lexical item and fluency is related to speed and accuracy. This explanation, however, is not enough to explain the productive dimension of lexical proficiency. Lastly, the model of lexical space does not clearly explain how to consider the categories of reception and production. One can assume that the axis of fluency could be part of production, but a clearer explanation of how to connect the model of lexical space to the receptive and productive dimensions of lexical proficiency is needed.

2.1.1.3 Distinction Between Breadth and Depth

As discussed above, viewing lexical proficiency through the frame of breadth and depth provides a simple and convenient means of clarifying different aspects of lexical knowledge. However, researchers have been unable to make a clear cut distinction between breadth and depth practically even though it has been possible to define each theoretically; therefore, it is not clear enough to establish exact definitions of breadth and depth or to determine their boundaries are (Milton, 2009).

Nevertheless, recently, Crossley, Salsbury, and McNamara (2011) have categorized breadth and depth of lexical proficiency according to particular measures of lexical proficiency: Breadth is measured by the Lexical Frequency Profile and lexical diversity and depth are measured by conceptual levels, word associations, polysemy, and semantic similarity. Crossley et al. (2011) explain as follows:

> Since both lexical diversity and word frequency indirectly assess how many words a learner knows, they are generally categorized as breadth of knowledge measures…. Unlike breadth of knowledge measures, depth of knowledge features are not based on the number or variety of words a learner produces, but on constraints at the phonemic, morphemic, and syntactic level (Qian, 2004) and deeper level features related to word associations. (p. 3~4)

In other words, Crossley et al. (2011) define the difference between breadth and depth with measures of lexical proficiency, allowing for a more clear-cut distinction. This provides general guidelines on how to view breadth and depth.

However, Crossley and collegues (2011)'s distinction, which is to categorize two dimensions of breadth and depth by measures of lexical proficiency, cannot provide a clear separation between breadth and depth. It is mainly because this approach is still based on a vague definition of breadth and depth. For instance, the Lexical Frequency Profile, one of the measures used in the study above, is categorized as breadth (Crossley et al., 2011). However, there still is a possibility to define the Lexical Frequency Profile as depth: the Lexical Frequency Profile is based on the concept that each word is used in different frequency which results indifferent levels of ease to learn each word. In other words, the Lexical Frequency Profile can be considered breadth as it deals with counting

words. At the same time, however, it can also be thought of as depth as it also deals with ease of acquiring and frequency of words.

One study that distinguishes between breadth and depth is Vermeer's (2001) who investigated the relationship between the two, using receptive vocabulary, description, and association tasks. In this study, Vermeer found that breadth of vocabulary knowledge is highly correlated with its depth. In other words, there is little difference between breadth and depth of vocabulary knowledge for all practical purposes, despite the theoretical distinction. Even though this research has been critiqued because there is no consensus among researchers for the relationship between the tasks used in this study and the two dimensions of breadth and depth, Vermeer's study grants some evidence that the distinction between breadth and depth is ambiguous.

2.1.2 The Theory of Receptive and Productive Dimensions

In addition to the dimensions of breadth and depth, there are two other dimensions that have played an essential role to explain lexical proficiency: receptive and productive dimensions. The receptive dimension of lexical proficiency focuses on one's ability to understand a word when listening or reading, while the productive dimension focuses on how one uses a word for expression when speaking or writing. Nation (2001) explains the distinction between receptive and productive vocabulary use as follows:

> Receptive vocabulary use involves perceiving the form of a word while listening or reading and retrieving its meaning. Productive vocabulary use involves wanting to express a meaning through speaking or writing and retrieving and producing the appropriate spoken or written word form. (p. 24)

The productive dimension is also called the "active" dimension, while the receptive dimension is labeled as "passive" (Milton, 2009, p. 13; Nation, 2001, p. 24). The basic assumption under this view is that people use different cognitive processes when passively understanding a word as compared to when actively using it. Regarding the distinction between receptive and productive vocabulary, it is generally assumed that one's receptive vocabulary knowledge is larger than his or her productive vocabulary knowledge (Milton, 2009) and that learning productive vocabulary is harder than learning receptive vocabulary (Nation, 2001).

2.1.2.1 Nation's Model (2001) of Lexical Proficiency

Nation's model (2001) of lexical proficiency defines lexical proficiency based on the receptive and productive dimensions (Table 2-1). Nation (2001) argues that knowing a word involves three different parts: form, meaning, and use. These three parts of lexical proficiency reflect both receptive and productive vocabulary knowledge. The first part, *form,* is comprised of three different components: spoken form, written form, and word parts. That is, when one can say "I know the form of a word," one needs to know what the word sounds like, how the word is spelled, and how the word is morphologically formed, e.g., affixes and a stem. The second part of knowing a word is meaning, which includes connecting form and meaning, concept and referents, and associations. Connecting form and meaning of a word is related to "how readily the learner can retrieve the meaning when seeing or hearing the word form or retrieve the word form when wishing to express the meaning" (Nation, 2001, p. 48). Concept and referents introduce about how one can relate a word to one's sense and references in reception or

production of that word. Regarding association, Nation argues that it is important to know the relationships among words and how to organize and group vocabulary words. The last part of knowing a word, *use*, is the concept of recognizing the conditions of when, where, and how one can use a word. *Use* involves grammatical functions, collocations, and constraints on use. That is, regarding knowing the use of a word, one needs to know how a vocabulary item grammatically fits into a sentence or a text (grammatical functions), what the word occurs with (collocations), and what sociolinguistic factors and constraints are related to using a word (constraints on use).

Table 2-1 What is Involved in Knowing a Word (Nation, 2001, p.27)

Form	Spoken	R	What does the word sound like?
		P	How is the word pronounced?
	Written	R	What does the word look like?
		P	How is the word written and spelled?
	Word parts	R	What parts are recognizable in this word?
		P	What word parts are needed to express the 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	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: In column 3, R = receptive knowledge, P = productive knowledge.

Nation's model of lexical proficiency has three advantages over the previous models of lexical proficiency. First, it provides a detailed explanation of what lexical proficiency is by categorizing the elements of vocabulary knowledge into small segments. Nation's model can be considered the first model that presents the idea that lexical proficiency includes not only simply knowing the meaning of a word but also understanding a host of complicated elements. Second, the model is practical and clear so that practitioners and L2 learners can easily understand the concept of lexical proficiency. In other words, based on understanding various dimensions of lexical proficiency as suggested in Nation's model, L2 learners and practitioners are able to develop effective and strategic plans for vocabulary learning. Third, the model explains all elements of lexical proficiency in the receptive and productive dimensions. In many cases, L2 learners do not have an exact understanding of the fact that the receptive vocabulary acquisition process is different from the productive process. Therefore, it can be said that this model is helpful to better grasp lexical proficiency vis-à-vis the two dimensions of production and reception.

As shown above, Nation's model offers a clear understanding of lexical proficiency. Nevertheless, Nation's model also contains three weaknesses. First is that the Nation's model of lexical proficiency does not explain how the concept of breadth and depth can be appropriately represented. In addition, Nation's model does not address the relationship between fluency and lexical proficiency. Also, Nation's model does not consider the developmental process (i.e., how lexical proficiency develops and grows). Overall, on one hand, Nation's model of lexical proficiency is very useful and practical for understanding lexical proficiency with the detailed explanation on the segmented

elements of a word. On the other hand, however, it still does not offer a comprehensive viewpoint on lexical proficiency, including the explanation on breadth and depth, fluency, and the developmental processes.

2.1.3 Summary and Necessity for a New Model of Lexical Proficiency

As discussed above, lexical proficiency has been thought of as being comprised of four different dimensions: breadth-depth and receptive-productive. Breadth and depth are related to the concept of quantity and quality, while the receptive and productive dimensions focus on functions and use of lexical proficiency. These four dimensions are practical and easy to understand. However, the theories regarding the four dimensions neither provide clear distinctions among dimensions nor explain all aspects of lexical proficiency. Therefore, there is a need to develop a new model to clarify lexical proficiency and to synthesize the previous theories of lexical proficiency.

2.1.4 The Cube: A New Model of Lexical Proficiency

In this section, I develop a new model of lexical proficiency, named the Cube. This new model is based on past theories of lexical proficiency which have been popularly used. This includes the breadth-depth theories with Henriksen's lexical proficiency model and Daller et al.'s lexical space model as well as the theory of receptive-productive dimensions with Nation's lexical proficiency model.

As Figure 2-2 shows, the Cube is three dimensional and defines one's lexical space. The Cube has three axes to show the development of the cube: breadth, depth, and fluency. The point where these three axes meet is the starting point where one's lexical space starts to grow. Each axis grows with the way of each arrow as above and

these three axes form three sides: reception, variety in production, and sophistication in production. Reception deals with only the receptive dimension of lexical proficiency while variety in production and sophistication in production encompass the productive dimension of lexical proficiency.

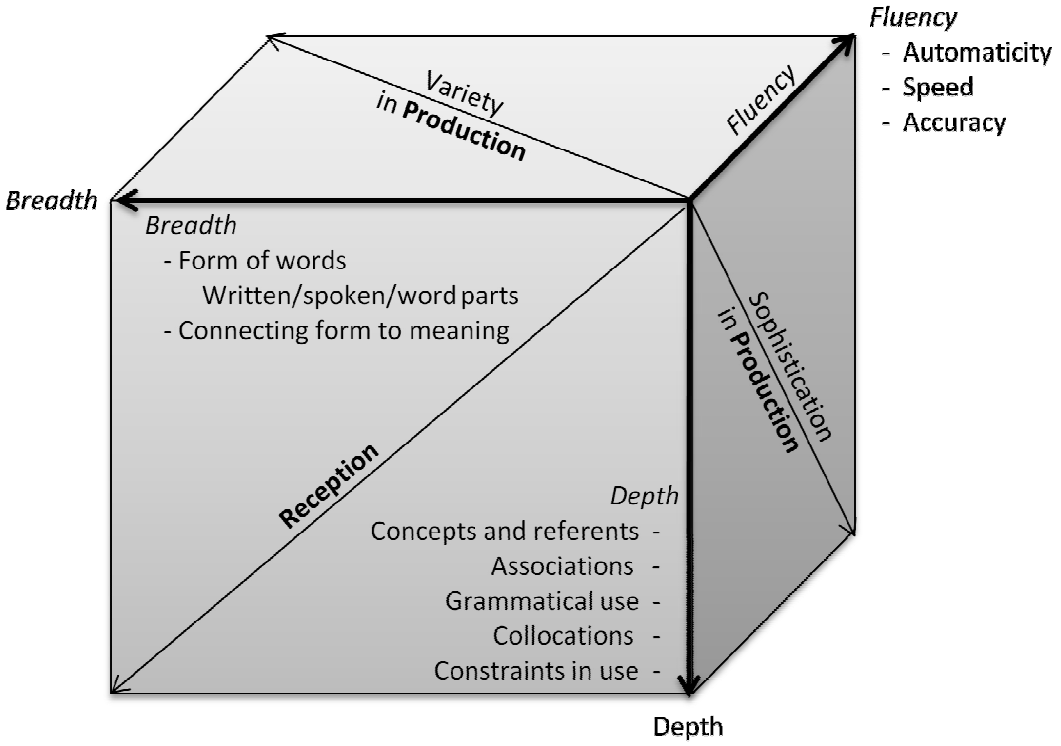

Figure 2-2 The Cube: A New Model Representing Lexical Proficiency

2.1.4.1 Three Axes of the Cube

There are three basic axes: *breadth*, *depth*, and *fluency*. These axes are the fundamental elements of the Cube. These are a reflection of Henriksen's (1999) explanation of lexical proficiency and Daller et al.'s (2007) model of lexical space. Axes provide a basic tool to explain how one's lexical space grows in a one dimensional view.

2.1.4.1.1 First Axis: Breadth

The first axis, breadth, allows a quantitative representation of a learner's lexical space. That is, breadth involves how many vocabulary words one knows. The axis of breadth has a thread of connection with other models which have explained breadth of lexical proficiency (e.g. partial to precise knowledge by Henriksen (1999) and breadth by Daller et al. (2007)). Nation (1999) explains that there are nine elements of lexical proficiency which are involved in knowing a word and these nine elements can be categorized into three segments, each of which has three elements. Of those nine elements, Daller et al. (2007) argue that breadth covers "form" and that "form and meaning" under the "meaning" segment fall into breadth. Based on this, the axis of breadth in the Cube covers how many vocabulary words one knows in terms of form and meaning.

2.1.4.1.2 Second Axis: Depth

Depth involves figuring out the identity of words in the relationship between words and contexts, while breadth is concerned only with its own form and meaning. Depth is the matter of how a word is used and recognized in a real situation and governed by which rules. In detail, the axis of depth can be defined, as Daller et al. (2007) explain, in the scope of how well one figures out a word's meaning and use in terms of Nation's definition of lexical proficiency. In terms of meaning, depth concerns how well one can connect a word's basic concept to relevant knowledge (concepts and referents) and can figure out the relationship between a word and other related words (association). The axis of depth also covers *use* of a word -- how to figure out the actual

usage and rules in using a word. In other words, the axis of depth defines the scope of how well one can figure out the grammatical pattern in which a word is used (grammatical use), how a word is used with other words as a group (collocations), and if there are any constraints to be considered for using a word (constraints in use).

The axis of depth grows on the base of development of breadth. This is because figuring out the form and meaning of a word is the first task to be done in vocabulary acquisition. Other elements of a word, such as association, grammatical use, or constraints in use, develop based on the form and meaning of a word.

2.1.4.1.3 Third Axis: Fluency

Fluency is the last axis. For fluency, the main focus is how one develops automaticity in producing a word. In the axis of fluency, the explanation of Daller et al. (2007) on fluency presents the main idea as follows:

> …this [fluency] is intended to define how readily and automatically a learner is able to use the words they know and the information they have on the use of these words. This might involve the speed and accuracy with which a word can be recognized or called to mind in speech or writing. (p. 8)

That is, as the dimension of fluency reflects how one can productively use a word in a certain situation; two key elements of fluency are speed and accuracy, which is linked to automaticity.

2.1.4.2 Three Sides of the Cube

There are three sides in the Cube. Each side is formed with two of three lines and shows one aspect of vocabulary development as more evolved than the axis. A side's function is to show the difference between reception and production.

2.1.4.2.1 First Side: Reception

The two axes of breadth and depth form *reception*, which plays a role in recognizing vocabulary words. The side of reception naturally forms with the growth of the axis of depth after the formation of breadth. Reception refers to the ability to recognize words in both written and spoken language. More specifically, reception defines one's lexical proficiency regarding how one passively understands and figures out words in the modes of reading and listening. The two axes of breadth and depth, composing the side of reception, are closely related with how well one understands a word at which level. In other words, when the axis of breadth grows, one is able to recognize the form and meaning of a word. On the other hand, with the growth of depth, one comes to understand the more complex functions of a word such as associations or grammatical usage. Therefore, without the axis of depth, it is nearly impossible to figure out the exact meaning of a sentence or the overall flow of a story while reading and listening without the development of the axis of depth, even though it is possible to understand the meaning of individual words with breadth. Namely, it is important that both axes of Breadth and Depth grow in balance for reception.

2.1.4.2.2 Second Side: Variety in Production

The second side is *variety in production*, which is formed with two axes, breadth and fluency. Unlike reception, this side considers how to productively use vocabulary. Reflecting breadth, variety in production quantitatively represents productive lexical proficiency. In other words, within variety in production, the important issue is how various vocabulary words can be produced in written or spoken forms, knowing the meaning of the word. If one can productively speak or write a word, recognizing its meaning satisfies the sufficient condition of variety in production. However, in variety in production, the more complex functions of lexical proficiency such as context or association are not considered.

2.1.4.2.3 Third Side: Sophistication in Production

The last side of the Cube, *sophistication in production*, is shaped with depth and fluency. This is another side of productive lexical proficiency, along with variety in production. Sophistication in production is the matter of how to use vocabulary, recognizing complex functions such as concepts and referents, associations, grammatical functions, collocations, and constraints of use. Sophistication in production is a matter of quality in productive lexical proficiency, while variety in production is about quantity in productive lexical proficiency. That is, how one uses a word with its precise usage and context is taken into account. Therefore, in order to expand sophistication in production, one should understand more complex levels of word functions in depth.

An important consideration for both variety in production and sophistication in production is fluency. As mentioned in the previous section, the side of fluency is

composed of two elements: speed and accuracy. When it comes to productive lexical proficiency, it is essential not only "what" words one productively uses but also how automatically one uses the words in terms of speed and accuracy. Automaticity of a word should be considered one vital element of productive lexical proficiency as different from how many words one knows in reception or uses in production. This can be better demonstrated through consideration of real conversation process. In a casual conversation, fluent speakers spend little time pulling out a certain word in order to express their ideas. That is, thinking and producing words take place almost at the same time; fluent speakers do not need to take a long time to say or understand one word. If it takes too much time to access and produce a word, even though one may receptively know the word and can produce the word, it can be said that the word is of little use in conversation. Namely, automaticity should be considered one important function of lexical proficiency aside from having knowledge of words because it is not very easy to actually use words which are not automated.

2.1.4.3 Distinction between Written and Spoken Forms of Language

The written form of lexical proficiency is visual and based on letters while the spoken form is acoustic and based on sound. These two forms of lexical proficiency may separately develop in different spaces. First of all, in terms of the developmental process of lexical proficiency, the written and spoken go through dissimilar processes. This becomes more obvious in the L1 developmental process; young babies acquire vocabulary words in spoken form first. They start to learn vocabulary in written form when they are 3-4 years old at the earliest or after 7 at the latest. Second, it is possible to

develop one's vocabulary only in one way. It is plausible to know only the spoken form of vocabulary without knowing the written form, and vice versa. If one has not been taught written language, one can rely on the spoken in living. Third, there is the case in which one knows only written forms of vocabulary, when he or she learns a foreign language. In this case, it is possible that one can read and write the foreign language, even though one cannot speak or understand it. To recap, the spoken and written forms of lexical proficiency build up separately in two different Cubes. Figure 2-3 presents the visual image of spoken and written forms of lexical proficiency developing in different spaces as follows:

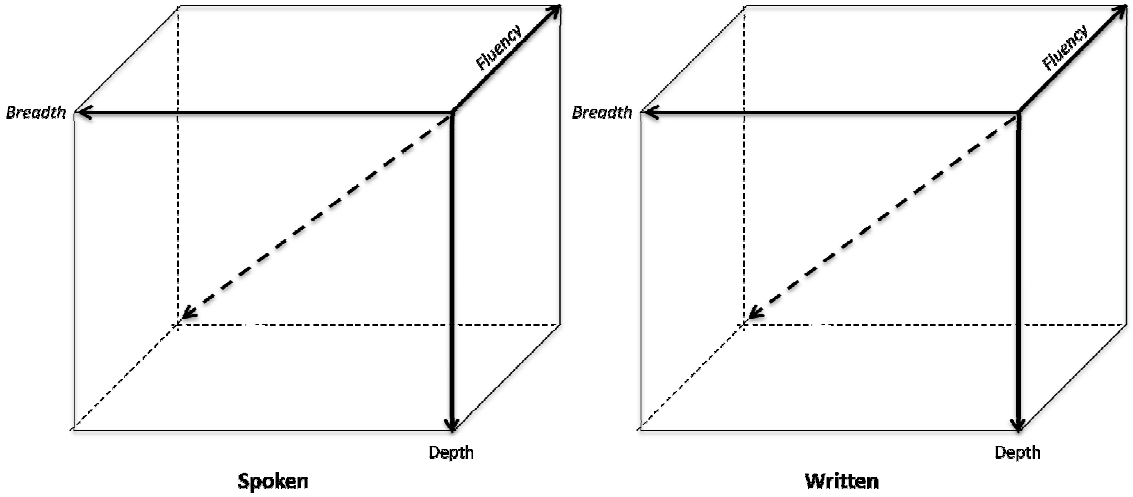

Figure 2-3 Spoken and Written Forms of Lexical Proficiency in Different Cubes

Spoken and written dimensions of lexical proficiency interact with each other and positively affect their mutual development, even though spoken and written forms develop in different spaces. Van Woerkum (2007) explains the relationship between spoken and written forms of language and how they affect one another:

> The underlying principle is quite simple. Words that are used often in everyday speaking and hearing are more familiar to people. More familiar words in written texts are recognized more quickly. This helps the reader to read a text faster and supports the efficiency of the reading process. (p. 189)

Van Woerkum's above explanation shows how the spoken form of language supports better understanding of the written form of language. Krashen (1989) also argues that people acquire vocabulary words from reading, that is, a written form of language, and use this vocabulary in their oral performance. This argument is supported by one of his theories, the Input Hypothesis (IH). Below is a part of his argument:

> The Input Hypothesis (IH) assumes that we acquire language by understanding messages. More precisely, comprehensible input is the essential environmental ingredient – a richly specified internal language acquisition device also makes a significant contribution to language acquisition. I argue that the best hypothesis is that competence in spelling and vocabulary is most efficiently attained by comprehensible input in the form of reading, a position argued by several others. (p. 440)

On the whole, the spoken and written forms of language are interconnected and mutually assist each other's development. Figure 2-4 presents how spoken and written forms of language are linked with one another. This leads us to the conclusion that both written and spoken forms should be dealt with in vocabulary learning simultaneously. In other words, it is essential to get balanced development between written and spoken forms, which facilitates effective growth of the Cube.

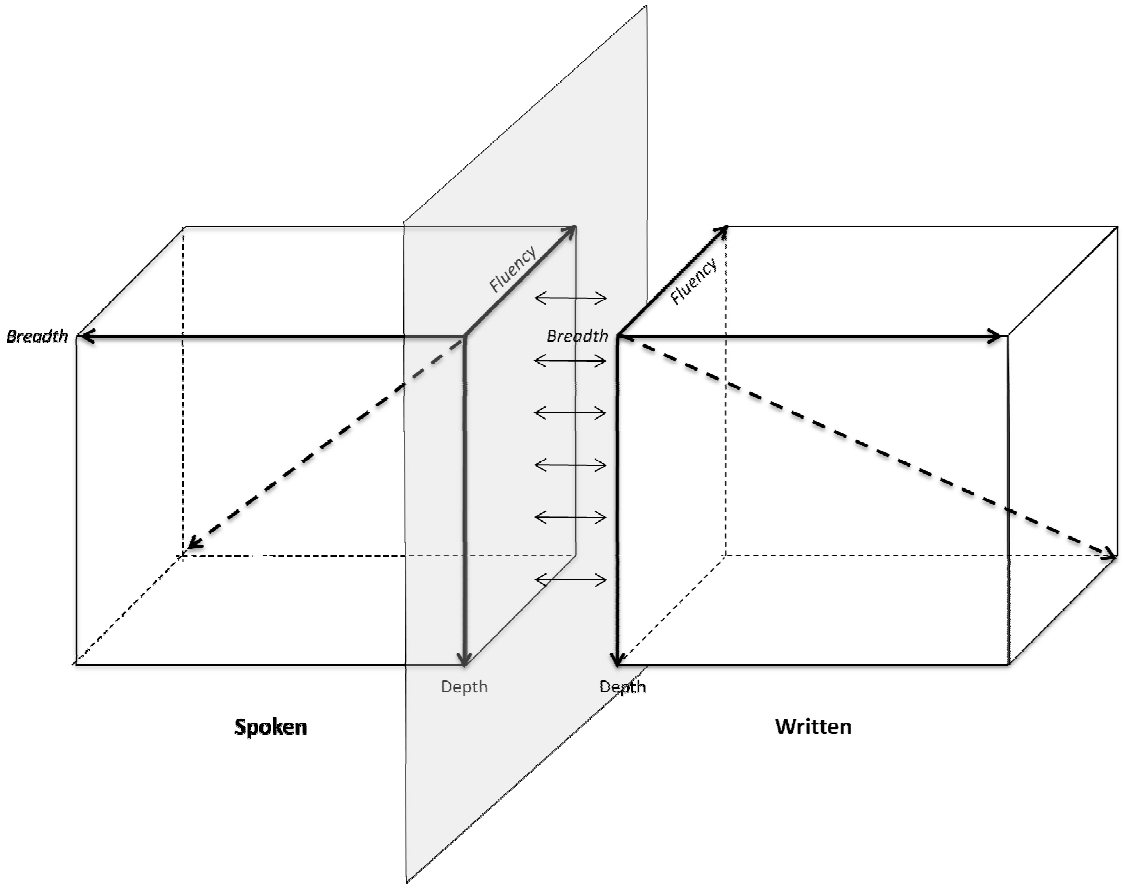

Figure 2-4 Mirror Image of Spoken and Written Forms of Lexical Proficiency

2.1.4.4 Summary: New model of lexical proficiency

In summary, the Cube, the new model of lexical proficiency, is developed synthesizing past theories and models of lexical proficiency. The Cube is three dimensional, operating with three axes of breadth, depth, and fluency as well as three sides of reception, variety in production, and sophistication in production. The three axes are reflect the model of lexical space by Daller et al. (2007). Including the model of Daller et al. (2007), the past theories addressing breadth and depth have shown the ambiguous cut between breadth and depth. However, in the Cube, the two dimensions of Breadth and Depth suggest clear distinction, importing Nation's model (2001) of lexical

proficiency. In addition to the three axes, the Cube contains three sides: reception, variety in production, and variety in sophistication. Reception is made with two axes of breadth and depth. That is, the axis of fluency is not considered in the side of reception. The second side, variety in production, is comprised of breadth and fluency, while the third side, sophistication in production, is made with depth and fluency. The written and spoken forms are separately developed in different spaces even though they affect each other positively.

2.2 Measures of Lexical Proficiency

There are two types of measures of lexical proficiency: lexical diversity and the Lexical Frequency Profile. Lexical diversity is a category of sets of measures while the Lexical Frequency Profileis a measure by itself.

2.2.1 Lexical Diversity

Lexical diversity refers to the variety of words that one uses in a productive language. Diversity is related to "the range of vocabulary and avoidance of repetition" (Malvern, Richards, Chipere, & Duran, 2004, p. 3). That is, lexical diversity deals with the size of vocabulary, depending on how many different words one produces in a corpus of a text. That is, the less often one repeats the same word and the more varied the words are, the higher the lexical diversity. Many researchers have used the concept of lexical diversity for their research, but they have often given different labels to indicate lexical diversity. For example, Engber (1995) and Read (2000) use "lexical variation" to mean lexical diversity in their studies. Laufer (2003) uses lexical diversity in his study, even

though he refers to it as “a combination of lexical variation and lexical sophistication” (p. 24).

Unlike other measures, lexical diversity has evolved along with many different measures, in order to overcome the limitation and pitfalls of single conceptualization. According to Malvern et al. (2004), the methods that have been used to measure lexical diversity can be separated into traditional approaches and mathematical approaches. The traditional approaches include two very basic measures of 1) types and 2) type-token ratio (TTR) and the other transformations of TTR (Malvern et al., 2004). However, traditional TTR measures of lexical diversity are seriously affected by sample size. To overcome the fundamental flaw of traditional measures of lexical diversity, a mathematical model of lexical diversity has been developed.

2.2.1.1 Tokens

The examination of tokens is one of the most popular and traditional ways to measure lexical diversity. A token is the total number of words, obtained by simply counting the numbers of words used in a text of productive language. Tokens represent how many words are produced in a certain amount of time. This measure is very simple and easy to use, given that no complex equations are needed. Tokens, however, cannot capture the details of lexical variety even though the number of words does present the absolute number of words produced. For example, it is theoretically possible that a person can have a high number of tokens simply by repeating only a few words. Therefore, the usefulness of tokens is limited; it is a good way to show the quantity of vocabulary use, but not robust enough for examining the quality of vocabulary use.

2.2.1.2 Types

Another traditional measure of lexical diversity is the examination of types, which involves counting how many unique words are used in a corpus of productive language. Typesis also called the number of different words (NDW) (Malvern et al., 2004). Malvern et al. (2004) argue that the examination of types is "the simplest measure available and clearly addresses an important aspect of diversity, namely the range of vocabulary deployed" (p. 16). That is, types is very easy to use for measuring lexical diversity. However, the result of types could be different, depending on the sample size. Malvern et al. (2004) suggest standardization of transcript length to overcome the limitation of NDW, as "Raw types has its use as a measure of the range of vocabulary in a language sample, but is limited when comparing the lexical diversity of different samples by being dependent on their size" (p.19). Overall, types is a good measure because of its convenience and ease of application, but it has limitations in that it is dependent on sample size.

2.2.1.3 Type-Token Ratio (TTR)

Another traditional measure of lexical diversity is Type-Token Ratio (TTR). TTR measures the proportion of the number of different words to the number of total words in a text. The number of whole words is called a token, while the number of different words is referred to as type. TTR can be calculated as the number of different words divided by the number of whole words in a text. Below is the equation for calculating TTR:

$$\text{TTR} = \frac{\text{type}}{\text{token}}$$

Because TTR is basically a proportion, it has values between 0 and 1. The value of 1 in TTR refers to the perfect diversity in lexical items used in productive language, while the value of 0 means no diversity in vocabulary. Therefore, the higher the TTR, the greater the lexical diversity.

TTR derives its strengths from two factors: ease of application and consideration of the sample size. First of all, TTR is easy to use. Calculation of TTR takes only three steps: 1) counting the number of whole words, 2) counting the number of different words, and 3) dividing the number of different words by the number of whole words. This is easy to apply, especially when it is considered that the estimate D(Duran, Malvern, Richards, & Chipere, 2004) is nearly impossible to calculate without the support of a calculator or special software programs. Secondly, TTR considers the size of the language sample, which makes TTR a better measure than types. Malvern et al. (2004) argued that TTR is more robust than types, which is expressed as NDW below:

> The assumption appears to be that because the number of different words is expressed as a proportion of the total number of words, the size of the language sample is therefore taken into account, and TTR will provide a more robust indication of lexical diversity than NDW. (p. 19)

In other words, TTR has its advantage in that TTR is calculated with the consideration of the number of whole words, while types does not take the number of whole words into consideration.

Due to these strengths of TTR, TTR has been used for many studies especially in the field of child language development and language impairment. For example, Templin (1957) conducted a cross-sectional study of the phonology, utterance length and

complexity, and vocabulary of 480 children at eight age points between three and eight years. In her study, she did not use the term of TTR, but reported types and tokens and introduced the concept of TTR by expressing "the proportion of different words used to all words" (p. 115). As a follow-up study, Miller (1981) computed mean TTRs using Templin's data and concluded "the consistency of this measure makes it enormously valuable as clinical tool" (p. 41).

While TTR has the strengths described above, it also has its limitations. First, TTR does not provide a consistent value for the different length of a given text. This is because it is greatly affected by sample size, as TTR is calculated with the number of whole words as its denominator. That is, depending on the length of a given text, the value of TTR would be different. This makes it hard to compare these values with different sample sizes. For example, let us say that two of the same values of TTR are given from two different samples with different lengths of text: the TTR is the same (0.6) for the two samples, but one is 60/100 from a child, while the other is 120/200 from an adult. In this case, it is difficult to judge who has better lexical diversity.

TTR has a second weakness in that it tends to gradually decrease as the sample size becomes larger. This is caused by TTR's inconsistency against the sample size. Simply speaking, it is inevitable that one repeats the same lexical items as a text gets longer. As shown in Figure 2-5, TTR drops down along with the number of tokens. In particular, TTR sharply decreases with the number of tokens from 0 to around 150, but then flattens out after the number of tokens becomes larger than 150. This indicates that the measure tends to lose sensitivity and discrimination as the sample size (the number of words) increases. Consequently, for children, TTR increases with developmental

transitions, i.e., TTR and production develop together. However, TTR does not discriminate for language learners who are able to produce 150 words in a given time period and those who have acquired a threshold with respect to fluency and production. In other words, it is hard to discriminate lexical diversity with the value of TTR in large sample sizes, especially in cases of those larger than 150 words.

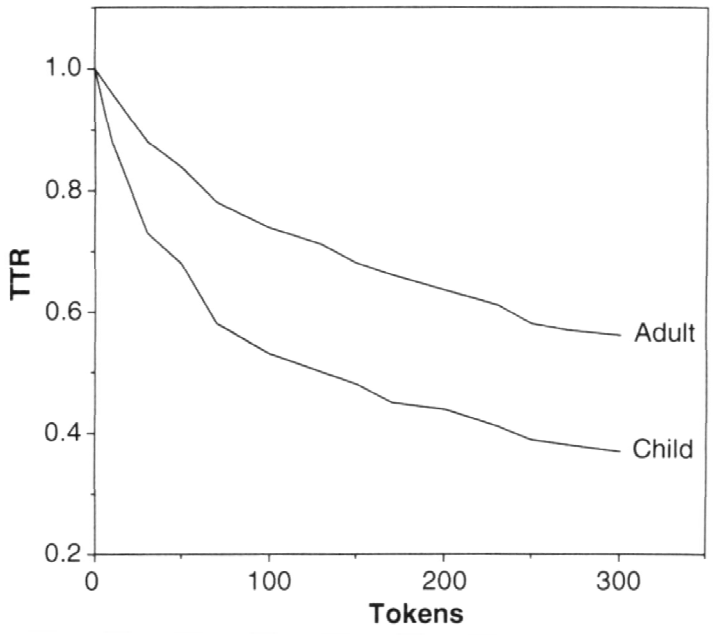

Figure 2-5 TTR Plotted Against Tokens for the Spoken Language of a Two-year-old and Academic Writing for an Adult (Malvern et al., 2004, p. 23)

The third weak point of TTR is that it does not provide any information about the repetition of particular words, for which Malvern et al. (2004) give the following explanation:

> As a simple illustration of the point, we can consider three imaginary texts of 40 tokens containing 20 types. In Text A, each of the 10 types is repeated twice; in Text B ten of the 20 types appear three times and the remaining ten once; Text C

> is more natural, with ten types occurring once, four twice, three occurring three times, two four times, and one type occurring five times. Raw TTR will not distinguish among these three transcripts as they all have the same overall numbers of types and of tokens producing a TTR of 0.5, but they are manifestly different in how they deploy the same vocabulary – they have different frequency distributions. (p. 31)

2.2.1.4 The Measurement D

As discussed in the previous section, the traditional TTR does not allow researchers to discriminate depending on the sample size. In addition, lexical diversity from the traditional TTR does not tell us how words are deployed in a text or the frequency of each word's occurrence. Thus, the measurement D was developed to address these problems.

First, to understand measurement D, one needs to figure out mathematical expression for ideal curves. This mathematical expression is a step used to generate the ideal curves with TTR as a function of tokens (Malvern et al., 2004). The equation used for generating the ideal curves is:

$$\mathrm{TTR} = \frac{\mathcal{D}}{\mathrm{N}}\left[\left(1 + 2\frac{\mathrm{N}}{\mathcal{D}}\right)^{\frac{1}{2}} - 1\right]$$

where N is the number of tokens and $\mathcal{D}$ is the estimate of constancy. The estimate $\mathcal{D}$ is distinguished from the measurement D, as "$\mathcal{D}$ is a particular value for best-fit between the ideal curves and those derived from real transcripts over the standard range of points of the TTR versus N curve drawn by a standardized procedure" (Malvern et al., 2004, p. 59).

Figure 2-6 shows the generated ideal curves based on the above equation. Each graph represents one sample. All plotted lines are placed between two extremes: TTR = 1 and TTR = 1/N. For the case of TTR = 1, the value for TTR does not change as a function of the number of tokens, which implies perfect lexical diversity. The graph of TTR=1/N refers to the case of least lexical diversity in a given sample. When we plot all the theoretically possible cases, it will look like Figure 2-6. The lexical diversity defined in Figure 2-6 is: "The nearer the graph of a language sample is to the line TTR = 1, the greater its lexical diversity; the nearer to the curve TTR = 1/N, the lower the lexical diversity" (Malvern et al., 2004, p. 48).

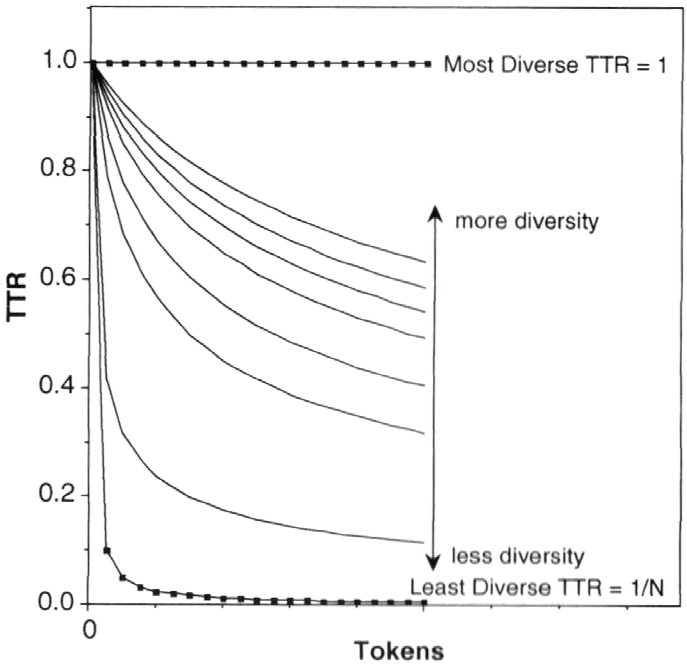

Figure 2-6 A Mathematical Model of Lexical Diversity

(Malvern et al., 2004, p. 48)

After generating the ideal curves with the above equation, the next step is producing the graphs from real data. According to Malvern et al. (2004), the two things needed to produce the graphs from real data are 1) "averaging a number of sub-samples of size 50 tokens drawn from throughout the whole sample" (p. 52) and 2) sampling without replacement, which is defined as "On selection the token is removed from the whole sample and is not available for reselection in the same sub-sample." (p. 53) The reason to obtain sub-samples of the size of 50 tokens is not only because the sub-samples of the size of 50 tokens reflect probabilistic theories but also because the number of 50 in token size is the point where the line TTR against tokens flattens (Malvern et al., 2004). Moreover, sampling without replacement is done when the graphs of TTR are being produced, as this method prevents "over-repetition" (p. 54) as well as "preserves the structure of the language" (p.53). When producing the graphs from real data is finished, the next step is standardizing. According to Malvern et al. (2004), there are a number of things that should be standardized:

> The values of N for which the curve is plotted; the number of trials of sub-samples used for the average at each point on the curve; and the method of random selection should all be fixed, so that every researcher makes the measurement in the same way. (p. 54)

The last step, which comes after standardizing, is to produce the measurement D (Figure 2-7). The most important part in this step is to find the best fit where the ideal curves best match the curves produced from the real data. Malvern et al. (2004) explain as below:

The program then needs to find the best fit between the ideal curves of theory and the curves drawn from empirical data by a curve-fitting procedure which adjusts the value of the parameter ($\mathcal{D}$) in the equation of ideal curves until a match is obtained between the actual curve for the transcript and the closest member of the family of curves represented by the mathematical model. This value of the parameter for best fit, $\mathcal{D}_{\text{best fit}}$ = D, is the index of the lexical diversity. (p. 56)

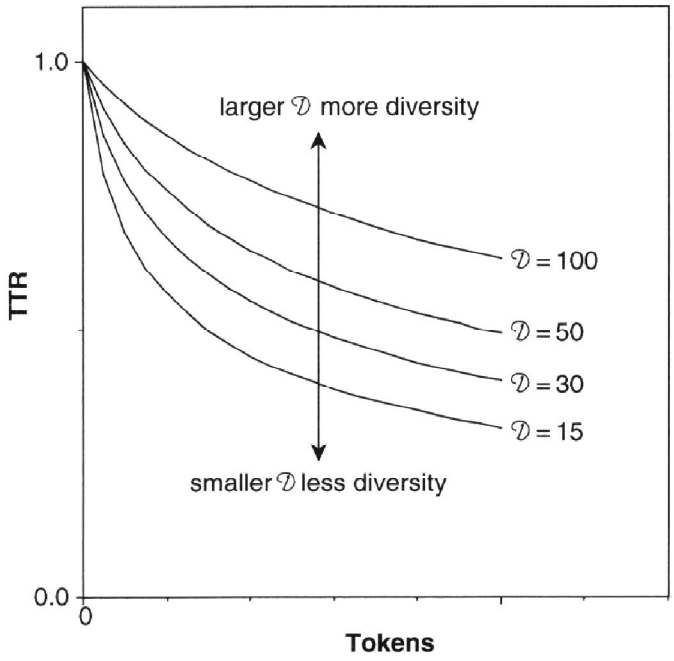

Figure 2-7 Ideal TTR Versus Token Curves Showing Increasing Diversity with Increasing D (Malvern et al., 2004, p. 52)

That is, when the best fit is found between the ideal curves and the curves from real data, the value of the measurement D is generated. The measurement D is interpreted as the higher the value of D, the greater the lexical diversity, and vice versa. Malvern et al. (2004) argue that "D provides a robust measure of lexical diversity which is not a function of sample size in the way raw TTR and its simple transformations are" (p. 60).

2.2.2 Lexical Frequency Profile

Another approach in response to the aforementioned limitations with the measures of lexical diversity was produced by Laufer and Nation (1995) who developed the Lexical Frequency Profile (LFP). The LFP was developed following the realization that all previous measures were neither consistent nor very objective. For example, type/token ratio is influenced by sample size in that it shows different values depending on the length of a sample (i.e., the longer the sample, the lower the TTR). A possible solution could be the use of a measure of lexical sophistication, which is the percentage of advanced words in a corpus of language, but it is not obvious what those "advanced" words are, and thus their selection would depend largely on the researcher's subjective assessment.

Given these drawbacks, Laufer and Nation (1995) developed the LFP to evaluate lexical proficiency in a completely new way that overcomes the pitfalls of previous measures. In other words, the LFP fundamentally has a different character from other measures in that the LFP makes a comparison between the actual word produced in a sample and a list of words. As Laufer and Nation (1995) explain, "The LFP shows the percentage of words a learner uses at different vocabulary frequency levels in one's writing – or put differently, the relative proportion of words from different frequency levels"(p. 311). In the LFP, there are four different word lists, each of which captures words at different levels of frequency: the first 1,000 most frequent words, the second most frequent 1,000 words, academic vocabulary, and less frequent words which do not appear in the first three categories. Laufer and Nation (1995) showed that the use of LFP comparisons produces a measure that is strongly correlated to the overall quality of

writing and discriminates different levels of writing as well as different levels of language proficiency. Therefore, the more advanced one's language competence, the greater use of less frequent words.

In summary, there are two popular measures of lexical proficiency: lexical diversity and the Lexical Frequency Profile. The first, lexical diversity, which indicates how various vocabulary words are used in a text, has developed in various ways: 1) tokens, the number of total words, and 2) types, the number of different words, are the very basic types of lexical diversity. Another traditional measure of lexical diversity is Type-Token Ratio (TTR), which is useful, easy to calculate, and better than types since, as a proportion, TTR respects the number of whole words. However, it is not stable, because the value of TTR depends on the sample size. In order to compensate for this flaw, the estimate D, a mathematical estimate of lexical diversity, was introduced by Malvern et al. (2004). D uses the ideal curves with the best fit of $\mathcal{D}$. The advantages of the measurement D is that D is not affected by the sample size as D uses random sampling without replacement and standardizing values of TTR. The second, the Lexical Frequency Profile developed by Laufer and Nation (1995), provides four levels of vocabulary word lists: the first 1,000 most frequent words, the second 1,000 most frequent words, the academic vocabulary, and the less frequent words which do not appear in the first three categories. The value of lexical proficiency is calculated by comparing a given text with these word lists. These measures of lexical proficiency all have different features and purposes. Therefore, it is essential to know the exact features of each measure, as well as its strong and weak points.

2.3 Relationship between Lexical Proficiency and L2 Oral Proficiency

Lexical proficiency is strongly correlated with productive language proficiency. For instance, lexical variation is highly related with the quality of L2 writing (Engber, 1995). The length of essay and lexical diversity are strong indicators of the quality of L2 writing (Mellor, 2011). L2 learners at different levels of proficiency have different lexical features: the more advanced in L2 proficiency, the more lexical variety appears in productive language (Ferris, 1994; Crossley, Salsbury, & McNamara, 2011). Additionally, more proficient L2 learners use more difficult and sophisticated words than less proficient learners in L2 writing (Laufer& Nation, 1995; Goodfellow, Lamy, & Jones, 2002). These studies show that the way that one uses vocabulary items is linked with L2 proficiency in production.

Ferris (1994) identifies the lexical characteristics that can be found in ESL students' writings by analyzing a corpus of 160 essays that were composed by students from four different L1 groups: Arabic, Chinese (Mandarin), Japanese, and Spanish. Three different graders rated each essay using a scale of 0 – 10, and the sum of these three ratings were calculated. After the rating process, the participants were divided into two groups based on their essay scores: an advanced group and a lower group. The differences in lexical and syntactic features were then compared between these two levels. The results indicate that the advanced group used a higher level of lexical and syntactic features than the lower group. That is, participants in the advanced group knew more vocabulary words than the ones in the lower group. Also, it was found that the advanced L2 learners were better able to associate the words with other lexical items, using synonymy and antonymy. The second important result of this study indicates that there is

a strong relationship between the holistic score for writing and other variables, such as number of words, synonymy/antonymy, or word length. In other words, the longer the essay is, the higher the score. Also, participants who have a higher score in writing tend to use more synonyms and antonyms as well as longer words.

Engber (1995) examined how lexical proficiency is related to the quality of L2 writing. In this study, 66 essays written by students of the Intensive English Program (IEP) at Indiana University were scored on a 6-point scale. Then, the relationship between the holistic scores and lexical richness measures were examined in terms of lexical variation, error-free variation, percentage of lexical error, and lexical density. The results indicate that lexical variation both with error ($p = .45$) and without error ($p = . 57$) is significantly related to the quality of L2 writing, with lexical error negatively correlated to the quality of L2 writing. That is, the fewer lexical errors that take place, the better the writing will be judged.

Goodfellow, Lamy, and Jones (2002) investigated the possibility of using the Lexical Frequency Profile for automatic writing feedback systems. To that end, 36 students' essays from an Open University French course were collected and evaluated. The participants were also administered a vocabulary test for comparison between the test results and lexical proficiency in L2 writing. To examine lexical proficiency in writing samples, the Lexical Frequency Profile (LFP; Laufer and Nation, 1995) was used as an adopted version for measuring participants' lexical proficiency in L2 writing. To adapt LFP to French, the authors generated three lists: the first 1,000 most frequent words, the second 1,000 most frequent, and the Academic Word List. Results showed that there is a strong relationship between one's lexical proficiency and the quality of writing in a

second language. Also, it turned out that one's lexical proficiency in L2 writing is significantly related to the vocabulary test results. The authors suggested that an automatic feedback system using the LFP would be useful for L2 writing students in the LFP provides feedback based on an existing list to which learners can refer in order to self-check their own level of lexical proficiency.

Mellor (2011) examined the relationship among essay length, lexical diversity, and the quality of essay. This study had two purposes: 1) to determine if a two-dimensional quantity/content model employing essay length and lexical diversity can predict human assessment of essays better than a single dimension of either quantity or content, and 2) to determine which measure of lexical diversity works best as the content dimension alongside the quantity dimension of essay length for this set of essays. Thirty-four college students who were third year English majors at a Japanese university participated in this study. They were given a prompt and asked to write an essay in 30 minutes. One native speaker rated all 34 essays as *good, above average, average, below average,* or *poor*. Each essay was then analyzed on two dimensions, quantity and content, each of which is typified as essay length in words and lexical diversity. Six measures were adopted for this study: TTR(100), Guiraud Index, Yule's K, the D estimate, Hapax (100), and estimate of Advanced Guiraud. The results indicate:

> Lexical diversity together with essay length can more accurately predict essay ratings than either feature alone with this set of essays…..Essay length is a very strong predictor as a single dimension….TTR(100), Yule's K, Hapax(100), and Advanced Guiraud perform similarly well in correlation and regression analyses. (P. 35-40)

2.4 Gaps Found in the Previous Studies

At least three limitations or gaps can be found in the previous studies listed above. First, many of these studies were conducted with the initial measures of lexical proficiency because more advanced measures did not exist at that time. Therefore, there is a need to revisit the research questions addressed using more recent and updated versions of lexical proficiency measures. Second, few studies actually examine the relationship among measures of lexical proficiency. Most recent studies that were conducted on the relationship between lexical proficiency and L2 productive language employ either the measures of lexical diversity or the Lexical Frequency Profile, but they do not examine the correlation between the measures of lexical diversity and the Lexical Frequency Profile. Third, most studies examining the relationship between lexical proficiency and L2 productive language proficiency have focused on L2 writing, while few have examined whether lexical proficiency can be correlated with L2 oral proficiency.

As an example of recent research on the relationship between lexical proficiency and L2 productive language, Yu (2009) explains how lexical diversity is related to the scores on writing and speaking proficiency exams. In detail, Yu examines whether there is a relationship between the quality of writing and a writing score, if the quality of writing is associated with the writing topics, whether the quality of writing is related to the participants' overall language competence, and whether there is any relationship between written discourse and spoken discourse in terms of quality. In this study, 200 compositions and 25 interviews were selected and rated by two raters. Estimate D was used for measuring lexical diversity of selected compositions. Yu (2009) concludes two

main points that 1) there is a significant relationship between estimate D and the quality of both L2 writing and speaking, and 2) the topic of one's essay can affect the outcome in terms of lexical proficiency, as well as the quality of writing (in other words, the more familiar the essay topic, the higher the lexical proficiency and the better the quality of the writing). Although Yu's (2009) work suggests interesting findings about the relationship between lexical proficiency and L2 productive language, it does not present any correlation between lexical diversity and the Lexical Frequency Profile.

2.5 Research Questions

Based on the gaps found in the previous studies, this study poses three related research questions:

1) How are different measures of lexical proficiency correlated with the holistic scores of the OEPT?
2) How are different measures of lexical proficiency correlated to each other? How are the measures of lexical diversity correlated to the Lexical Frequency Profile?
3) Would examinees with different L1 backgrounds show dissimilar patterns in measures of lexical proficiency in L2?

CHAPTER 3. METHODOLOGY

3.1 The Oral English Proficiency Test (OEPT)

In order to investigate the research questions proposed in the previous chapter, the data of the Oral English Proficiency Test (OEPT) were employed and analyzed in this study. The OEPT is a computer-based test operated by the Oral English Proficiency Program (OEPP) at Purdue University. The OEPT is used to screen prospective teaching assistants whose first language is not English. The examinees' impromptu responses to 12 questions are recorded and rated by trained raters. To be certified for becoming a teaching assistant, examinees are required to get a score of 50, 55 or 60. The internal consistency coefficients are very high across the eight items (0.96 to 0.98). Among the 12 items given to the OEPT examinees, the Compare and Contrast item, in which the examinees are asked to make a comparison between two different topics, was analyzed in this study.

During the entire course of the OEPT text, a narration is provided as the items are presented on the screen. The instructions are given on the first screen of each individual item. When the narration is finished, the examinees proceed to the next screen by clicking a button. Then, the actual item is presented on the next screen, accompanied by narration. When the narration comes to the end, the preparation time automatically starts, and a

countdown clock is presented on the screen. The examinees can click the "Record Now" button when they are ready or, if they spend the entire time preparing, the recording automatically begins. When the examinees finish answering the item, they can stop recording by clicking the "Finish" button. If they do not complete the response before time runs out, the recording automatically stops. All items are presented in a fixed sequence. The examinees manage the amount of time spent preparing for each item within the given time limit.

3.2 Rating

The examinees' responses to the items, which are recorded during the test, are assigned to at least two raters who are trained to rate on the OEPT scale. The scale ranges from 35 to 60. The examinees who are given 35, 40, or 45 on the OEPT fail to pass and are assigned to an instructional course in the Oral English Proficiency Program. On the other hand, the examinees who pass the test with 50, 55, or 60 are deemed ready to be assigned to a teaching assistant position. The inter-rater reliability ranges from 0.76 to 0.84.

3.3 The Sample

Roughly 500 examinees, with 40 different L1 backgrounds, take the OEPT every year. The main population of examinees included speakers of Mandarin (30%), Korean (15%), Hindi (10%), Spanish (5%), Marathi (2.5%), Bengali (2.5%), Telugu (2.5%), and Russian (2.5%). The other language groups not included represent less than 2.5% of the population annually. For the present study, Mandarin, Korean, and Hindi, which make up the three largest sub groups, have been selected. Table 3-1 presents the details of

examinees: 303 samples were selected as a total with 100 Korean, 111 Mandarin, 67 Hindi, and 25 English L1 Native.

Table 3-1 The Sample

OEPT score / L2 groups	35	40	45	50	55	60	Sub total
Korean	25	25	25	25	-	-	100
Mandarin	22	25	25	25	14	-	111
Hindi	-	-	-	25	5	17	67
English	25						25
Total							303

It should be noted that each group obtained a different range of the OEPT scores, not corresponding to the entire score range: Korean scores range from 35 to 50, Mandarin 35 to 55, and Hindi 50 to 60. The group differences are considered a limitation of this study but are characteristic of the examinee population at Purdue. For comparison, the English L1 group with 25 samples was recruited from the OEPT staff, instructors, and professors.

3.4 The Measures of Lexical Proficiency

Table 3-2 presents the measures of lexical proficiency employed in this study to analyze vocabulary use of L2 examinees. In this study, two dimensions of lexical proficiency were examined: variety in spoken production and sophistication in spoken production. In order to measure each dimension, traditional measures were selected: the

measures of lexical diversity and the Lexical Frequency Profile. The measures of lexical diversity were selected to measure variety in spoken production, while the LFP was chosen to measure sophistication in spoken production. This is mainly because each measure of lexical diversity and the LFP reflects the characteristics of each lexical proficiency dimension: the dimension of variety in production quantitatively deals with lexical proficiency, and the measures of lexical diversity do so as well. The LFP shows information regarding word frequency, which is the level of easiness of each word; consequently, the LFP is considered closer to the dimension of sophistication in production rather than variety in production. However, it is hard to conclude that these measures selected for the current study best represent each dimension of lexical proficiency. This is not only because each measure contains its own drawbacks but also because lexical diversity is partly overlapped with the LFP: both lexical diversity and the LFP are based on counting lexical items. This fact is considered a limitation of this study.

For the measure of lexical diversity, four traditional measures were selected: tokens, types, type-token ratio (TTR), and D. The estimation of tokens counts the total number of words produced in an examinee's speech, while the estimation of types counts the number of different words used. TTR refers to types divided by tokens, which implies the ratio of unique words to the number of total words. The estimation of D is the mathematical model of TTR, calculated as a random-sampling of tokens from the original transcription and standardizing TTR.

The Lexical Frequency Profile was measured with K1 types, K2 types, AWL types, and Off-list types. The estimation of K1 types was done by counting the number of types which overlap with the list of the first 1,000 most frequent words, while the

estimation of K2 types is the number of types overlapping with the list of the second 1,000 most frequent words. The estimation of AWL types was done by counting types which fall into the Academic Word List, and the estimation of Off-list types indicates the number of types which are not included in the above three categories.

Table 3-2 Variables: Measures of Lexical Proficiency

Lexical Proficiency	Measures		How to calculate
Variety in Spoken Production	Lexical diversity	Tokens	Total number of words
		Types	Number of unique words
		Type-token ratio	Types / Tokens
		D	$\mathrm{TTR} = \frac{D}{\mathrm{N}}\left[\left(1 + 2\frac{\mathrm{N}}{D}\right)^{\frac{1}{2}} - 1\right]$
Sophistication in Spoken Production	Lexical Frequency Profile	K1 types (K1Typ)	Number of words used in the first 1000 most frequent words
		K2 types (K2Typ)	Number of words used in the second 1000 most frequent words
		AWL types (AWLTyp)	Number of types included in the Academic Word List
		Off-list types (OffTyp)	Number of types not included in three lists above

It should be noted that in this study, the LFP was calculated by counting the actual number of types, not by the percentage in each word list. The main reason is that, as one of the research purposes, I intended to examine the actual change of vocabulary size in four different levels of word frequency through oral proficiency development. The proportion does not provide the exact information of the concrete vocabulary size.

Two software programs, VocabProfile 3.0 and Coh-Metrix 3.0, were used for vocabulary analyses. VocabProfile 3.0 is a program that analyzes lexical characteristics of text based on the four levels of word lists by frequency. Coh-Metrix 3.0 is another computational program used to analyze the cohesion of a text using linguistic and discourse indices. Both programs are accessible online for free. In this study, the Lexical Frequency Profile was examined by using VocabProfile 3.0, while the measures of lexical diversity were analyzed with Coh-Metrix 3.0.

3.5 Statistical Procedure

Two statistical procedures were included: Spearman rank order correlation and descriptive statistics. First of all, the correlation coefficient among the OEPT scores and the variables of lexical proficiency were calculated through Spearman rank order correlations. This is because the major statistical focus in this step was on the individual correlations among all variables, answering the first and second research questions. Specifically, the Spearman rank order was selected in this step, as the OEPT scores are ordinal. Secondly, the descriptive statistics on eight variables of lexical proficiency were employed to observe the difference between different L2 groups as well as between each sub OEPT score groups. The values of the descriptive statistics included mean, standard deviation, standard error, range, 99% confidence interval, degrees of freedom, t-value, and p-value.

3.6 The change in the numbers of samples

There were some changes in the numbers of samples while the Spearman rank order correlation coefficient was run because the value of 0 was found on the

measurement of D. Technically, the value of 0 in D (D0s) indicates that there is lexical diversity. However, this is impossible because if someone speaks even only one sentence, there should exist tokens and types, which means that lexical diversity can be calculated. In this case, D0s were produced as VocD makes 0 of D when the number of tokens is below 50 or even at 50 – 100 (Malvern at al., 2007). This is considered a flaw of D, and the issue of D0s will be discussed in detail in the discussion section.

For the first time running the analysis, the initial number of samples was 278 including 100 Korean, 111 Mandarin, and 67 Hindi. However, the samples with the values of 0 in the measurement of D (D0s) were found for some OEPT sub groups: 13 of D0s were found in the group of Korean 35 (The Korean group with 35 of the OEPT score), 3 in Korean 40, 1 in Korean 50, 3 in Hindi 50, and 3 in Hindi 55. After D0s were found, the samples which contained D0s were excluded as 0 of D is not a proper value to represent lexical diversity. That is, if a sample included 0 of D, the other values of lexical proficiency measurements were excluded in running correlation coefficients as the number of samples were equivalent through the variables.

As a result, 255 samples remained in the statistical analysis: 83 Korean, 111 Mandarin, and 61 Hindi. Table 3-3presents the details: the numbers in parentheses are the initial number of samples while the numbers without parentheses are the final numbers after all examines with D0s were removed. However, for running the descriptive data analysis, there was no change in the number of examinees except for the measurement of D: for D, all D0s were removed as above.

The native speakers of English were not included in running the Spearman rank order correlation coefficient, as the OEPT is originally designed for international students whose first language is not English.

Table 3-3 The Changed Samples After Removing D0s

OEPT score / L2 groups	35	40	45	50	55	60	Sub total
Korean	12 (25)	22 (25)	25	24 (25)	-	-	83 (100)
Mandarin	22	25	25	25	14	-	111
Hindi	-	-	-	22 (25)	22 (25)	17	61 (67)
English	21 (25)						21 (25)
Total							276 (303)

CHAPTER 4. RESULTS

4.1 Research Question 1

Q1. How are different measures of lexical proficiency correlated with the holistic scores of the OEPT?

The Spearman correlation matrix in Table 4-1 demonstrates the correlation coefficients among the OEPT scores and the variables of lexical proficiency. The notable findings are as follows: there were moderately strong positive correlations for the OEPT scores and types ($r_{OEPT,\ Types} = 0.32$), D ($r_{OEPT,\ D} = 0.35$), and AWL types ($r_{OEPT,\ AWLTyp} = 0.39$). Also, the OEPT scores had a moderate positive correlations with K1 types ($r_{OEPT,\ K1Typ} = 0.25$) and K2 types ($r_{OEPT,\ K2Typ} = 0.22$). However, the OEPT scores had a weak correlation with tokens ($r_{OEPT,\ Tokens} = 0.18$), TTR ($r_{OEPT,\ TTR} = 0.16$), and Off-list types ($r_{OEPT,Off\ Typ} = 0.05$).[1]

Of interest is that different results were obtained when the Spearman rank order was rerun without the Mandarin groups: as shown in Table 4-2, the correlation coefficients increased for 8 out of 9 variables. In detail, there were strong positive correlations for the OEPT scores and D ($r_{OEPT,\ D} = 0.46$) and AWL types ($r_{OEPT,\ AWLTyp} = 0.43$). Also, moderately strong positive correlations were on types ($r_{OEPT,\ Types} = 0.38$), K1

[1] K1 types (the 1st 1000 most frequent words), K2 types (the 2nd 1000 most frequent words), AWL types (Academic Word List), Off types (words not included in three lists above)

types ($r_{OEPT,\ K1Typ} = 0.33$) and K2 types ($r_{OEPT,\ K2Typ} = 0.33$) with the OEPT scores, while moderate correlation was found on TTR ($r_{OEPT,\ TTR} = 0.26$). However, the OEPT scores still remained to have a weak negative correlation with tokens ($r_{OEPT,\ Tokens} = 0.18$) and Off-list types ($r_{OEPT,\ Off\ Typ} = 0.07$).

Figure 4-1 clearly presents how r values changed through the selected variables when Mandarin was excluded. The r value of types increased from 0.32 to 0.38, TTR from 0.16 to 0.26, D from 0.35 to 0.46, K1 types from 0.25 to 0.33, K2 types from 0.25 to 0.33, and AWL types from 0.39 to 0.43. Types, D, and AWL types moved their area from moderately strong to strong whileK1 types and K2 types became moderately strong from moderate. On the other hand, there were little change of r values for tokens and Off-list types.

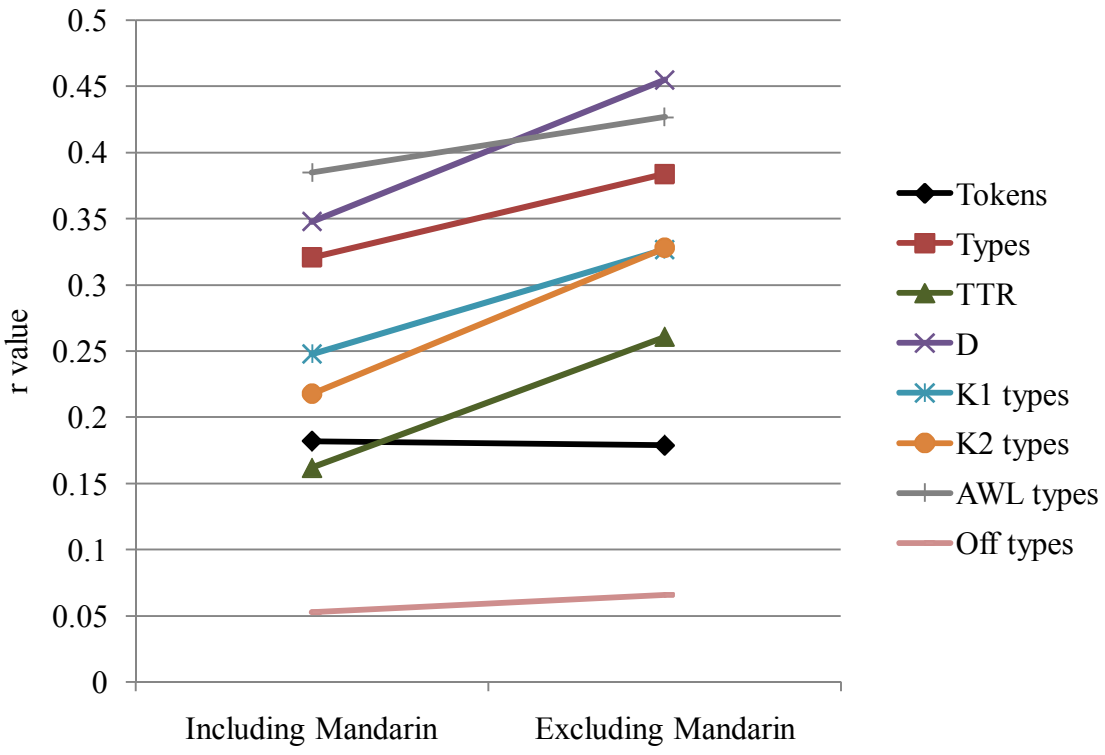

Figure 4-1 Change of *r* Values When Mandarin is Excluded

Table 4-1 1st Running of Spearman Rank Order Correlation Coefficient: Variables of Lexical Proficiency and OEPT scores Korean (n = 83), Mandarin (n = 111), Hindi (n = 61)

	1	2	3	4	5	6	7	8	9
1. OEPT	1.00								
2. Tokens	0.18*	1.00							
3. Types	0.32**	0.88**	1.00						
4. TTR	0.16*	-0.61**	-0.20*	1.00					
5. D	0.35**	0.17*	0.52**	0.51**	1.00				
6. K1Typ	0.25**	0.87**	0.94**	-0.27**	0.45**	1.00			
7. K2Typ	0.22**	0.41**	0.51**	-0.01	0.27**	0.38**	1.00		
8. AWLtyp	0.39**	0.31**	0.38**	-0.02	0.19*	0.24**	0.20*	1.00	
9. OffTyp	0.05	0.33**	0.44**	0.04	0.34**	0.27**	0.19*	0.05	1.00

*Note: * p< .05, ** p<.001*

Table 4-2 2nd Running of Spearman Rank Order Correlation Coefficient: Variables of Lexical Proficiency and OEPT scores Korean (n = 83), Mandarin (n = 0), Hindi (n = 61)

	1	2	3	4	5	6	7	8	9
1. OEPT	1.00								
2. Tokens	0.18*	1.00							
3. Types	0.38**	0.89**	1.00						
4. TTR	0.26*	-0.62**	-0.21*	1.00					
5. D	0.46**	0.15*	0.52**	0.55**	1.00				
6. K1Typ	0.33**	0.87**	0.95**	-0.26*	0.47**	1.00			
7. K2Typ	0.33**	0.42**	0.52**	-0.004	0.27*	0.38**	1.00		
8. AWLtyp	0.43**	0.40**	0.49**	-0.06	0.24*	0.34**	0.33**	1.00	
9. OffTyp	0.07	0.41**	0.51**	0.02	0.36**	0.34**	0.22*	0.13	1.00

*Note: * p< .05, ** p<.001*

4.2 Research Question 2

Q2. How are different measures of lexical proficiency correlated to each other? How are the measures of lexical diversity correlated to the Lexical Frequency Profile?

As found with the first research question, people who have a higher score on an L1 speech test tend to use more words and display higher lexical diversity. The question that follows, then, is if people with higher scores on the OEPT have better lexical diversity, what kind of words do they use more than others? Do they use easier and more frequent words? Or do they use more difficult words than speakers with lower OEPT scores? The answers for the above questions are laid on the relationship between lexical diversity and the Lexical Frequency Profile.

In order to answer the second research questions, Spearman rank order analysis was completed for all the L1 language groups: Korean, Mandarin, and Hindi. Analysis was done in three different cases: 1) correlations among the measures of lexical diversity, 2) correlations among the indices of the Lexical Frequency Profile, and 3) correlations between lexical diversity and the Lexical Frequency Profile. After the analyses were finished with the above three different types of correlations in lexical proficiency, an analysis to look into the details on the last part, correlations between lexical diversity and Lexical Frequency Profile, was needed. Accordingly, descriptive statistics were carried out for the Lexical Frequency Profile under each OEPT sub-score group as well.

4.2.1 Correlations Among the Measures of Lexical Diversity

Table 4-3 presents the correlation coefficients among the variables of lexical diversity, which is provided by the Spearman rank order correlation. A very strong positive correlation was found between tokens and types ($r_{Tokens, Types}$= 0.89). Strong positive correlations were found between D and types ($r_{D, Types}$ = 0.52) and between D and TTR ($r_{D, TTR}$ = 0.51). Also, a strong negative correlation was found between TTR and tokens ($r_{TTR, Tokens}$ = - 0.61). There is not a significant relationship between TTR and types ($r_{TTR, Types}$ = - 0.20) and between D and tokens ($r_{D, Tokens}$ = 0.17).

Table 4-3 Correlation Coefficients among the Measures of Lexical Diversity

	1	2	3	4
1. Tokens	1.00			
2. Types	0.88**	1.00		
3. TTR	-0.61**	-0.20*	1.00	
4. D	0.17*	0.52**	0.51**	1.00

*Note: * p< .05, ** p<.001*

4.2.2 Correlations Among the Measures of the Lexical Frequency Profile

As shown in Table 4-4, the Lexical Frequency Profile had weak correlations between each other in general, except for the correlation between K1 types and K2 types: K1 types had a moderately strong positive correlation with K2 Types ($r_{K1Typ, K2Typ}$ =

0.38). In other words, using more frequent words does not guarantee using more less-frequently-used words such as AWL types or Off-list words, and vice versa, even though there is a moderate correlation between using the first 1,000 frequent words and the second 1,000 words.

Table 4-4 Correlation Coefficients among the Measures of LFP

	1	2	3	4
1. K1Typ	1.00			
2. K2Typ	0.38**	1.00		
3. AWLtyp	0.24**	0.20*	1.00	
4. OffTyp	0.27**	0.19*	0.05	1.00

*Note: * p< .05, ** p<.001*

4.2.3 Correlations Between Lexical Diversity and the Lexical Frequency Profile

In terms of the second research questions, the relationships between lexical diversity and the Lexical Frequency Profile are the most interesting. As Table 4-5 presents, each variable of lexical diversity shows a different pattern of correlation with the Lexical Frequency Profile.

Tokens show strong or extremely strong correlations with the Lexical Frequency Profile. In terms of the correlations between tokens and the Lexical Frequency Profile, K1 types are strongly correlated with tokens ($r_{\text{Tokens, K1Typ}} = 0.87$). K2 types and tokens

have a moderately strong correlation ($r_{\text{Tokens, K2Typ}} = 0.42$). AWL types and Off-list types are moderately correlated with tokens ($r_{\text{Tokens, AWLtyp}} = 0.40$, $r_{\text{Tokens, OFfTyp}} = 0.41$).

Table 4-5 Correlation Coefficients between Lexical Diversity and LFP

		Lexical Diversity			
		Tokens	Types	TTR	D
Lexical Frequency Profile	K1Typ	0.87**	0.95**	-0.26*	0.47**
	K2Typ	0.42**	0.52**	-0.004	0.27*
	AWLtyp	0.40**	0.49**	-0.06	0.24*
	OffTyp	0.41**	0.51**	0.02	0.36**

*Note: * p< .05, ** p<.001*

Types builds a similar, but stronger pattern of correlation with the Lexical Frequency Profile than tokens does: K1 types have an extremely strong correlation with types ($r_{\text{Types, K1Typ}} = 0.95$). The other three variables of the Lexical Frequency Profile, which are K2 types, AWL types, and Off-list types, have a strong correlation with types ($r_{\text{Types, K2Typ}} = 0.52$, $r_{\text{Types, AWLTyp}} = 0.49$, $r_{\text{Types, OffTyp}} = 0.51$).

TTR and D are in weak or moderate correlation with the Lexical Frequency Profile, while tokens and types have strong correlations with the LFP. TTR shows weak negative correlations with all four variables of the Lexical Frequency Profile ($r_{\text{TTR, K1Typ}} = -0.26$).However, the correlations between TTR and the lexical frequency variable, except for K1 types, are excluded from analysis because of their high p-value over 0.05.

D and the Lexical Frequency Profile show all the positive correlations, which are around a strong or moderate area ($r_{D, K1Typ}$= 0.47, $r_{D, K2Typ}$= 0.27, $r_{D, AWLTyp}$= 0.24, $r_{D, OffTyp}$= 0.36).

The notable finding regarding the relationship between lexical diversity and the Lexical Frequency Profile is thatK1 types have an extraordinarily strong correlation with tokens and types as the correlations are almost close to 1 ($r_{Tokens, K1Typ}$= 0.87, $r_{Types, K1Typ}$ = 0.95). Also, K2 types have the second strongest correlation, among the variables of Lexical Frequency Profile, with tokens and types ($r_{Tokens, K2Typ}$= 0.42, $r_{Types, K2Typ}$ = 0.52). Comparatively, in terms of AWL types and Off-list types, the correlations with tokens and types are low, under 0.5 ($r_{Tokens, AWLtyp}$ = 0.40, $r_{Tokens, OFfTyp}$ = 0.49, $r_{Types, AWLTyp}$ = 0.41, $r_{Types, OffTyp}$ = 0.51).

Figure 4-2 shows the detailed relationship between types and the indices of the Lexical Frequency Profile in terms of how the values of the LFP indices change through OEPT scores: As OEPT scores increase from 35 to 60, the overall number of types also increases. However, when comparing each LFP index, it turned out that K1 types occupy the largest portion of types through all OEPT score groups. In other words, the number of types increases mostly in the area of K1 types while OEPT scores increase. Table 4-6 also presents more detailed information on the descriptive statistics of the LFP indices for each OEPT score group.

This fact implies that when OEPT test takers speak more words and have a higher score, the major change of tokens and types happen in the first 1,000 frequent words list rather than in less frequent words. In other words, people who are good at L2 speaking

make their speech better and longer than people with poor speaking by using more frequent words rather than less frequent words.

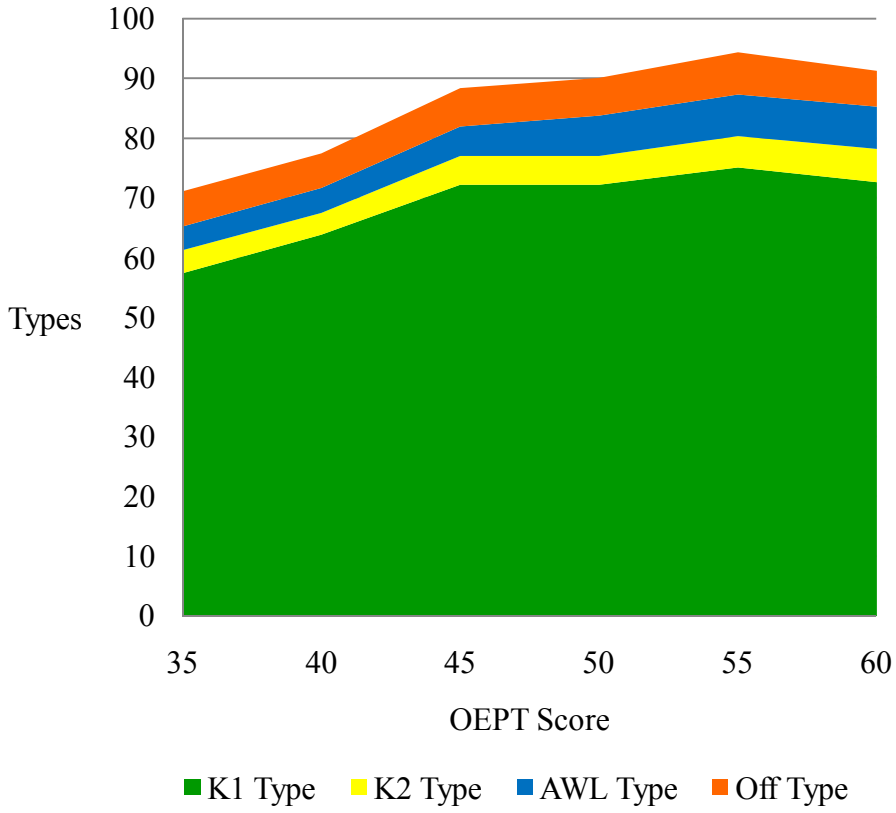

Figure 4-2 The Change of LFP Values through OEPT Scores

Table 4-6 Descriptive Statistics with LFP Indices among OEPT Score Groups

OEPT	LFP index	N	Mean	SD	Std Err	Min	Max	99% CL Mean		DF	t Value	Pr > |t|
35	K1 Type	47	57.45	17.78	2.59	27	96	50.48	64.42	46	22.15	<.0001
	K2 Type	47	3.87	2.37	0.35	0	10	2.95	4.80	46	11.23	<.0001
	AWL Type	47	3.98	2.12	0.31	0	10	3.15	4.81	46	12.86	<.0001
	Off Type	47	5.89	3.29	0.48	1	14	4.60	7.18	46	12.27	<.0001
40	K1 Type	50	63.88	13.39	1.89	36	91	58.80	68.96	49	33.73	<.0001
	K2 Type	50	3.60	1.91	0.27	0	8	2.88	4.32	49	13.36	<.0001
	AWL Type	50	4.26	2.0	0.28	1	9	3.50	5.02	49	15.08	<.0001
	Off Type	50	5.74	3.13	0.44	1	12	4.55	6.93	49	12.97	<.0001
45	K1 Type	50	72.24	13.30	1.88	45	105	67.20	77.28	49	38.42	<.0001
	K2 Type	50	4.80	2.40	0.34	1	13	3.89	5.71	49	14.15	<.0001
	AWL Type	50	4.94	2.45	0.35	1	13	4.01	5.87	49	14.24	<.0001
	Off Type	50	6.42	3.47	0.49	1	14	5.11	7.73	49	13.1	<.0001
50	K1 Type	75	72.25	17.92	2.07	22	108	66.78	77.73	74	34.91	<.0001
	K2 Type	75	4.84	2.46	0.28	1	12	4.09	5.59	74	17.07	<.0001
	AWL Type	75	6.68	3.28	0.38	1	16	5.68	7.68	74	17.64	<.0001
	Off Type	75	6.33	3.35	0.39	1	16	5.31	7.36	74	16.35	<.0001
55	K1 Type	39	75.15	17.95	2.87	33	108	67.36	82.95	38	26.15	<.0001
	K2 Type	39	5.26	2.11	0.34	1	10	4.34	6.17	38	15.55	<.0001
	AWL Type	39	6.92	3.26	0.52	1	14	5.51	8.34	38	13.28	<.0001
	Off Type	39	7.13	3.90	0.63	1	16	5.43	8.82	38	11.41	<.0001
60	K1 Type	17	72.71	17.98	4.36	50	112	59.97	85.44	16	16.67	<.0001
	K2 Type	17	5.53	2.13	0.52	2	11	4.02	7.04	16	10.73	<.0001
	AWL Type	17	7.06	3.01	0.73	3	14	4.93	9.19	16	9.67	<.0001
	Off Type	17	6.06	2.90	0.70	2	12	4.00	8.12	16	8.6	<.0001

4.3 Research Question 3

Q3. Would examinees with different L1 backgrounds show dissimilar patterns in measures of lexical proficiency in L2?

Under the first research question about the relationship between OEPT scores and measures of lexical proficiency, an interesting result was found: when the Spearman correlation was run without Mandarin, the overall correlation coefficients increased for most selected variables. This implies the possibility that examinees with different L1 backgrounds present different patterns in the measures of lexical proficiency. To look into the details among different L1 backgrounds, descriptive statistics were run for all measures of lexical proficiency except for Off-list types, which had weak correlations with OEPT scores ($r_{\text{OEPT, Offtyp}} = 0.05$).

The samples which contained D0s were included through each variable for the descriptive statistical analysis, even though the value of 0 was completely removed only in the measurement of D. This is because, unlike running correlation coefficients, it is not necessary to maintain an identitical sample size for all the variables.

To better illustrate the big picture as well as comparisons of the different L1 groups in each measure of lexical proficiency, the line graphs for each measure were generated with each L1 sub group's average values on the variables. Tables 4-7, 4-8, 4-9, and 4-10show the detailed statistical numbers used to generate each of the respective graphs.

Table 4-7 Descriptive Statistics of Measures of Lexical Diversity

Index	Group	N	Mean	SD	SE	Min	Max	Lower 99% CI	Upper 99% CI	DF	t Value	Pr>|t|
Token	Korean (35)	25	115.1	48.80	9.76	51	211	87.8	142.4	24	11.79	<.0001
	Korean (40)	25	152.0	45.37	9.07	82	274	126.7	177.4	24	16.76	<.0001
	Korean (45)	25	184.3	54.03	10.81	117	315	154.1	214.5	24	17.05	<.0001
	Korean (50)	25	164.0	50.01	10.00	43	264	136.0	192.0	24	16.40	<.0001
	Mandarin (35)	22	178.5	41.52	8.85	118	277	153.4	203.5	21	20.16	<.0001
	Mandarin (40)	25	178.5	35.24	7.05	104	243	158.8	198.2	24	25.33	<.0001
	Mandarin (45)	25	202.7	40.02	8.00	117	269	180.3	225.1	24	25.32	<.0001
	Mandarin (50)	25	211.0	48.57	9.71	104	311	183.8	238.1	24	21.72	<.0001
	Mandarin (55)	14	207.6	40.54	10.83	118	280	175.0	240.3	13	19.17	<.0001
	Hindi (50)	25	189.8	67.91	13.58	83	307	151.8	227.7	24	13.97	<.0001
	Hindi (55)	25	186.7	70.64	14.13	51	306	147.2	226.2	24	13.22	<.0001
	Hindi (60)	17	178.7	56.01	13.58	103	319	139.0	218.4	16	13.16	<.0001
	L1 English	25	169.6	76.56	15.31	41	344	126.7	212.4	24	11.07	<.0001
Types	Korean (35)	25	57.7	18.39	3.68	32	102	47.4	68.0	24	15.69	<.0001
	Korean (40)	25	72.6	16.07	3.21	46	110	63.7	81.6	24	22.60	<.0001
	Korean (45)	25	84.6	17.25	3.45	56	131	75.0	94.3	24	24.53	<.0001
	Korean (50)	25	78.7	16.34	3.27	34	101	69.6	87.9	24	24.09	<.0001
	Mandarin (35)	22	84.6	15.36	3.28	60	124	75.3	93.9	21	25.83	<.0001
	Mandarin (40)	25	82.4	12.74	2.55	59	104	75.2	89.5	24	32.33	<.0001
	Mandarin (45)	25	92.3	14.46	2.89	66	121	84.2	100.4	24	31.92	<.0001
	Mandarin (50)	25	98.1	19.10	3.82	54	129	87.4	108.8	24	25.68	<.0001
	Mandarin (55)	14	95.5	15.69	4.19	62	120	82.9	108.1	13	22.77	<.0001
	Hindi (50)	25	93.5	23.79	4.76	48	126	80.2	106.8	24	19.64	<.0001
	Hindi (55)	25	94.0	25.71	5.14	39	139	79.6	108.4	24	18.28	<.0001
	Hindi (60)	17	91.4	19.28	4.68	68	128	77.8	105.1	16	19.55	<.0001
	L1 English	25	85.9	31.43	6.29	32	163	68.3	103.5	24	13.67	<.0001

Table 4-8 Descriptive Statistics of Measures of Lexical Diversity

Index	Group	N	Mean	SD	SE	Min	Max	Lower 99% CI	Upper 99%CI	DF	t Value	Pr>|t|
TTR	Korean (35)	25	0.53	0.11	0.02	0.33	0.76	0.47	0.59	24	24.38	<.0001
	Korean (40)	25	0.49	0.08	0.02	0.37	0.64	0.45	0.54	24	31.56	<.0001
	Korean (45)	25	0.47	0.06	0.01	0.32	0.59	0.44	0.51	24	37.30	<.0001
	Korean (50)	25	0.50	0.09	0.02	0.58	0.79	0.45	0.55	24	27.63	<.0001
	Mandarin (35)	22	0.48	0.07	0.01	0.34	0.63	0.44	0.52	21	34.30	<.0001
	Mandarin (40)	25	0.47	0.06	0.01	0.39	0.62	0.44	0.50	24	40.43	<.0001
	Mandarin (45)	25	0.46	0.05	0.01	0.37	0.6	0.43	0.49	24	42.49	<.0001
	Mandarin (50)	25	0.47	0.05	0.01	0.63	0.39	0.44	0.50	24	43.60	<.0001
	Mandarin (55)	14	0.47	0.04	0.01	0.4	0.53	0.43	0.50	13	44.64	<.0001
	Hindi (50)	25	0.51	0.07	0.01	0.4	0.68	0.47	0.56	24	34.44	<.0001
	Hindi (55)	25	0.53	0.09	0.02	0.39	0.76	0.48	0.58	24	29.72	<.0001
	Hindi (60)	17	0.53	0.07	0.02	0.4	0.7	0.47	0.58	16	29.11	<.0001
	L1 English	25	0.54	0.08	0.02	0.41	0.8	0.49	0.58	24	32.23	<.0001
D	Korean (35)	12	43.18	13.03	3.76	22.86	69.61	31.50	54.87	11	11.48	<.0001
	Korean (40)	22	48.88	12.94	2.76	28.72	69.79	41.07	56.69	21	17.72	<.0001
	Korean (45)	25	51.10	12.49	2.50	30.02	73.92	44.12	58.09	24	20.46	<.0001
	Korean (50)	24	54.41	14.04	2.87	32.88	85.98	46.36	62.46	23	18.98	<.0001
	Mandarin (35)	22	53.96	13.80	2.94	33.25	90.73	45.63	62.28	21	18.34	<.0001
	Mandarin (40)	25	49.20	13.00	2.60	26.22	90.48	41.93	56.47	24	18.93	<.0001
	Mandarin (45)	25	52.13	10.29	2.06	30.50	74.69	46.37	57.88	24	25.33	<.0001
	Mandarin (50)	25	58.55	15.67	3.13	28.56	102.89	49.78	67.31	24	18.68	<.0001
	Mandarin (55)	14	55.09	13.38	3.57	31.06	85.98	44.32	65.86	13	15.41	<.0001
	Hindi (50)	22	65.73	12.04	2.57	40.87	84.66	58.46	73.00	21	25.61	<.0001
	Hindi (55)	22	66.31	11.93	2.54	46.96	94.99	59.11	73.51	21	26.08	<.0001
	Hindi (60)	17	65.80	20.14	4.89	37.93	113.48	51.53	80.07	16	13.47	<.0001
	L1 English	21	68.37	14.00	3.05	45.66	92.38	59.68	77.06	20	22.38	<.0001

Table 4-9 Descriptive Statistics For K1 Type and K2 Type of Lexical Frequency Profile

Index	Group	N	Mean	SD	SE	Min	Max	Lower 99% CI	Upper 99% CI	DF	t Value	Pr>|t|
K1 Type	Korean (35)	25	46.20	14.22	2.84	27	79	38.24	54.16	24	16.24	<.0001
	Korean (40)	25	59.68	14.72	2.94	36	91	51.44	67.92	24	20.27	<.0001
	Korean (45)	25	68.00	13.33	2.67	45	101	60.54	75.46	24	25.50	<.0001
	Korean (50)	25	63.60	14.14	2.83	22	85	55.69	71.51	24	22.50	<.0001
	Mandarin (35)	22	70.23	11.82	2.52	49	96	63.09	77.36	21	27.87	<.0001
	Mandarin (40)	25	68.08	10.62	2.12	49	88	62.14	74.02	24	32.06	<.0001
	Mandarin (45)	25	76.48	12.07	2.41	58	105	69.73	83.23	24	31.68	<.0001
	Mandarin (50)	25	80.36	17.08	3.42	45	108	70.80	89.92	24	23.52	<.0001
	Mandarin (55)	14	77.86	13.54	3.62	47	100	66.96	88.75	13	21.52	<.0001
	Hindi (50)	25	72.80	18.76	3.75	38	103	62.31	83.29	24	19.40	<.0001
	Hindi (55)	25	73.64	20.10	4.02	33	108	62.40	84.88	24	18.32	<.0001
	Hindi (60)	17	72.71	17.98	4.36	50	112	59.97	85.44	16	16.67	<.0001
	L1 English	25	68.4	25.80	5.16	27	134	54.01	82.87	24	13.26	<.0001
K2 Type	Korean (35)	25	2.84	1.97	0.39	0	8	1.74	3.94	24	7.20	<.0001
	Korean (40)	25	3.20	1.73	0.35	0	7	2.23	4.17	24	9.24	<.0001
	Korean (45)	25	4.12	2.09	0.42	1	10	2.95	5.29	24	9.87	<.0001
	Korean (50)	25	4.60	2.55	0.51	1	12	3.17	6.03	24	9.02	<.0001
	Mandarin (35)	22	5.05	2.26	0.48	1	10	3.68	6.41	21	10.49	<.0001
	Mandarin (40)	25	4.00	2.02	0.40	1	8	2.87	5.13	24	9.90	<.0001
	Mandarin (45)	25	5.48	2.54	0.51	1	13	4.06	6.90	24	10.81	<.0001
	Mandarin (50)	25	4.76	1.71	0.34	1	8	3.80	5.72	24	13.88	<.0001
	Mandarin (55)	14	5.43	1.40	0.37	4	9	4.30	6.55	13	14.52	<.0001
	Hindi (50)	25	5.16	3.00	0.60	1	12	3.48	6.84	24	8.61	<.0001
	Hindi (55)	25	5.16	2.44	0.49	1	10	3.79	6.53	24	10.56	<.0001
	Hindi (60)	17	5.53	2.12	0.52	2	11	4.02	7.03	16	10.73	<.0001
	L1 English	25	4.68	2.81	0.56	0	11	3.11	6.25	24	8.33	<.0001

Table 4-10 Descriptive Statistics for AWL Type and Off Type of Lexical Frequency Profile

Index	Group	N	Mean	SD	SE	Min	Max	Lower 99% CI	Upper 99%CI	DF	t Value	Pr>|t|
AWL Type	Korean (35)	25	3.68	2.25	0.45	0	10	2.42	4.94	24	8.18	<.0001
	Korean (40)	25	4.20	2.02	0.40	2	9	3.07	5.33	24	10.39	<.0001
	Korean (45)	25	5.24	2.68	0.54	2	13	3.74	6.74	24	9.77	<.0001
	Korean (50)	25	5.40	2.69	0.54	2	12	3.89	6.91	24	10.03	<.0001
	Mandarin (35)	22	4.32	1.96	0.42	1	9	3.13	5.50	21	10.33	<.0001
	Mandarin (40)	25	4.32	2.01	0.40	1	8	3.19	5.45	24	10.72	<.0001
	Mandarin (45)	25	4.64	2.22	0.44	1	11	3.40	5.88	24	10.47	<.0001
	Mandarin (50)	25	6.60	3.33	0.67	1	13	4.74	8.46	24	9.91	<.0001
	Mandarin (55)	14	5.36	3.08	0.82	1	10	2.88	7.84	13	6.51	<.0001
	Hindi (50)	25	8.04	3.35	0.67	3	16	6.17	9.91	24	12.01	<.0001
	Hindi (55)	25	7.80	3.07	0.61	3	14	6.08	9.52	24	2.71	<.0001
	Hindi (60)	17	7.06	3.01	0.73	3	14	4.93	9.19	16	9.67	<.0001
	L1 English	25	6.36	3.04	0.61	1	13	4.66	8.06	24	10.46	<.0001
Off Type	Korean (35)	25	5.04	3.10	0.62	1	13	3.30	6.78	24	8.12	<.0001
	Korean (40)	25	5.56	2.84	0.12	1	10	3.97	7.15	24	9.77	<.0001
	Korean (45)	25	7.16	4.08	0.82	1	14	4.88	9.44	24	8.78	<.0001
	Korean (50)	25	5.16	3.05	0.61	1	13	3.45	6.87	24	8.46	<.0001
	Mandarin (35)	22	6.86	3.30	0.70	1	14	4.87	8.86	21	9.76	<.0001
	Mandarin (40)	25	5.92	3.44	0.69	1	12	4.00	7.84	24	8.61	<.0001
	Mandarin (45)	25	5.68	2.59	0.52	2	12	4.23	7.13	24	10.95	<.0001
	Mandarin (50)	25	6.40	2.94	0.59	1	12	4.75	8.05	24	10.87	<.0001
	Mandarin (55)	14	6.93	3.56	0.95	2	15	4.06	9.80	13	7.28	<.0001
	Hindi (50)	25	7.44	3.74	0.75	1	16	5.35	9.53	24	9.94	<.0001
	Hindi (55)	25	7.24	4.15	0.83	1	16	4.92	9.56	24	8.73	<.0001
	Hindi (60)	17	6.06	2.90	0.70	2	12	4.00	8.12	16	8.60	<.0001
	L1 English	25	6.44	4.04	0.81	1	15	4.18	8.70	24	7.97	<.0001

4.3.1 Lexical Diversity

4.3.1.1 Tokens

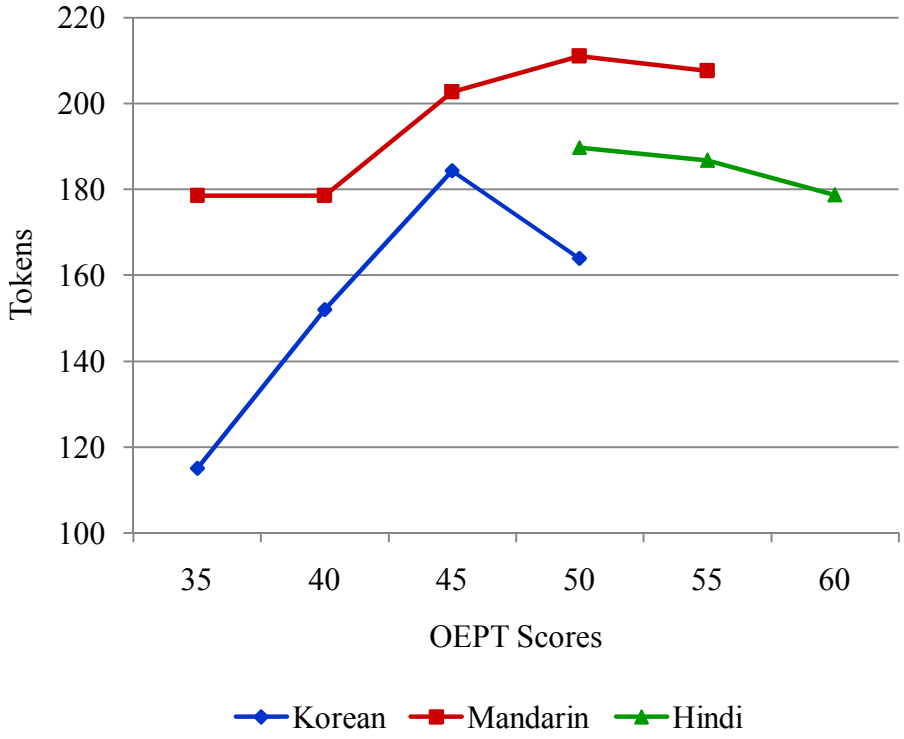

Figure 4-3 The Pattern of Tokens in Three L1 Groups

As shown in Figure 4-3, in the overall picture of the relationship between tokens and OEPT scores, Mandarin and Korean show the pattern in general that the higher the OEPT scores, the more tokens produced. That is, there is a tendency that people speak more words when the OEPT scores are higher: simply speaking, at higher score levels more words are produced. On the other hand, Hindi makes a slightly decreasing line, even though Hindi groups are placed higher than Korean and do not produce a big difference in tokens among its sub OEPT score groups. In terms of the relationships among three different L1 groups of Mandarin, Korean, and Hindi, the Mandarin line is

placed higher than the other two groups. This means that Mandarin groups tend to produce more tokens than other L1 groups of Korean and Hindi. Interestingly, Mandarin 35 and 40 achieve much higher numbers of tokens than Korean 35 and 40, as well as produce a similar number of tokens as Hindi groups which have OEPT scores over 50.

4.3.1.2 Types

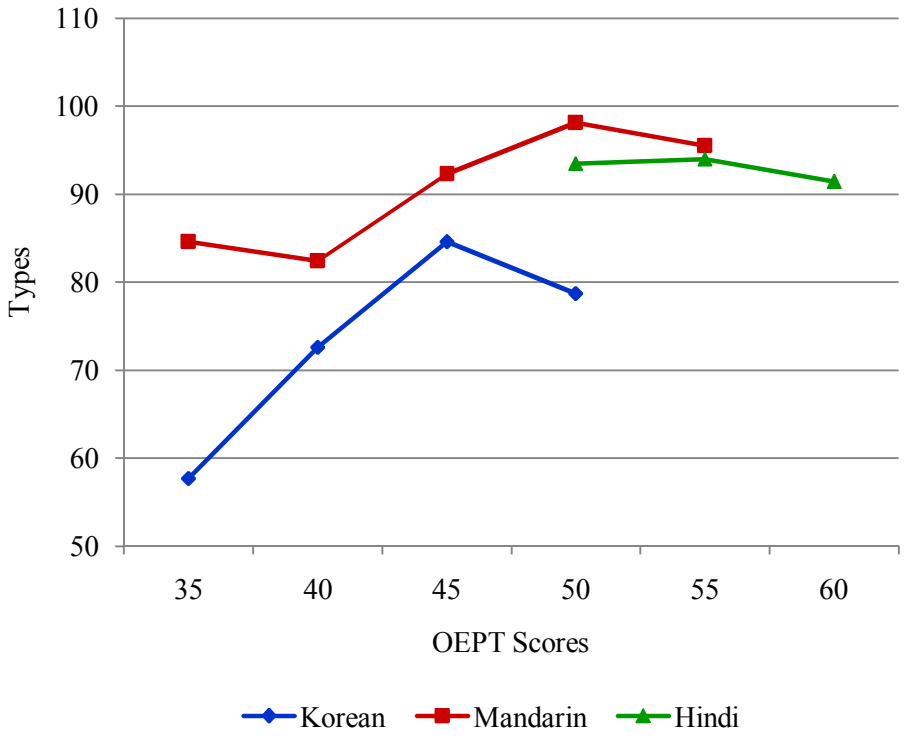

Figure 4-4 The Pattern of Types in Three L1 Groups

As shown in Figure 4-4, types create similar patterns to the tokens. Overall, all three L1 groups generate increasing lines of types in general. In other words, L2 speakers generally tend to speak more unique words in a second language when they obtain higher scores in a speaking test. The Mandarin line is placed higher than the other two L1 groups as was found for tokens. However, the difference between Mandarin and other

two L1 groups becomes smaller than the case of tokens: Mandarin 35 and 40 are located lower than Hindi groups. Also, Mandarin became closer to Hindi 50, compared to the case of tokens, and even Mandarin 55 almost overlaps with Hindi 55.

4.3.1.3 TTR

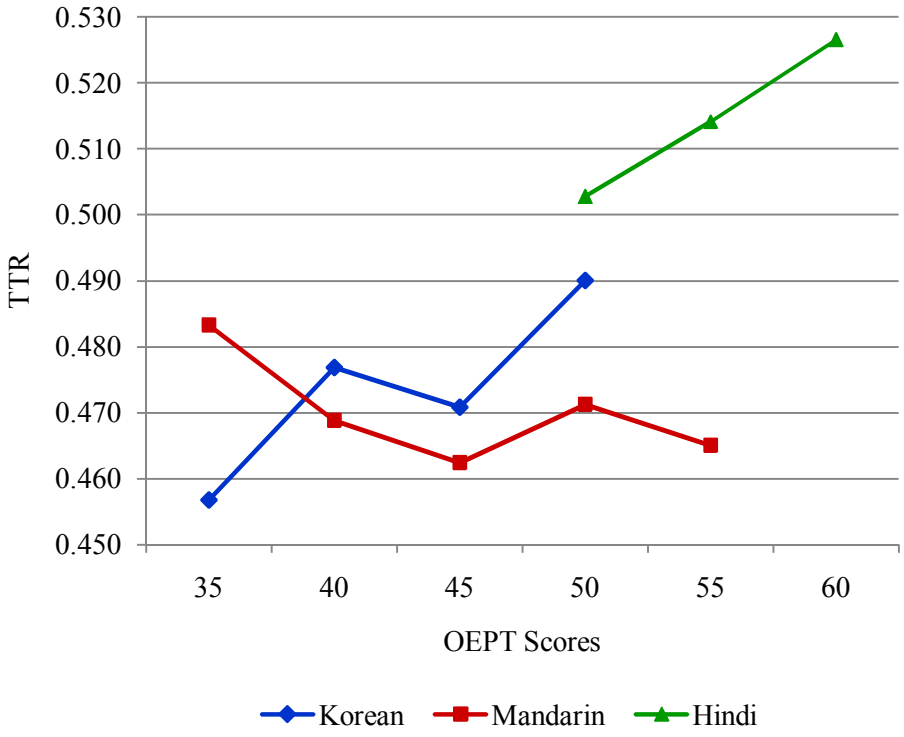

Figure 4-5 The Pattern of TTR in Three L1 Groups

Figure 4-5 shows that TTR has different patterns in each L1 group. Korean and Hindi create an increasing trend together in general, while Mandarin produces a decreasing line. In other words, TTR has a discrimination power for the Korean and Hindi groups, but not for the Mandarin group.

4.3.1.4 D

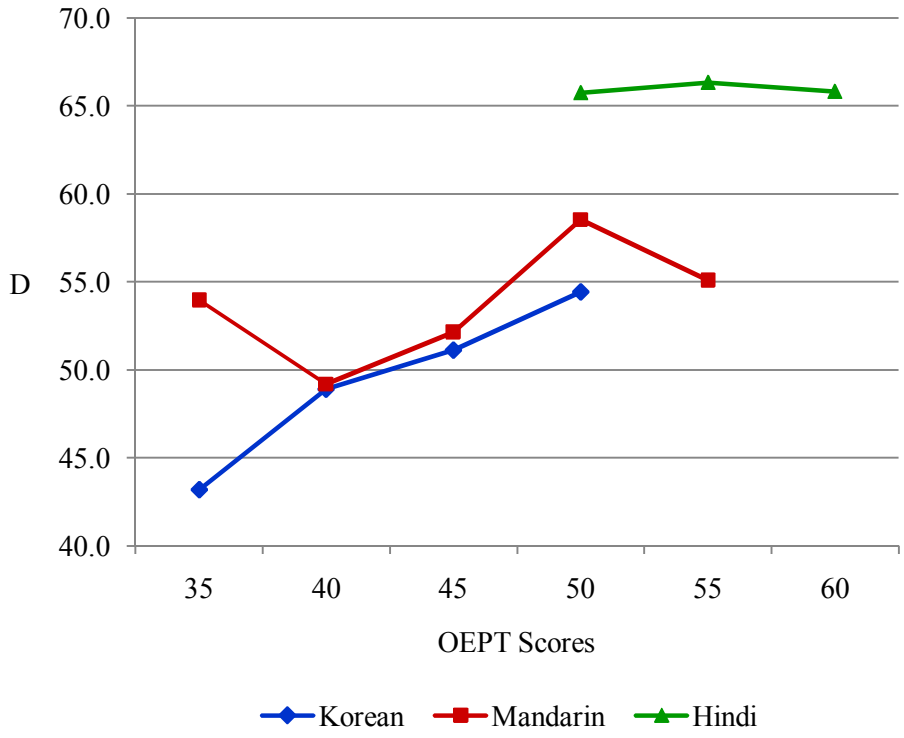

Figure 4-6 The Pattern of D in Three L1 Groups

Figure 4-6 demonstrates that an increasing pattern of D was generated as OEPT scores rise in general. This implies that when there is a higher score in an L2 speaking test, there is likely to be more various lexical items. However, D shows dissimilar patterns, depending on the different L1 background. Korean shows a clear increasing trend of D from low to high OEPT score groups, which appear linier. In the Mandarin group, on the other hand, D appears to discriminate well at levels 40, 45, and 50, but not at the extremes of 35 and 55, showing a somewhat up-and-down pattern: Mandarin 35 is comparatively high while Mandarin 55 is low. Hindi is located much higher than Mandarin as well as Korean, implying that Hindi has much higher lexical diversity in

speech than the other two groups. However, the line of Hindi is flat within itself, indicating that D has little discrimination of OEPT scores at the higher levels suggesting a threshold for D.

Compared to the cases of tokens and types, the difference between Mandarin and Korean became narrower. Except for Mandarin 35, the other sub groups in Mandarin are almost overlapped with Korean. Mandarin 35 produced D, which is much higher than Korean 35 and even similar to the other 50 or 55 OEPT score groups.

4.3.1.5 Comparing four variables of lexical diversity

In terms of the relationship between OEPT scores and lexical diversity, four variables of lexical diversity, which are tokens, types, TTR, and D, demonstrate similar patterns in general. However, the Mandarin group shows dissimilar patterns compared to the Korean and Hindi groups through all four variables of lexical diversity. First of all, the comparable part among tokens, types, TTR, and D is that all four variables present an increasing pattern. That is, the value of each variable in lexical diversity rises when OEPT scores increase. In other words, generally speaking, when one has a low L2 speaking score, it is likely that the diversity of words used is narrow. On the other hand, the possibility of using more diverse lexical items would be higher when one's L2 speaking score becomes higher.

Four variables of lexical diversity – tokens, types, TTR, and D – also show difference especially in terms of how the Mandarin group works in each variable. Namely, Mandarin produced a much higher value of tokens than Korean or Hindi did. However, in types, the difference between Mandarin and the other two groups becomes smaller. In

TTR, Mandarin showed a decreasing trend while an increasing trend was generated in Korean and Hindi. In D, the pattern of Mandarin was up and down, which made the trend relatively flat.

Among all four variables of lexical diversity which are tokens, types, TTR, and D, D demonstrated the most stable pattern, resulting in the best discrimination and prediction for OEPT scores. This is not only because D has a strong correlation with OEPT scores in both including and excluding Mandarin, but also because in the analysis of descriptive statistics, D demonstrated the increasing pattern with OEPT scores increasing. The difference between Mandarin and the other two L1 groups of Korean and Hindi was the smallest, compared to the cases of Korean and Hindi as well.

4.3.2 Lexical Frequency Profile

The measure, Lexical Frequency Profile (LFP), is invented by Laufer and Nation (1995). LFP is composed of four indices: K1 types, K2 types, AWL types, and Off types. K1 types refer to the number of words included in the list of 1st 1000 most frequent words. Similarly, K2 types is the number of words included in the list of 2nd 1000 most frequent words while AWL types is the number of words in Academic Word List. Off types is the number of words which do not appear in the above three lists.

4.3.2.1 K1 Types

Figure 4-7 presents that all three groups of Mandarin, Korean, and Hindi generate the overall increasing pattern of K1 types (1st 1000 most frequent words) even though there are some points decreasing, as the general pattern of K1 types is similar to the pattern of tokens. This demonstrates that the higher the OEPT score, the more unique the

words used from the first 1,000 frequent word list. Comparing three L1 groups, Mandarin is placed much higher than Korean and Hindi. This is parallel to the results in tokens as well.

K1 types generally has discrimination through all sub groups of OEPT scores. However, the discrimination becomes weak when the OEPT score is over 50: the Korean 50 and Mandarin 55 come down, and Hindi generates a flat line which has a weak discrimination.

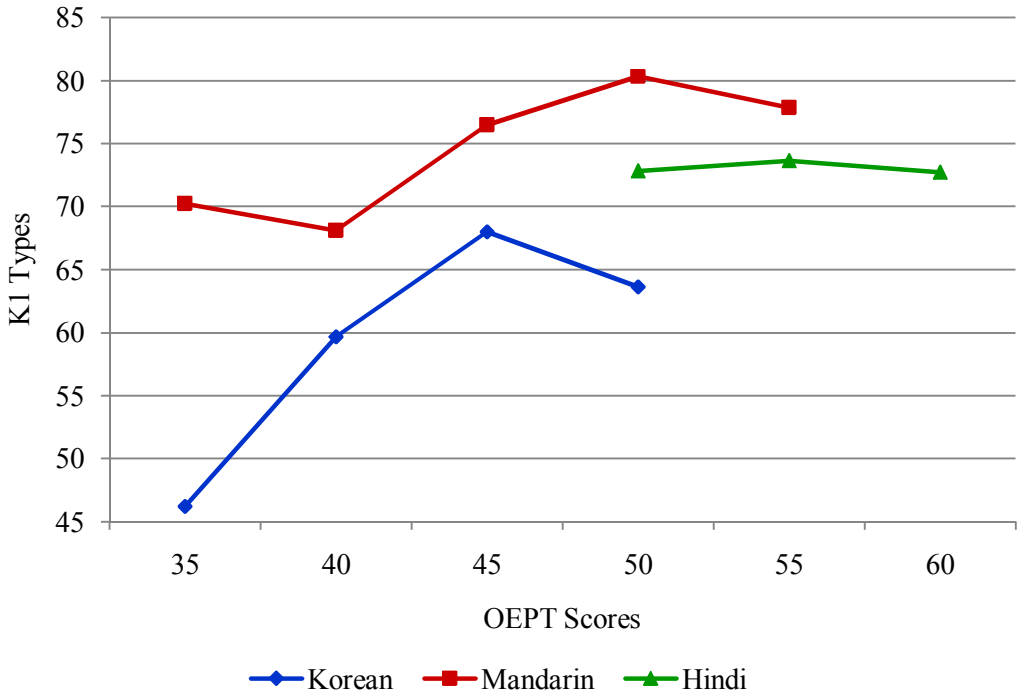

Figure 4-7 The Pattern of K1 Types in Three L1 Groups

As in the variable of tokens, Mandarin 35 also shows an exceptional pattern which does not fit the overall trend: Mandarin 35 is greatly higher than Korean 35 as well as well as is similar to Hindi. In other words, Mandarin 35 produces the unexpectedly high value of K1 types, even though their OEPT score is the lowest.

4.3.2.2 K2 Types

Figure 4-8 shows that the values of K2 types (2nd 1000 most frequent words) generally increase as the OEPT scores increase, even though Mandarin shows a considerably different pattern than the other two L1 groups. First of all, Korean and Hindi display an increasing trend of K2 types in which there is no decreasing point. This means that K2 types allows good discrimination in the groups of Korean and Hindi. On the other hand, the line of Mandarin creates the severe up-and-down pattern through the OEPT scores as though Mandarin still generates more K2 types than Korean and Hindi.

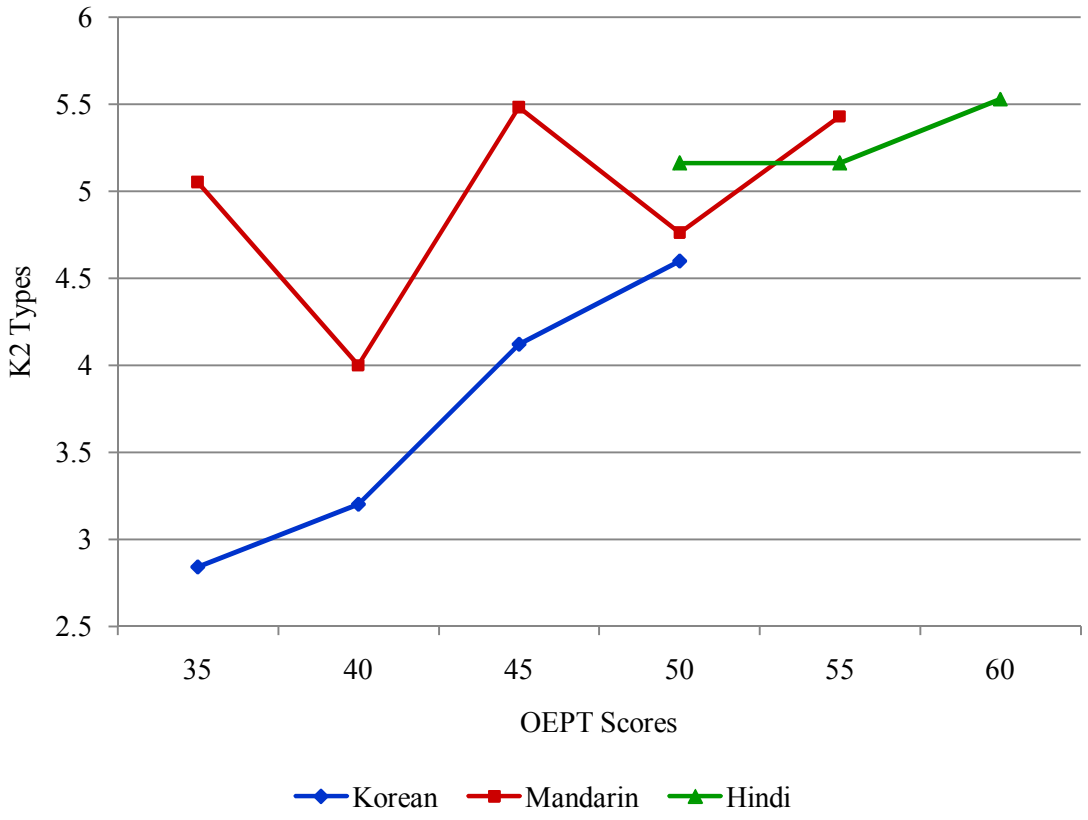

Figure 4-8 The Pattern of K2 Types in Three L1 Groups

4.3.2.3 AWL Types

As shown in Figure 4-9, in terms of AWL types (Academic Word List), the general trend is that the value of AWL types becomes higher when OEPT scores increase. However, each L1 group shows a slightly different pattern: Korean shows the most stable increasing pattern. Mandarin also shows arising tendency even though the line is almost flat between the OEPT scores of 35 and 55 and down at 55. Hindi shows the decreasing pattern of AWL types. That is, the increasing pattern collapses when the OEPT score is over 50.

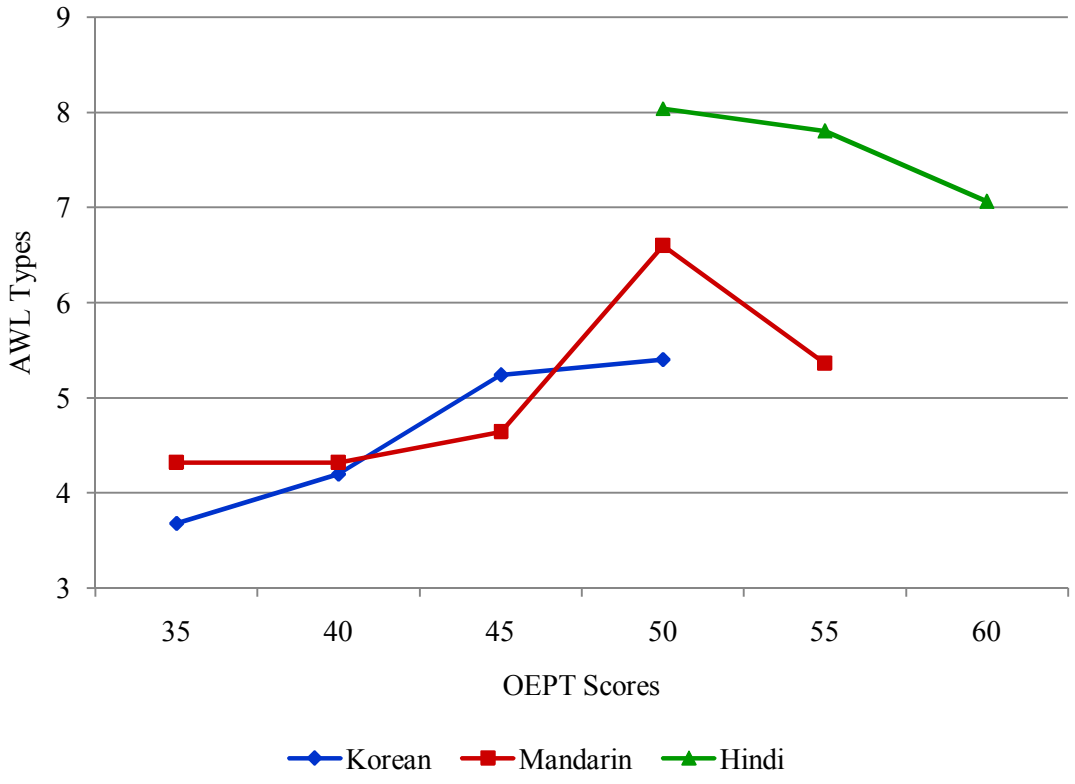

Figure 4-9 The Pattern of AWL Types (Academic Word List) in 3 L1 Groups

4.3.2.4 Comparison Among the Measures of the Lexical Frequency Profile

Three different variables of LFP, which are K1 types, K2 types, and AWL types, present similar patterns in terms of the relationship between each variable and OEPT scores. The overall tendency is that the value of each variable increases when the OEPT score rises. This reveals that people use more unique words when they have better speech proficiency in a second language test. However, the overall increasing pattern collapses when the OEPT scores are over 50. The OEPT sub groups of 50, 55, and 60 made much higher values of each variable than other low OEPT score groups, but in 50, 55, and 60, the values became flat or even slightly decreased. In other words, the discrimination power of the LFP becomes weak for the high OEPT score groups over 50.

Interestingly, the Mandarin group shows a dissimilar trend in each LFP variable. In K1 types, the line of Mandarin was located much higher than the Korean and Hindi groups, which says that Mandarin produced more K1 types than the other two L1 groups. However, in K2 type, Mandarin created large ups and downs through the OEPT scores. In AWL types, the overall increasing pattern was back in the Mandarin line, which even almost overlapped with Korean. In other words, Mandarin shows the pattern that Mandarin uses more unique words than other groups (K1 types) especially when it comes to the first 1,000 frequent words. However, when it comes to the Academic Word List, which is less frequently used, Mandarin shows a similar pattern with the other groups.

CHAPTER 5. DISCUSSION

5.1 Relationship Between Lexical Diversity and L2 Oral Proficiency

The results show that advanced L2 speakers tended to use more diverse words in their speech in a given time than the beginners, as types and D have strong positive correlations with the holistic score of the OEPT. On the other hand, the total numbers of words (tokens) are weakly correlated to OEPT scores. These results indicate that employing various words is more important than simply increasing the number of words used in speech. That is, lexical diversity, whether one can access various words and expressions, is a good indicator to discriminate the advanced learners from the beginners: beginners seem to repeat the same words over again with a limited pool of lexical items, while advanced speakers hold a larger collection of lexical items accessible.

5.2 Do Advanced L2 Learners Use More Infrequent Words Than the Beginners?

The assumption which has been accepted by researchers is that L2 advanced learners use more infrequent words in production than beginners do. According to Laufer and Nation (1995), "The better a learner is, the more likely they are to use more infrequent vocabulary in production" (p. 316). The current study also backed up this assertion: the advanced speakers tend to use more words in the Academic Word List than beginners ($r_{\text{OEPT, AWLtyp}} = 0.39$).This assumption gives L2 learners and educators the

impression that the main focus of L2 vocabulary learning for improving oral proficiency should be expanding to more difficult words.

However, a different story can be found in the big picture. When it comes to the correlations between lexical diversity and the Lexical Frequency Profile as well as the descriptive statistics of the Lexical Frequency Profile, a contradictory conclusion can be drawn. That is, in terms of the relationship between lexical diversity and the Lexical Frequency Profile, types, which was highly correlated with OEPT scores, was in extremely high correlation with K1 types ($r_{\text{K1Types, Types}} = 0.95$), compared to the correlations with other variables of the LFP ($r_{\text{K2Types, Types}} = 0.58$, $r_{\text{AWLTypes, Types}} = 0.42$, $r_{\text{OffTypes, Types}} = 0.47$). This indicates that the major change of types, while L2 advanced speakers use more types (the number of unique words) than beginners, occurs in the first 1,000 most frequent words. Figure 4-2 shows that the proportion of each LFP measure backs this up: while it is true that the advanced learners use more AWL words than the beginners, it was the words in the first 1,000 words that the L2 speakers employ in their speech most frequently. In other words, L2 advanced speakers express their ideas employing frequent words as a rule rather than infrequent ones. These results of the current research indicate that when it comes to vocabulary learning for oral proficiency, the frequent words may be a more important focus in L2 vocabulary learning than infrequent words.*

If the above results are assumed to be true, why and how has the idea that advanced L2 students use more infrequent words in productive language been accepted?

---

* LFP (Lexical Frequency Profile) is composed of four indices: K1 types (1st 1000 most frequent words), K2 types (2nd 1000 most frequent words), AWL types (Academic word list), and Off types (Words not listed in above three).

This is mainly because in terms of lexical proficiency development, writing and speaking tend not to be separately considered, rather being treated together in the boundary of productive language, even though speaking is obviously different from writing regarding its context of use. For example, Laufer and Nation (1995) demonstrated that L2 advanced learners use more infrequent words such as AWL words, but this study was based on L2 writing.

It is natural that the proportion of infrequent words becomes higher in writing due to the characteristics of writing: writing is often done under formal and academic contexts, which need a special structure. By contrast, speaking happens in comparatively casual situations where speed of production is considered. From this speaking condition, as the results show in the current research, it is possible to communicate at an advanced level, employing mostly easy and frequent words. The different features on lexical use between writing and speaking should be applied to L2 learners' vocabulary learning strategies as well as educators' pedagogical approaches.

The important implication for L2 educators and learners, especially at the beginner level, can be drawn from the above results and discussions: if they want to improve their lexical proficiency in L2 speaking, they may have to pay more attention to the first 1,000 more frequent words rather than less frequent words. This sounds simple and logical. Nevertheless, why do many L2 learners still struggle in growing their lexical proficiency in speaking? One of the reasons is that in many cases L2 learners are not aware of the fact that they don't know a certain word. More specifically, many L2 learners misunderstand their current level of lexical proficiency. This is mainly because they believe that they have acquired a certain word when they know only part of the

word's dimensions. For example, people misunderstand that they know a word only when they can receptively recognize that word in a written form.

Their misunderstanding on their own lexical proficiency as well as the process of word learning misleads them to the wrong strategies of vocabulary learning: when they feel they know the word, they stop exploring the same word on other dimensions such as its spoken form or productive dimension. Importantly, even though the easy words should be focused on in teaching and learning in L2 speaking, people put their focus on infrequent words, because of their biased beliefs: 1) advanced students use more difficult words and 2) learners themselves already know the easy words. These biased beliefs should be clearly rechecked in the different contexts between writing and speaking.

In order to prevent misconceptions regarding vocabulary knowledge, understanding the basic structure of a word is essential. First of all, the receptive dimension of a word should be considered a different element from productive dimension. This is because reception goes through a dissimilar process from production. (Nation, 2001). In addition, a written form of a word should be distinctively taught and learned from a spoken form, as knowing the written dimension of a word does not guarantee knowing its spoken dimension.

5.3 Different L1 Background and Lexical Proficiency

The results of this study indicate that OEPT examinees with different L1 backgrounds tend to have different paths of lexical proficiency development. First, Korean is the group to show the most predictable path of L2 lexical proficiency development, compared to Mandarin and Hindi: there was a clear tendency of increasing

values of lexical proficiency measures when the OEPT scores increase. That is, the low OEPT score group in Korean used a lower total number of words as well as had lower lexical diversity. On the other hand, the high OEPT score group employed more words in speech with higher lexical diversity.

The Mandarin group, for all OEPT score groups, used more lexical items in OEPT tests than the Hindi and Korean did: Mandarin used a larger number of words (tokens), more different numbers of words (types), and more of the first 1,000 frequent words (K1 types) than Korean and Hindi in each OEPT score group. From this result, it is assumed that Mandarin speakers have their own special training, which is distinctive from Korean or Hindi, to increase their absolute numbers of vocabulary words used in speech. However, when it comes to the lexical proficiency development inside the group of Mandarin speakers, there is little difference of lexical diversity between low and high score groups in the OEPT. Especially in D, which is considered one of the strong predictors of OEPT scores, Mandarin did not show a special pattern of lexical diversity even though the expected outcome is the increasing pattern of D. This is interesting as Mandarin showed an increasing trend of tokens through OEPT scores. Mandarin's dissimilar trend of lexical proficiency development should be investigated through further studies.

Hindi is the only group to show no scores below 50 in OEPT score results, compared to Korean and Mandarin. That is, all Hindi examinees passed the OEPT test with test results over the OEPT cut-off score, which is 50. Of interest is that there is little difference of lexical proficiency across the Hindi score levels. That is, there is neither strong increasing nor decreasing patterns of lexical proficiency in Hindi. The assumption

which can be drawn from the above result is that the degree of how much lexical proficiency impacts oral fluency becomes less significant when L2 proficiency reaches a certain level, which can be categorized as advanced.

Why does each L1 group show different development patterns of lexical proficiency? The Three Circles Model by Kachru (1988) partly explains this question. According to Kachru, there are three circles which conceptualize the territories depending on the ways English is acquired and used: the Inner Circle, the Outer Circle, and the Expanding Circle. In the Inner Circle, English is used as the dominant first language, e.g., UK, USA, or Canada. On the other hand, in the Outer Circle, English has been used as a second language as a result of colonialism, while in the Expanding Circle, English is learned as a foreign language as a consequence of globalization. In accordance with Kachru's Three Circles Model, Korea and China belong to the Expanding Circle, while India is in the Outer Circle: Chinese students learn English as a foreign language while Hindi students learn English as a second language. That is, Hindi students are naturally exposed to English, using English in school systems and in daily life. On the other hand, Chinese and Korean students learn English as one subjects in school curricula, which makes the exposure limited and motivation weak in learning English. This difference of learning English – whether as a second or foreign language – affects the development of lexical proficiency as well as L2 oral proficiency.

However, even though Kachru's Three Circles Model shows some reasons for the discrepancy among the different L1 groups in terms of lexical proficiency development, still a question remains: Why do Korean and Mandarin present different trends? For this question, it is assumed that the difference in lexical proficiency development between

Korean and Mandarin results from different language learning styles or strategies. In this regard, Reid (1987) found that there is difference in language learning styles among different language backgrounds: Korean students significantly preferred the visual learning style than Chinese students, while Chinese students have a preference for the auditory learning style, compared to Korean. In visual learning style, reading or studying charts is the favored way of language learning, while in auditory learning style, listening to lectures or audiotapes is more preferred. In addition, according to Hong-Nam and Leavel (2007), Chinese preferred more social strategies in L2 learning than Korean, while Korean favored metacognitive strategy than Chinese. In other words, Chinese students may like to learn L2 by cooperating with others while Korean students prefer planning and monitoring in L2 learning. The results of the studies above partly explain how Korean and Mandarin are different from each other regarding language learning styles or strategies, and it can be assumed that the different language learning styles or strategies can affect dissimilar patterns of lexical proficiency development. However, in order to further examine this assumption, additional study needs to be conducted in terms of the relationship between language learning style and lexical proficiency development in different L1 backgrounds.

5.4 Lexical Diversity and Lexical Efficiency Through Different L1 Background

When Hindi is compared to Mandarin, it is obvious that the matter of length in a speech is not the essential factor for receiving a favorable evaluation: More important is lexical diversity. The whole Hindi group whose OEPT score range is the highest with scores of 50 to 60 produced a smaller total number of words (tokens) than Mandarin with

35 to 55 of OEPT scores. Although Hindi uses a lower number of lexical items, Hindi outweighs Mandarin in terms of how diverse vocabulary words are used ($r_{OEPT, D}$= 0.35). This is to show that employing diverse words is more essential than simply producing many words.

Lexical efficiency, however, is more essential than lexical diversity. Lexical efficiency is how efficient vocabulary words are employed and arranged in accordance with a topic given. Comparison between Mandarin 35 and the Hindi group supports this idea. That is, Mandarin 35 shows that regardless of securing lexical diversity with the high value of D, the OEPT scores can be low: lexical diversity with D in Mandarin 35 is similar with Korean 50 and Mandarin 55. On the other hand, in Hindi 50 and 55, there are 6 samples which obtained high OEPT scores even though they used a small number of words--as few as 50 tokens. This shows that one can have a good evaluation on speaking when he or she directly gets to the point, regardless of lexical diversity as well as the number of words produced.

Lexical efficiency is related not only to the difference between writing and speaking but also to cultural difference. First, lexical efficiency differentiates writing from speaking. That is, in writing, lexical efficiency has less emphasis than in speaking, as vocabulary use depends on the writing structure which contains an introduction, body, and conclusion. By contrast, in speaking, it is important to touch the core promptly by efficiently selecting key words which are relevant to the topic. Second, lexical efficiency is also related to cultural difference in terms of the way to deliver a speech. Generally speaking, in Western culture, direct speech in which a main idea is followed by an explanation is preferred. In Eastern culture, on the other hand, preference is on indirect

speech, where the background is explained first, and then the conclusion or main idea is suggested at the end. This explains why Mandarin 35 did acquire the low OEPT score even though Mandarin 35 has a high value of lexical diversity: Mandarin 35 failed to have lexical efficiency.

5.5 The Sensitivity of D

It turned out that D is one of the strongest predictors for OEPT scores ($r_{OEPT, D}$= 0.35). In the meantime, a note of caution is that 0 values of D (D0s) were found for 23 samples out of 278. This is abnormal as 0 of D means that lexical diversity does not exist for the sample. The reason for D0s is that those 23 samples had too few tokens. It has already been established that tokens should be at least 50 to run the software for D. This is because the software randomly picks up 35 to 50 tokens for plotting the TTR against tokens as not replacing the tokens for the next random sampling (Malvern at al., 2004). In the current study, however, D0s were found not only with the tokens around 50 but also with the tokens under 100. This indicates that D is affected by the sample size; hence, it is not a proper measure of lexical diversity to be employed for some cases of free speech, in which it is possible that the sample could be extremely short.

5.6 Relationship Between the Findings and the Cube

The Cube, as a new model of lexical proficiency, is shaped as three dimensional as it is comprised of three axes (breadth, depth, and fluency) and three sides (reception, variety in spoken production, and sophistication in spoken production). Among these three axes and three sides, two sides of the Cube were examined in this study: *variety in spoken production* and *sophistication in spoken production*. Variety in spoken production was measured with

four indices of lexical diversity (tokens, types, TTR, and D) as the measure of lexical diversity corresponds with the quantitative characteristics of the side, variety in spoken production. Similarly, LFP was selected for sophistication in spoken production as the side, sophistication in spoken production, is related to "how" to use vocabulary as the matter of quality in vocabulary use.

The results of this research imply three important points in terms of the Cube. First, L2 speech proficiency is strongly related to the side of variety in spoken production. In other words, people with advanced L2 speech proficiency tend to quantitatively produce more lexical words in their speech. Second, in terms of sophistication in spoken production which is related to quality of vocabulary, advanced L2 speakers tend to speak using more frequent words, which are considered more familiar and easier, than less frequent words. Third, depending on a speaker's L1 background, the Cube could develop in a different pattern. In this study, three different L1 backgrounds were examined, and each different L1 group show dissimilar patterns of vocabulary acquisition in the Cube.

5.7 The Cube as a Better Representation of Lexical Proficiency

The Cube is a better representation of lexical proficiency on three grounds. First, the Cube is the first model of lexical proficiency to combine two continuums of breadth-depth and reception-production while previous models of lexical proficiency focus on either one of two continuums. Second, the Cube draws a sharp dividing line between breadth and depth which has been ambiguously defined in previous studies, by employing some parts of Nation (2001)'s definition in lexical proficiency. Third, the Cube reflects the developmental difference between writing and speaking by visually dividing two areas of written and

spoken lexical proficiency. By contrast, none of the previous lexical proficiency models have reflected the difference between speaking and writing in L2 development even though the research results have proved that written lexical proficiency differently developed from spoken area.

As discussed above, the Cube better corresponds to lexical proficiency than any other previous models. Nevertheless, there is a need to make a practical version of the Cube for L2 learners and educators. This is because the Cube is a combination of theoretical and practical grounds, containing not only the outcome of vocabulary acquisition (practical) but also developmental process of lexical proficiency (theoretical). In actual sites of L2 learning and teaching, the theoretical bases in the Cube do not necessarily have to be dealt with: L2 learners and educators can better focus on the target outcomes of vocabulary learning as it is more effective in setting goals of learning and teaching. For this reason, the practical version of the Cube, which excludes developmental process of vocabulary acquisition, is introduced in the section of pedagogy as well.

CHAPTER 6. PEDAGOGICAL IMPLICATIONS

Based on the results of the current research, six practical steps to increase the power of lexical proficiency in L2 speech are suggested as pedagogical implications.

The first is to realize the fact that knowing a word is not a simple process, as a word has diverse dimensions. L2 learners often do not realize the above facts and tend to believe they know a word when they know the word's written form and its meaning. However, knowing a word is not that straightforward. To be more specific, when it comes to lexical proficiency, two conventions need to be considered: whether it is spoken or written and whether it is receptive or productive. Also, as Nation (2001) pointed out, three basic elements are needed for knowing a word: form, meaning, and use. Table 6-1 shows the simple way to understand the basic four categories to know a word.

Table 6-1Structure of Lexical Proficiency

		Written	Spoken
Receptive	Form		
	Meaning		
	Use		
Productive	Form		
	Meaning		
	Use		

With two conventions of lexical proficiency, four areas are constructed: written-receptive, spoken-receptive, written-productive, and spoken-productive. Each area of lexical proficiency has to be focused on and planned separately in terms of L2 vocabulary learning and teaching.

Table 6-1 for understanding lexical proficiency contains the equivalent facets with Nation's model (2001) of "What is involved in knowing a word" (p.27): both have three main components of knowing a word as form, meaning, and use. On the contrary, there also exists a big difference between the table above and Nation's definition of knowing a word: in the table above, the written and spoken areas are separated through all three basic components of form, meaning, and use, unlike Nation, who put the spoken and written areas under the element of "form" as its sub-categories. The vocabulary acquisition process is dissimilar in spoken and written areas. Even though one can read or write a word, it is possible that he or she cannot understand it in listening nor speak it. Therefore, spoken and written areas should be separately dealt with through the basic word elements of form, meaning, and use.

As a second step, clear target areas of lexical proficiency to improve should be selected among the above four. Ideally, the best case will be that all four areas are checked. Importantly, L2 learners and educators should be aware that each area of lexical proficiency is dissimilar in terms of its developmental process, hence needing different approaches to teaching and learning. In this case, the productive-spoken lexical proficiency becomes the target area as the current research is about lexical proficiency in L2 speech.

The third step is to decide "which words" need to be focused on in L2 vocabulary learning. Often L2 learners tend to focus on difficult and less-frequent words in their vocabulary learning rather than easy and more-frequent words. This is because they consider themselves to already know easy and frequent words, even though they know the words only partially, for example, on the receptive-written area. Consequently they try to learn more difficult words to expand their vocabulary knowledge rather than spending more time on more-frequent words. As the current study shows, however, advanced L2 speakers express their ideas and enhance their speaking fluency, mainly employing the first 2,000 frequent words. Of course advanced L2 speakers also produce more words on the Academic Word List, but the overall rate of AWL was low, below 1% out of the total number of different words they used in their speech. In other words, it is possible to make L2 speech primarily using the first 2,000 most frequent words. That is, it is significant to notice that in terms of the level of lexical productivity in L2 speech, the level of L2 learners depends on speed--how fast one can recall and produce vocabulary words rather than how many difficult words are known. Namely, if a learner is considered a beginner in L2 speech, he or she should be encouraged to revisit the first 1,000 words learned and the second 1,000 words learned which the learner believes are already known.

The fourth step is to practice the target lexical items by speaking out orally. This sounds very simple, but many L2 learners fail to focus. On oral production, this happens as they misunderstand the difference of the developmental processes between written and spoken areas. Even though an L2 learn has written vocabulary knowledge for a certain word, it is not easy to use the word in the spoken area, if the learner is not accustomed to

activating that word in the spoken area of lexical proficiency. Therefore, L2 learners have to speak target words out loud if they desire to use the words in speech. This is not only because they can hear the sound by repeatedly speaking out the words but also because they can practice more accurate pronunciation.

This "speaking out" strategy can be carried out when L2 learners learn a specific list of target words. Namely, they can speak out the target words repeatedly as they think about the meaning and use of the words. Also, the strategy of "speaking out" can be done even while reading. This might have a similar effect with reading out loud. However, speaking out words should be more than reading out loud as many studies have supported the fact that reading out loud is effective for incidental vocabulary learning. Importantly, L2 learners need to have a clear sense of purpose to find the unfamiliar words in reading and make it possible to use them in speech by practicing. This sense of purpose to expand the spoken-productive area of lexical proficiency will help L2 learners do active vocabulary learning while reading rather than passively waiting to pick up words incidentally.

The fifth step is to practice making example sentences using the target words. It is obvious that the receptive area of lexical proficiency is different from the productive area in terms of the development process. This is akin to the example of driving: whether the driver has knowledge about the car or driving is dissimilar from whether the driver can actually drive the car. That is, the receptive knowledge about a car does not guarantee the active ability to drive a car, so the driver has to practice driving until he or she can drive skillfully. Likewise, one needs to do copious amounts of practice in productively speaking target words in sentences in order to expand receptive word knowledge into the

productive area: using the words by making sentences. Being able to use the word in a sentence means that one understands the spoken form and meaning as well as how to "use" the word. Therefore, L2 learners need to repeatedly use target vocabulary words in sentences repetedly until they feel they can automatically recall and produce the words in speech without hesitation.

As a sixth step, one needs to do association training to increase lexical diversity in the spoken-productive area. Increasing lexical diversity means increasing the range and diversity of words used in speech. In other words, when lexical diversity is high in speech, one can recall and use various words with the same meaning about a certain topic rather than repeating the same words over again. Augmenting lexical diversity is highly correlated with association. Nation (2001) explains several ways to enhance associations: finding substitutes, explaining connections, making word maps, classifying words, finding opposites, suggesting causes or effects, suggesting associations, and finding examples. Of course, in order to increase spoken-productive lexical proficiency, it is important that the above activities should be carried out orally.

Last but not least, increasing lexical efficiency should be focused on in vocabulary learning for speech as well. Lexical efficiency refers to how to efficiently structure vocabulary words chosen in accordance with the speech topic given. In other words, lexical efficiency is related to the process of arranging vocabulary words in speech. Increasing lexical efficiency is important as the structure of speech affects the speakers' attention as well as favorableness of the speech. That is, in order to increase lexical efficiency, it is essential to practice speaking the key points first, which is followed by the relevant background.

CHAPTER 7. SUGGESTIONS FOR FUTURE RESEARCH

There are three suggestions for the future research. First, the difference in lexical proficiency development between L2 writing and speaking needs to be investigated. One of the assumptions in the current research was that L2 spoken lexical proficiency develops differently from L2 writing. The limitation, however, is that this assumption was drawn by comparing different studies which were conducted with the samples collected in different settings. In order to verify this assumption, overcoming its fundamental limitation, a comparative study on lexical proficiency development in L2 writing and speaking should be conducted with writing and speaking samples collected from the same participants.

Secondly, the relationship between L2 lexical proficiency development and language learning style needs to be investigated. Especially, it should be examined whether there is any difference in language learning styles with different L1 backgrounds, and if any, how the dissimilar L2 learning styles derived from different L1 backgrounds affect lexical proficiency development. In particular, in the current research, Mandarin showed different patterns of lexical proficiency development, compared to Korean and Hindi. Consequently, it should be investigated how Mandarin's different language learning style brings the different path of lexical proficiency development. In addition, the investigation should look at why Mandarin is different from Korean in terms of

lexical proficiency development, although they both are part of Asian culture as neighboring countries.

The third topic that is recommended for future investigation is how the topic of speech is related to L2 lexical proficiency. In the current study, the speech samples analyzed were the responses to newspaper item: in this item, the OEPT examinees express whether they agree or disagree with a newspaper article. This is considered one of the limitations of this study as it is possible that the result can be different depending on what kind of topic the examinees deal with in a speech. Therefore, in a future study, it is recommended to investigate how speech topics can affect lexical proficiency as well as oral proficiency.

CHAPTER 8. CONCLUSION

Second language proficiency development is strongly correlated with lexical proficiency. Vocabulary acquisition is not only about knowing a word's meaning, but also involves containing complex dimensions. Thus, one cannot say that he or she knows a word if one knows only its meaning. Nevertheless, many L2 learners or educators tend to think that learning vocabulary is easy and merely a small part of language learning. This misunderstanding of vocabulary acquisition is a minor issue in terms of receptive/written area of L2 learning such as reading. However, when it comes to the spoken/productive area, L2 speaking, this is problematic because there is no guarantee that one can use the word in speech even though the word is understood in reading.

From the conclusion above, L2 learners and educators should have a clear note that vocabulary learning for speech should be separately treated from receptive or written lexical proficiency. In detail, the focus for L2 spoken vocabulary learning should be on expanding the range of productively usable lexical items to increase oral fluency. At the same time, increasing lexical efficiency in accordance with a topic should be an educational focus as well. Additionally, it is important to recheck the easy words, which are considered already "known" if those easy words are easily accessible in speech.

In terms of measures of lexical proficiency, various measures were employed in the current research, including measures of lexical diversity and the Lexical Frequency

Profile. Each measure has its own function for measuring lexical proficiency as well as for predicting the OEPT scores. Especially, it turned out that types, D, and AWL types have a moderately strong correlation with OEPT scores. However, as the results show, no single measure was able to perfectly predict the OEPT scores. That is to say, it is important to consider various measures of lexical proficiency together as they are complementary.

REFERENCES

REFERENCES

Aitchison, J. (1994). *Words in the mind. An introduction to the mental lexicon.* 2nd edition. Oxford: Blackwell.

Almeida, J. (1984). The interplay of cohesion and coherence in native and non-native written academic discourse.Dissertation Abstracts International, 46(5), 1263A.

Anderson, R. C., & Freebody, P. (1981).Vocabulary Knowledge. In J.T. Guthrie (ed.) *Comprehension and Teaching: Research Reviews* (p. 77-117). Newark, DE: International Reading Association.

Brown, C., & Payne, M. (1994). Five Essential Steps of Processes in Vocabulary learning. Paper presented at TESOL Convention, Baltimore, MD.

Carlson, S. (1988). Cultural differences in writing and reasoning skills. In A. Purves (Ed.), Writing across languages and cultures: Issues in contrastive rhetoric (pp. 227-260). Newbury Park, CA: Sage.

Carroll, J. B. (1964). *Language and thought.* Englewood Cliffs, NJ : Prentice-Hall.

Connor, U. (1984).A study of cohesion and coherence in English as a second language students' writing. Papers in Linguistics, 17(1-4), 301-316.

Crossley. S. A.,& McNamara. D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing.*Journal of Second Language Writing, 18*, 119-135.

Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2010). Predicting lexical proficiency in language learner texts using computational indices.*Language Testing, 28*(4), 561-580.

Crossley, S. A., Salsbury, T., &McNamara, D. S. (2011).Predicting the proficiency level of language learners using lexical indices. Language Testing, 1-21, doi:10.1177/0265532211419331.

Daller, H., Milton, J., &Treffers-Daller, J. (2007).*Modelling andassessing vocabulary knowledge*. Cambridge: Cambridge University Press.

Duran, P., Malvern, D., Richards, B., &Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics, 25*(2), 220-242.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Engber, C. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. Journal *of Second Language Writing, 4*(2), 139-155.

Ferris, D. R. (1994).Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency, *TESOL Quarterly*, 28(2), 414-20.

Goodfellow, R., Lamy, M., & Jones, G. (2002).Assessing learners' writing using lexical frequency.*ReCALL, 14*(1), 133-145. DOI: 10.1017/S0958344002001118.

Harris, V., & Snow, D. (2004). *Classic pathfinder: Doing it for themselves: Focus on learning strategies and vocabulary building*. London: CILT.

Hafernik, J. (1990). Relationships among English writing experience, contrastive rhetoric, andEnglish expository prose of L1 and L2 college writers. Dissertation Abstracts International, 51(12), 4007A.

Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21(2), 303-317.

Hong-Nam, K. & Leavell. A. G. (2007). A Comparative study of language learning strategy use in an EFL context: Monolingual Korean and bilingual Korean-Chinese university students. *Asia Pacific Education Review* 8 (1), 71−88.

Hu, Z., Brown, D., & Brown, L. (1982).Some linguistic differences in the written English of Chinese and Australian students.Language Learning and Communication, 1(1), 39-49.

Indrasuta, C. (1988). Narrative styles in the writing of Thai and American students. In A. Purves (Ed.), Writingacrosslanguagesandcultures:Issuesin contrastiverhetoric (pp. 206-226). Newbury Park, CA: Sage.

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: a developmental perspective. *Working Papers, 53*, 61-79.

Kachru, B. B. (1988). The sacred cows of English. *English Today*, *4(16)*, 3-8.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the Input Hypothesis, *The Modern Language Journal*, 73(4), 440-464.

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*(3), 307-322.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? *Applied Linguistics, 19*, 255-271.

Laufer, B. (2003). The influence of L2 on L1 collocational knowledge and on L1 lexical diversity in free written expression.In V. Cook (Ed.), *Effect of the second language on the first* (pp. 19-31).Clevedon: Multilingual Matters.

Linnarud, M. (1986). Lexis in composition:Aperformanceanalysis of Swedish. Lund, Sweden: Liber Forlag Malmo.

Mahmoud, A. (1982). A functional analysis of written compositions of Egyptian students of English and the implications of the notional functional syllabus for the teaching of writing.Dissertation Abstracts International, 44(5), 1439A.

Malvern, D. D., Richards, B. J., Ghipere, N., & Duran, P. (2004).*Lexical diversity and language development: Quantification and assessment.*Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.

Mellor, A. (2011). Essay length, lexical diversity, and automatic essay scoring.*Memoirs of the Osaka Institute of Technology, Series B 55*(2), 1-44.

Milton, J. (2009). *Measuring second language vocabulary acquisition*.Clevedon, UK: Multilingual Matters.

Muncie, J. (2002). Process writing and vocabulary development: comparing Lexical Frequency Profile across drafts. System, 30, 225-235.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Oi, K. (1984). Cross-cultural differences in rhetorical patterning: A study of Japanese and English. Dissertation Abstracts International, 45(8), 2511A.

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10, 355-371.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Reid, J. (1988). Quantitative differences in English prose written by Arabic, Chinese, Spanish, and English students. Dissertation Abstracts International, 50(3), 672A.

Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. TESOL Quarterly, 27(4), 657-675.

Van Woerkum, C. (2007). Orality and the process of writing. *Journal of Technical Writing and Communication*, 37(2), 183-201. doi: 10.2190/j181-tw15-7215-6271.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics, 22*, 217-234.

Yu, G. (2009). Lexical diversity in writing and speaking task performances.*Applied Linguistics, 31*(2), 236-259.

VITA

VITA

YUNJUNG YOU

EDUCATION

Ph.D. Purdue University, West Lafayette, IN December 2014
- Major areas: Teaching English as a Second Language, Language Testing, Second Language Writing, Corpus Linguistics
- Dissertation: "Relationship between lexical proficiency and L2 oral proficiency"
- Advisor: Dr. April Ginther

M.S.Ed. Purdue University, West Lafayette, IN December 2007
- Major areas: Educational Psychology, Gifted Education and Talent Development
- Advisor: Dr. Marcia Gentry

B.A. Ewha Womans University, Seoul, Korea February 2003
- Major areas: Elementary Education
- Advisor: Dr. Min-kyeong Kim

LICENSES AND CERTIFICATES

2013	Minnesota Teaching License Area: Elementary Generalist for K - 6	Department of Education Minnesota, US
2013	ESL Teaching Certificate	Purdue University West Lafayette, IN
2011	Indiana Teaching License Area: Elementary Generalist(Primary/Intermediate)	Department of Education Indiana, US
2003	Elementary Teacher License	Department of Education Seoul, Korea
2010	CPR/AED/First Aid – certified	American Red Cross

2010 Passed all the Praxis exams required for the Indiana math teacher license (for middle and high schools)

EXPERIENCE

Minnesota State University **Mankato, MN, U.S. A.**
Department of English
Full Time Faculty August, 2014 – present

- Teach 3 courses per semester and develop teaching materials for new course assignments
    - ENG 101: Composition for non-native English speakers
    - ENG 207: Technical Communication for STEM
- Utilize online communication tool for teaching – Desire2Learn (D2L)
- Advise undergraduate and graduate students academically
- Support undergraduate student research
- Actively serve on departmental committees and meetings -- English Department meetings, TESL Track meetings, Technology committees, meetings for Brazilian cohort
- Support Intensive English language Institute (IELI) by participating in meetings of instructors and advising international students on a regular basis
- Investigate into developing methods to effectively measure student learning in undergraduate education in English with the purpose of increasing student retention

Purdue University **West Lafayette, IN, U.S. A.**
Department of English
Graduate Instructor August, 2011 – May, 2013

- Teach ENG 106 for First-Year Composition for native English speakers as well as ENG 106-I, First-Year Composition for international students only
    - Designing the syllabus, assignments, class materials and evaluating assignments
    - Actively employing Blackboard to promote communication with students and to provide relevant materials
    - Reviewing the students' papers and giving comments and feedback on them
    - Differentiating the content to meet the individual needs by having one-on-one individual conferences on a regular basis
    - Instructor evaluation:
        - Sp 2013 - **4.8** / 5.0, Fall 2012 - **4.9** / 5.0, Sp 2012 - **4.7** / 5.0, Fall 2011 - **4.4** / 5.0
    - Course evaluation:
        - Sp 2013 - **4.6** / 5.0, Fall 2012 - **4.7** / 5.0, Sp 2012 - **4.7** / 5.0, Fall 2011 - **4.3** / 5.0

Seoul Yangwon Elementary School **Seoul, Korea**
English teacher 2009, 2003 – 2004

- Teach English in English and/or Korean from 1st to 6th graders
- Develop teaching and testing materials / design lesson plans on a daily basis
- Develop a school curriculum for English
- Supervise English native speaker teachers in English classes
- Run Summer English camps for 3rd to 6th graders
- In charge of "Classroom Improvement Research" program with emphasis on ESL education, selected by Department of Education, Seoul, Korea.
    - Research topic: "Improving basic communication skills in English as a second language" (Grant: $ 1000)
    - Collecting and analyzing data based on observation, survey, and interview with students
    - Publishing the final report in the Journal for K-12 teachers

International Center at Purdue University **West Lafayette, IN**
Korean Teacher 2010, 2008

- Teach Korean for beginners (2010)
- Teach Korean for intermediate learners (2008)
    - Designing the syllabi, class materials, and evaluation materials
    - Developing the curriculum for Korean classes by coming up with the overall sequence across the all levels of Korean classes
    - Discussing the entry and exit levels for each class with the director as well as the other Korean teachers of the International Center

Ewha Womans University **Seoul, Korea**
Research Assistant 2004, 2002

- Participate in Professor Min-kyeong Kim's research (2004)
    - Recruiting the research participants from the K-12 teachers
    - Collecting data with the survey questionnaire
    - Organizing the collected data in Excel Spread sheets
- Assist Professor Min-kyeong Kim's book publishing (2002)
    - Book title: "How Parents Can Utilize the Internet for Kids in Education"
    - Designing the purpose and actual content of the book

Ewha Elementary School **Seoul, Korea**
Intern Teacher 2002

- Teach 3rd graders, collaborating with the home room teacher
    - Assisting the home room teacher to design and create the class materials
    - Observing class activities and writing the daily report

PUBLICATIONS

You, Y. (submitted). Lexical proficiency in L2 writing: Theories, Measures, and Empirical Studies.

Maliborska, V., & **You, Y.** (manuscript in preparation). Group and individual L2 writing conferences: Instructor and student perspectives.

You, Y. (manuscript in preparation). Relationship between lexical proficiency and the quality of L2 spoken language.

You, Y. (manuscript in preparation). What is lexical proficiency: A new model based on past theories.

ACADEMIC PRESENTATIONS

Maliborska, V., & **You, Y.** (November 2014). Group and individual L2 writing conferences: Instructor and student perspectives, *Symposium of Second Language Writing*, Phoenix, AZ.

You, Y. (October 2013). Lexical Proficiency and L2 Oral Proficiency, *MwALT*, University of Michigan, Lansing, MI.

You, Y. (April 2013). Relationships between Lexical Proficiency and L2 Oral Proficiency, Poster presentation. *2013SLA Graduate Symposium*, The University of IOWA, Iowa City, IOWA.

You, Y. (March 2013). Relationships between Lexical Proficiency and L2 Oral Proficiency. *SLS/ESL Graduate Symposium*, Purdue University, West Lafayette, IN.

Maliborska, V. &**You, Y**. (March 2013). Individual conferences in the ESL writing classroom: What do students and teachers expect? *SLS/ESL Graduate Symposium*, Purdue University, West Lafayette, IN.

Maliborska, V., &**You, Y**. (March 2013). Toward understanding the role of individual conferences in L2 composition courses: Teacher and student's perspectives. *TESOL Doctoral Forum*, Dallas, TX.

You, Y. (October 2012). Understanding lexical proficiency in L2 writing. *INTESOL Conference,* Indianapolis, IN.

You, Y. (September 2012). Measuring lexical proficiency in L2 writing: Various issues and suggestions for future research. *Symposium of Second Language Writing,* Purdue University, West Lafayette, IN.

You, Y. (November 2011). Factors in vocabulary acquisition through reading, poster presentation. *INTESOL Conference,* Indianapolis, IN.

HONORS AND AWARDS

2013, May	Summer PRF for Research Grant ($ 2,975.06)	Purdue University West Lafayette, IN
2013, Feb	Quintilian award for the top ten percent of Fall 2012 instructor evaluations	Purdue University West Lafayette, IN
2011, Apr	1st place in Walter J. Johnson Award for the best graduate essay in English as a Second Language, entitled "Factors in Vocabulary Acquisition through Reading" ($ 100)	Purdue University West Lafayette, IN
2003, Mar	Teacher in charge of "Class Improvement Research" (Grant - $ 1000)	Department of Education Seoul, Korea
2002, Sep	1st place in Graduation Portfolio Competition	Ewha Womans University Seoul, Korea
2002, Feb	Academic Scholarship	Ewha Womans University Seoul, Korea

PROFESSIONAL AFFILIATIONS

American Association for Applied Linguistics (AAAL)

Teachers of English to Speakers of Other Languages (TESOL)

Indiana Teachers of English to Speakers of Other Languages (INTESOL)

International Language Testing Association (ILTA)

SERVICE

TESOL Proposal Reviewer, (2013). TESOL International Convention, Portland, OR

- Selected as a proposal reviewer for TESOL 2014 International Convention
- Reviewing 30 proposals after completing the training to be a reviewer through the program provided by TESOL

Treasurer, (2012 – 2013). ESL Graduate Student Organization, Purdue University, West Lafayette, IN

- Managing the fund of ESL Graduate Student Organization across incomes and outcomes
- Promoting fundraising to increase the fund
- Cooperating with Business Office of Student Organization (BOSO)

LANGUAGES

Korean	First language	Fluent in spoken and written
English	Second language	Fluent in spoken and written
Chinese	Second language	Beginner in spoken / intermediate in written

Relevant Courses during Graduate Programs

Practicum for Teaching College Students

- ENGL 50200 Practicum of Teaching Composition in English as a Second Language
- ENGL 50500 Teach First-Year Composition (I)
- ENGL 50500 Teach First-Year Composition (II)

Research Methodology

- EDPS 53300 Introduction to Research Methodology
- ENGL 61800 Quantitative Research
- ENGL 61900 Qualitative Research
- STAT 50100 Experimental Statistics I
- STAT 50200 Experimental Statistics II
- EDPS 63100 Research Design Seminar

English as a Second Language

- ENGL 51600 Theoretical Foundations in Teaching English as a Second Language
- ENGL 51800 Curriculum Development in Teaching English as a Second Language
- ENGL 58900 Lexical Proficiency in ESL
- ENGL 59000 Psychological Constructs in ESL
- ENGL 63000 Second Language Writing
- ENGL 67400 Language Testing

Language Studies

- ENGL 50600 Introduction to Linguistics
- ENGL 56500 Sociolinguistics
- EDCI 53000 English Language Development

Educational Psychology and Curriculum Design

- EDPS 53000 Advanced Educational Psychology
- EDPS 53100 Introduction to Measurement and Evaluation
- EDPS 59100 Human Growth and Development
- EDPS 54000 Gifted, Creative, and Talented Children
- EDPS 54200 Curriculum and Program Development in Gifted Education
- EDPS 54300 Teaching Thinking skills
- EDPS 54500 Social and Affective Development of Gifted Students
- CDFS 60100 Advanced Child Development