

## Lexical Richness in L2 Writing: How Much Vocabulary Do L2 Learners Need to Use?

Sun-Hee Kwon

(Pusan University of Foreign Studies)

**Kwon, Sun-Hee. (2009). Lexical richness in L2 writing: How much vocabulary do L2 learners need to use? *English Teaching*, 64(3), 155-174.**

While there has been accumulation of research into how L2 vocabulary knowledge is related to L2 writing, research on how many words L2 learners actually use in their writing has been poorly done. The goals of the present study are to examine lexical richness in the L2 texts with a comparison of L1 texts, and also to investigate how L2 lexical richness correlates with holistic ratings of L2 texts. Data was collected from three sources: essays of one-hundred twenty-two university students (N=122) for L2 intermediate proficiency texts, fifty TWE samples (N=50) for L2 high proficiency texts, and fifty sample articles from *Time* magazine (N=50) for L1 texts. Lexical richness of the texts was examined through measures of lexical diversity, lexical sophistication, and lexical density. The data analysis has shown that the D value, a measure of lexical diversity, was the only factor that distinguished different levels of texts. The measure of LFP, a measure of lexical sophistication, indicated that both L1 or L2 texts contained over 83% of words at the 2,000 most frequent word level. Lexical density has been shown to be a non-significant factor for holistic ratings of writing. Pedagogical implications and limitations of the present study are discussed.

### I. INTRODUCTION

Students often claim that what makes writing most difficult in second or foreign language (hereafter L2) is their lack of L2 vocabulary (Aliakbari, 2002; Leki & Carson, 1994; Raimes, 1985; Saville-Troike, 1984; Silva, 1992). L2 writers' heavy concern with vocabulary also enters into the judgment of the quality of the L2 writing; Grobe (1981) mentioned that what teachers perceived as good writing is closely associated with vocabulary diversity and Astika (1993) suggested that vocabulary proficiency is perhaps the best indicator of overall composition quality. Also Engber (1995) argued that the

diversity of lexical choice and the correctness of lexical form have a significant effect on the readers of timed- tasks written by L2 writers.

The question is then why L2 learners experience such a burden for vocabulary knowledge when writing, in spite of their relatively successful achievement in L2 academic contexts (i.e. passing the placement test, a high score in TOEFL). It might be simply that those who report their lack of vocabulary do not have sufficient vocabulary knowledge themselves. Nation (2001), however, suggested several possible reasons why vocabulary use in writing (or speaking) is more challenging than in reading (or listening): Productive use of words requires 1) more linguistic knowledge (the amount of knowledge explanation), 2) more practice (the practice explanation), 3) stronger semantic association (the access explanation), and 4) higher motivation to use (the motivation explanation). Given that even native-speaking (hereafter L1) writers tend to have difficulty retrieving the right words to express ideas, successful performance with rich and sophisticated vocabulary in L2 writing will always be a demanding job.

An attempt to better explain vocabulary use in language skills has considered a variety of issues, such as valid assessment of the discrete dimensions of lexical knowledge, or developing effective tasks for improving use of target words. In particular, recent studies have shown an increased interest in the relationship of productive use of vocabulary and L2 writing: Studies on L2 learners' productive vocabulary development in process writing (Muncie, 2002), the effect of explicit productive vocabulary instruction on ESL learners' writing (Lee, 2003), and the relationship between lexical retrieval and writing in ESL learners (Van Gelderen, Snellings, & De Glopper, 2004). All of these studies suggest that productive use of L2 vocabulary can influence L2 learners' writing process and product positively, or vice versa, writing itself may be a supportive way to improve productive lexical knowledge.

While the previous studies have suggested a number of important aspects for consideration with regard to productive use of vocabulary in writing, there remains a fundamental question; how much vocabulary do L2 learners need to make their writing good? This question presupposes the question of how many words, and how sophisticated words L2 learners actually use in their writing, and how their vocabulary use is different from that of L1 writers. To address these questions, this study examines L2 learners' vocabulary use through measuring a variety of lexical features. Also it explores the size of gap which exists between L2 texts and L1 texts in terms of vocabulary use. Lastly, it examines the correlation of vocabulary use and holistic scores of written essays from L2 learners.

## II. BACKGROUND

### 1. Vocabulary Use in L2 Writing

Examining vocabulary as one of factors affecting L2 writing has been less common than measuring grammatical accuracy or sentence complexity. Despite that, a number of studies have argued that there exists an on-going interaction between lexical knowledge and writing process (Cumming & Riazi, 2000; Hayes & Chenowith, 2001; Zimmermann, 2000), and in the case of L2 writing, poor lexical knowledge can be a detrimental factor to affect the overall writing process (Askita, 1993; Bardovi-Harlig & Bofman, 1989; Engber, 1995; Santos, 1988). Bardovi-Harlig and Bofman (1989) analyzed texts written by groups of ESL students at an American university. These students were divided into two groups: those below and above the minimum level of TOEFL required for admission to the university. They found that the students' syntactic expression was not significantly different in both groups, but the groups differed in their command of English morphology, suggesting this aspect of English is an area on which to focus teaching. A study by Askita (1993) reported in her study that vocabulary proficiency accounted for the largest amount of variance in total composition scores of ESL college students. Santos (1988) investigated the reactions of professors to errors in timed writings of nonnative-speaking students. The result indicated that professors generally judged the content and language independently, except for lexical errors, which were judged the most serious language errors. Santos suggested that the use of wrong words can often obscure the meaning of the essay, and this significantly affect reader judgment of the quality of an essay.

Vocabulary use in writing was also the focus in a study by Engber (1995). Engber conducted analysis of correlations between lexical proficiency and quality of the timed essay. She examined the relationship between various lexical measures of a piece of writing-lexical variation (diversity) with error, lexical variation without error, percentage of lexical error, lexical density, and holistic scores of writing quality. Engber found that there were significant correlations between lexical variation with error and a holistic measure of writing quality, and between lexical variation without error and a holistic measure of writing quality. In addition, error-free lexical variation was correlated best with the holistic writing score. These results indicate that the readers (raters) gave higher scores to writers who were able to use a variety of lexical resources correctly.

While the previous studies above have suggested that there exist positive relations between lexical knowledge and writing performance by L2 learners, these studies mainly focus on a relatively greater influence of lexical knowledge on L2 learners' writing than that of other linguistic features. For example, a study by Bardovi-Harlig and Bofman, and a study by Askita succeeded in addressing the importance of lexical knowledge on the

quality of writing, but they appeared to involve little consideration of what sub-features of vocabulary knowledge could be related to good writing. A study by Engber appeared to have been concerned with this issue, examining participants' lexical knowledge through a variety of lexical measures. However, it was limited to investigating how many various words learners produced, without examining the difficulty level of these words, which is another crucial aspect that may affect the L2 writing quality. Furthermore, Engber was not concerned with the lexical differences that would exist between L2 writing and L1 writing. These facts raise the need for detailed exploration of the lexical features of L2 writing, with a comparison of L1 writing.

## 2. Measures of Lexical Richness

Lexical features can be investigated by using several different measures, such as type-token ratio (TTR), word length, and lexical errors. While there are differing views on how to define sub-features involved in lexical knowledge, the broad consensus is that good writing depends on three features; how diverse words are, how difficult words are, and how many content words are produced in writing. That is, good writing has a greater lexical diversity, a greater lexical sophistication, and a greater lexical density.

Lexical diversity is often called lexical variation, or TTR. It is reasonable to expect that more proficient writers have a larger vocabulary that allows them to avoid repetition by using synonyms, and other kinds of related words. Engber (1995) found a significant relationship between lexical diversity and holistic ratings, and Linnarud (1986) found a significant difference between L2 writers and L1 writers. However, Hyltensam (1988) found no difference between L2 and near-L1 writers, and Laufer (1994) found no effect for instruction with advanced learners on this measure.

The lexical diversity is calculated by the ratio of different lexical words in the text (type) to all the lexical words used in the text (token), including repetitions of words that are used more once. The calculation of lexical diversity is commonly done by using TTR, but using TTR has been recently questioned due to its being sensitive to the text length. Currently, Malvern and Richard (2002) developed a new measure of lexical diversity based on TTR. This measure, *D* value, is not dependent on the text length since it is calculated by taking and averaging TTR for numerous random samples from the whole set of a text. *D*-value has been proved to be a relatively reliable and practical measure (Daller & Xue, 2007; Tidball & Treffers-Daller, 2007).

Lexical sophistication is related to the proportion of relatively difficult and low-frequency words in the text. This is another important aspect of lexical richness because it indicates the extent to which L2 learners use words in a precise and sophisticated manner. Linnarud (1986) found that L1 writers used significantly more

sophisticated vocabulary than second language writers, but found only a small correlation between the ratio of sophisticated words and holistic ratings of the writings.

A measure of lexical sophistication can be done by the Lexical Frequency Profile (LFP) that Laufer (1995) devised for measuring productive vocabulary. The LFP shows the percentages of words that learners use from different frequency levels in their writing. For example, given a composition consisting of 200 words, the LFP is calculated as follows: among 200 words, 150 words belong to the first 1,000 most frequent words, 20 to the second 2,000, 20 to the academic word list, and 10 are not in any list. Then numbers are converted to percentages, from out of the total 200 words. The entire calculation is done by a computer program which matches vocabulary frequency lists based on Nation's vocabulary list (1984).

Lexical density refers to a proportion of content words, compared to function words, in the text. Content words are lexical words such as nouns, verbs, adjectives, and adverbs, while function words are grammatical words such as prepositions, determiners, or auxiliaries. Lexical density is one of the major characteristics that distinguishes written form from spoken form and a higher lexical density reflects that ideas are presented in a more concentrated way (Read, 2000). Waller (1993) found that texts with a lexical density above 50 percent were mostly written by L1 writers, suggesting that lexical density is considered a crucial component for good writing. However, Hyltenstam (1988) showed that lexical density may not be the best measure for writing quality since it is possible to get a high density score with a small vocabulary.

### III. RESEARCH DESIGN

#### 1. Research Questions

The research questions are presented as follows:

- 1) To what degree do L1 texts and L2 texts differ in terms of lexical diversity, lexical sophistication, and lexical density?
- 2) To what degree do L2 intermediate proficiency texts and L2 high proficiency texts differ in terms of lexical diversity, lexical sophistication, and lexical density?
- 3) To what degree are measures of lexical diversity, lexical sophistication, and lexical density related to the holistic ratings of L2 texts?

#### 2. Data Collection

- 1) L2 Texts

### (1) Essays Written by L2 University Students

One-hundred twenty-two essays of L2 university students (N=122) were collected in order to explore lexical richness of L2 intermediate proficiency texts. L2 university students who participated in the study consisted of four intact classes enrolled in English reading courses at a university in South Korea. The participants in Class A and B were all first-semester freshmen majoring in English Language, and the participants in Class C and D were juniors majoring in English Language when the study was conducted. All of the students in these four groups were considered to be at the L2 Intermediate proficiency group level, since the average score of their essays was 2.99 out of 6. The freshmen in Class A and B were assigned to L2 Intermediate-Low texts (N=70) and juniors in Class C and D were assigned to L2 Intermediate-High texts (N=52), based on their average writing scores (M=2.84 for the freshman group, M=3.19 for the junior group). Participants in each group were assumed to have had a similar English education background since most of them had been taught English as a compulsory subject for at least six years before entering the university in South Korea. However, the participants were considered to vary in their general English proficiency from upper-beginning to advanced level. Both groups of L2 Intermediate-Low and L2 Intermediate-High texts were taught by the same instructor in the same instructional way during the time of the study.

During class, participants in each group were requested to write a short essay using different writing prompts. The participants in the group of L2 Intermediate-Low texts were asked to write an essay about what sorts of Korean food they would recommend for their foreign friends, and the participants in the group of L2 Intermediate-High texts were asked to write about whether they agree or disagree with same sex marriage. These writing prompts had been selected for two reasons; first, they had been discussed as reading topics during class, and second, both required a similar writing style (persuasive writing). The participants were asked to do this task individually for 30 minutes, during class time, and submit it to the instructor upon task completion.

### (2) TWE sample essays

Fifty sample essays (N=50) from TOEFL writing test (TWE), which scored highest in the test, were obtained through publicly available resources, such as TWE sample essay books. These TWE essays were selected to examine lexical richness of L2 learners who have high proficiency in the writing. The topics of selected essays were mostly about issues in daily life, such as the topics of *why go to university?* *money & success*, asking the essays to persuade readers, as the topics given in L2 texts.

## 2) L1 Texts

Fifty essays/articles in English (N=50) were randomly selected from the online version of *Time* magazine (www.time.com). *Time* magazine is an authentic text, mostly written for and by native speakers of English, and essays/articles of this magazine use formal writing styles. The topics of articles selected varied from political issues to drama reviews. These articles were examined as a baseline data of L1 texts in comparing lexical richness of L2 texts and L1 texts.

## 3. Data Analysis

### 1) Lexical Diversity

Measurement for lexical diversity in this study was carried out using *D*-values (Malvern & Richards, 2002). It has been argued that when compared with TTR, which is the typical measure of lexical diversity, *D*-values are relatively independent of sample size, thus allowing valid comparisons between varying quantities of linguistic data. To obtain *D*-values, the collected essays from three different sources—L2 participants, TWE samples, and *Time* magazine—were condensed into text files and put into the CHAT (Codes for the Human Analysis of Transcripts) (MacWhinney, 2000a, 2000b) format to apply CLAN (Computerized Language Analysis) command *vocd* that produces the value for *D*. A higher score for *D* represents greater lexical diversity.

### 2) Lexical Sophistication

To examine the extent to which L2 texts and L1 texts included sophisticated words, data was analyzed using the Lexical Frequency Profile (Laufer & Nation, 1995). LFP shows the percentage of words at different word frequency levels, and the values of LFP in this study were used to measure five different frequency levels: 1,000, 2,000, 3,000, 5,000, and off-list levels, based on British National Corpus (BNC). Essays obtained from three different sources were formatted into text files, and analyzed using through *vocabprofile* program (Cobb, 2002) which automatically calculates the percentages for word frequency levels.

### 3) Lexical Density

Lexical density is calculated by the proportion of content words in the text. Content words include nouns, verbs, adjectives, and adverbs, and the higher percentage of content words the text has, the greater its lexical density. The lexical density of collected essays was also calculated using the *vocabprofile* program, which indicated the percentages for content words in the written texts.

#### 4) L2 Essays Scoring

The L2 essays written by participants in this study were scored based on the rubric of TWE. It consists of six scales from 1 to 6, and each with a number of definitions of how well competence in writing, for both the rhetorical and syntactic categories is demonstrated. The rubric used in this study was exactly the same as the one used for the official TWE test. Two native speakers of English rated the essays using this rubric. Upon completion the rating, the results were cross-checked. In the case of discrepancies in the ratings, the essay in question was discussed by the raters, and a single score was agreed upon for on the essay.

## IV. FINDINGS

### 1. A Comparison of Lexical Richness in L2 Texts and L1 Texts

Table 1 shows the means and standard deviations for D value, LFP, and lexical density for the three different texts—L2 Intermediate, L2 High, and L1 texts. The descriptive measures show differences between the texts in the expected direction; the mean scores of the L1 texts were higher than those of the other two texts, and the mean scores of the L2 High texts were higher than those of the L2 Intermediate texts on most of lexical measures. For the 1,000 values of LFP, however, it appeared that the levels of texts were not related to the lexical outcomes; while L1 texts produced the lowest percentage of words at 1000 frequency level, L2 High texts produced the highest percentage of words at 1000 frequency level, suggesting that essays by L2 High texts contained more easy word than those of the L2 Intermediate texts. In Figure 1, a comparison of D values and ratios of 1,000 + 2,000 in the LFP for three different texts were presented, to show the differences clearly. The measures of lexical density also show that there was no consistent relation between the levels of texts and lexical outcomes; that is, L1 texts and L2 Intermediate texts yielded similar levels of lexical density, along with far less values from L2 High texts. It suggests that L2 High texts included less content words than the other two texts. As a follow-up test, one-way ANOVAs (Table 2) were conducted to see if there were statistically significant differences between the three texts on different lexical measures.



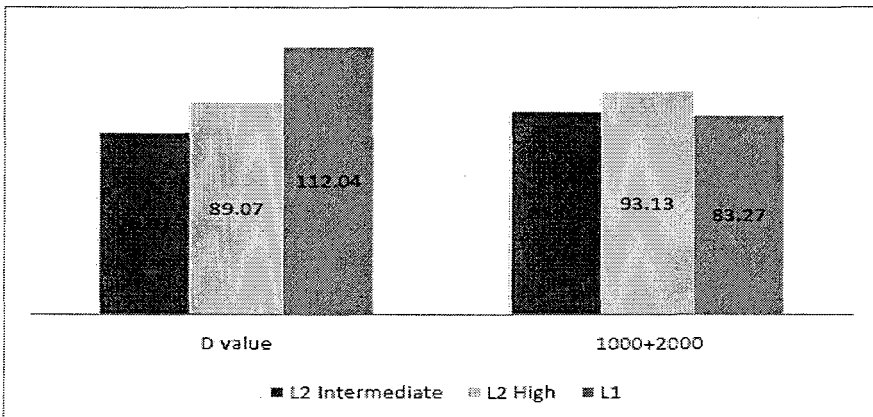
**TABLE 1**  
**Descriptive Statistics of Lexical Richness of L2 Intermediate, L2 High, and L1 Texts**

	L2 Texts				L1 Texts	
	Intermediate		High		Articles in <i>Time</i> magazine (N=50)	
	L2 participants (N=122)	TWE samples (N=50)	M	SD		
	M	SD	M	SD	M	SD
D value	76.07	25.43	89.07	23.50	112.04	22.23
1000	77.77	9.09	85.24	4.36	74.24	5.56
2000	7.25	2.88	7.89	2.64	9.03	1.96
3000	1.61	1.56	2.50	1.40	3.21	1.78
5000	2.56	1.68	.63	.54	1.20	1.02
Density	.59	.05	.51	.03	.58	.04

**TABLE 2**  
**ANOVA on Each Measure of Lexical Richness for L2 Intermediate, L2 High, and L1 Texts**

	df	F	Sig.
D value	2	39.04	.000
1000	2	28.34	.000
2000	2	8.017	.000
3000	2	19.23	.000
5000	2	41.70	.000
Density	2	52.72	.000

**FIGURE 1**  
**A Comparison of D values and Ratios of 1000+2000 in the LFP for L2 Intermediate, L2 High and L1 Texts**



Since the measures of ANOVA showed that there were statistically significant differences between the three texts on all of lexical measures, the post hoc Scheffe tests were run to examine which comparisons contributed to the main differences. The post hoc tests in Table 3 show that the main differences were mostly from the gap between L2 texts and L1 texts. That is, L1 texts performed significantly better than the L2 Intermediate texts on all the lexical measures, except for lexical density. In comparison of L2 High texts and L1 texts, L1 texts performed better than L2 High texts on the D value (L1 text at higher value), at 1,000 frequent level (L1 text with low percentage) and on lexical density (L1 text at higher value) at statistically significant levels. However, it seems that L2 High texts and L1 texts did not differ on 2,000, 3,000, and 5,000. These results indicate that in general, L1 texts contained more sophisticated words than L2 High texts, but as the word frequency increases, the gap between the two texts decreases.

**TABLE 3**  
Scheffe Post Hoc Comparisons of L2 Intermediate, L2 High, and L1 Texts

	Groups	Diff.	SD	Sig.
D value	L2 Intermediate – L2 High	-13.22	4.08	.006
	L2 Intermediate – L1	-35.97	4.09	.000
	L2 High – L1	-22.74	4.86	.000
1000	L2 Intermediate – L2 High	-7.46	1.25	.000
	L2 Intermediate – L1	3.41	1.25	.027
	L2 High – L1	10.88	1.49	.000
2000	L2 Intermediate – L2 High	-.63	.44	.357
	L2 Intermediate – L1	-1.77	.44	.000
	L2 High – L1	-1.13	.52	.103
3000	L2 Intermediate – L2 High	-.88	.26	.005
	L2 Intermediate – L1	-.15	.26	.000
	L2 High – L1	-.70	.31	.087
5000	L2 Intermediate – L2 High	1.92	.22	.000
	L2 Intermediate – L1	1.35	.22	.000
	L2 High – L1	-.56	.27	.120
Density	L2 Intermediate – L2 High	.08	.01	.000
	L2 Intermediate – L1	.01	.01	.218
	L2 High – L1	-.06	.01	.000

## 2. A Comparison of Lexical Richness in L2 Texts

To compare the differences among the L2 texts, the L2 Intermediate proficiency texts were classified again into two subtexts: the L2 Intermediate-Low texts from 70 freshmen, the L2 Intermediate-High texts from 52 Juniors. Thus, three groups in total, such as the L2 Intermediate-Low texts (L2 IL), the L2 Intermediate-High texts (L2 IH), and L2 High texts (L2 High) were compared with one another using the different lexical measures. Table 4 shows the means and standard deviations for D value, LFP, and lexical density for each of the three texts. The results show differences between the texts in the expected direction on only D value. That is, on only D value measure, the mean score of the L2 High texts was higher than those of the other two texts, and the mean scores of the L2 Intermediate-High texts were higher than that of the L2 Intermediate-Low texts. For other measures, such as from 1,000 to 5,000, and density measure, it appeared that there was no consistent relation between the levels of texts and lexical outcomes. In particular, it is interesting to note that L2 Intermediate-Low texts performed better than, or at least similar to, the other two texts on almost all measures, except for D value. This is shown more clearly in Figure 2 with a comparison of D values and ratios of 1,000 + 2,000 in the LFP for three different texts. As a follow-up test, one-way ANOVAs (Table 5) were conducted to see if there were statistically significant differences between the three texts on different lexical measures.

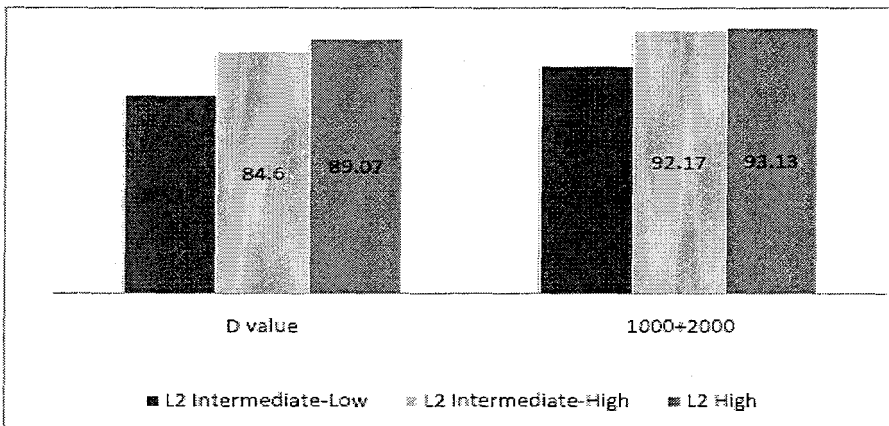
**TABLE 4**  
**Descriptive Statistics of Lexical Richness of L2 Texts**

	L2 Texts					
	Intermediate				High	
	Intermediate-Low (L2 IL, N=70)		Intermediate-High (L2 IH, N=52)		TWE samples (L2 High, N=50)	
	M	SD	M	SD	M	SD
D value	69.73	29.69	84.60	14.53	89.07	23.50
1000	71.88	7.25	85.69	3.26	85.24	4.36
2000	7.83	3.35	6.48	1.84	7.89	2.64
3000	2.01	1.94	1.07	.54	2.50	1.40
5000	3.07	1.96	1.88	.83	.63	.54
Density	.60	.05	.58	.03	.51	.03

**TABLE 5**  
ANOVA on Each Measure of Lexical Richness for L2 Texts

	df	F	Sig.
D value	2	10.96	.000
1000	2	126.60	.000
2000	2	4.51	.012
3000	2	12.32	.000
5000	2	46.42	.000
Density	2	55.53	.000

**FIGURE 2**  
A Comparison of D value and Ratios of 1000+2000 in the LFP for L2 IL, L2 IH, and L2 High Texts



The measures of ANOVA showed that there were statistically significant differences between the three texts on all of the lexical measures, thus, the post hoc Scheffe tests were run. The post hoc tests in Table 6 show interesting results; in a comparison of the L2 Intermediate-High texts and the L2 High texts, there is no statistically significant difference between the texts on D-value and at 1,000 frequency level. This indicates that L2 Intermediate-High and L2 High texts produced diverse words at a similar level, and a similar percentage of words at 1,000 frequency word levels. In a comparison of L2 Intermediate-Low and L2 High texts, L2 High texts performed better than the L2 Intermediate-Low texts on D value, but, for the other measures, L2 Intermediate-Low texts performed better than, or were similar to the L2 High texts. This was the same case in a comparison of L2 Intermediate-Low and L2 Intermediate-High; L2 Intermediate-High texts performed better than the L2 Intermediate-Low on D value, but, for the other measures, L2 Intermediate-Low texts performed better than the L2 Intermediate-High texts.

These results suggest that except for D value, there is no consistent relation between the levels of texts and lexical outcomes in L2 texts.

**TABLE 6**  
**Scheffe Post Hoc Comparisons for L2 Texts**

	Groups	Diff.	SD	Sig.
D value	L2 IL – L2 IH	-14.87	4.42	.004
	L2 IL – L2 High	-19.56	4.47	.000
	L2 IH – L2 High	-4.69	4.78	.619
1000	L2 IL – L2 IH	-13.80	1.00	.000
	L2 IL – L2 High	-13.35	1.01	.000
	L2 IH – L2 High	.45	1.08	.918
2000	L2 IL – L2 IH	1.35	.50	.030
	L2 IL – L2 High	-.06	.51	.992
	L2 IH – L2 High	-1.41	.54	.038
3000	L2 IL – L2 IH	.93	.27	.003
	L2 IL – L2 High	-.48	.27	.207
	L2 IH – L2 High	-1.42	.29	.000
5000	L2 IL – L2 IH	1.18	.25	.000
	L2 IL – L2 High	2.43	.25	.000
	L2 IH – L2 High	1.24	.27	.000
Density	L2 IL – L2 IH	.03	.01	.016
	L2 IL – L2 High	.09	.01	.000
	L2 IH – L2 High	.06	.01	.000

### 3. Lexical Richness and Holistic Scores of L2 Texts

To investigate the relations between lexical knowledge and the quality of L2 texts, first, the mean scores of the two L2 Intermediate texts were compared. The average scores of L2 Intermediate-Low and L2 Intermediate-High texts were 2.84 for L2 IL, and 3.19 for L2 IH, out of a total possible score of 6. An ANOVA analysis (Table 7) shows that there was a statistically significant difference between the two texts regarding the holistic ratings of the texts. Next, to examine how much lexical knowledge contributes to the quality of L2 texts, the correlations between the holistic ratings of each of the L2 IL, and L2 IH texts were examined along with the other lexical measures. The results in Table 8 show that for both L2 IL, and L2 IH texts, the correlations exist only for D value. This suggests that L2 texts having more diverse words may have great probability for

obtaining higher scores in holistic rating. In terms of the significance of the correlation between the 3,000 frequency level and the holistic ratings for the L2 Intermediate-Low text, it can be assumed that a higher percentage of words at the 3,000 frequency level may contribute to the higher score of writing. However, this should be argued to the limited degree since only the L2 Intermediate-Low texts yielded this result.

**TABLE 7**  
**A Comparisons of Holistic Ratings of L2 Texts**

	Holistic Ratings of L2 Texts		ANOVA	
	M	SD	F	Sig.
L2 Intermediate-Low (N=70)	2.84	.87	5.37	.022
L2 Intermediate-High (N=52)	3.19	.74		

**TABLE 8**  
**Correlation of Lexical Richness and Holistic Ratings of L2 Texts**

		D value	1000	2000	3000	5000	Density
Holistic Ratings of L2 Texts	L2 Intermediate-Low (N=70)	.597**	-.087	-.014	.332**	-.005	-.027
	L2 Intermediate-High (N=52)	.458**	.145	-.107	-.222	.031	.038

\*\* Correlation is significant at the 0.01 level (2-tailed).

## V. DISCUSSIONS & CONCLUSION

### 1. Lexical Richness in L2 Texts and L1 Texts

It has been shown in this study that the differences between L2 texts and L1 texts are significant in terms of lexical diversity and lexical sophistication. The D value that represents lexical diversity is the highest for L1 texts, followed by L2 High proficiency texts and L2 Intermediate proficiency texts. For lexical sophistication, the L1 texts showed the lowest percentage of the first most frequent 1,000 words, suggesting that the essays written by L1 writers contain fewer easy words than those written by L2 writers. In terms of lexical density, however, both L1 and L2 texts showed higher than .50. This indicated that both L1 and L2 texts contained content words at a moderately high level.

The results from the comparisons of L1 texts and L2 texts in this study partially support the previous studies. First, this research demonstrates the statistically significant

differences between L1 texts and L2 texts in terms of lexical diversity and lexical sophistication, which add strong evidence that there exists a variety of lexical knowledge available to L1 and L2 writers (Arnaud, 1984, cited from Read (2000); Linnarud, 1986). Next, in this study, lexical density, as a lexical feature is, not shown to be a factor that distinguishes how L1 and L2 texts differ in their use of lexical features. This evidence supports the findings from Hyltenstam (1988), but it is inconsistent with findings from Linnarud (1986), and Waller (1993). It seems that lexical density was not as strong a factor as lexical diversity and lexical sophistication with regard to the difference it created between L1 texts and L2 texts in this study.

Another significant finding is that despite significant lexical differences between L1 texts and L2 texts, around 83% of words in the L1 texts belonged to the most frequent 2,000 words. It appears that in this study, the L1 texts had more sophisticated vocabularies than the L2 texts, but the ratios of words beyond the most frequent 2,000 word level in the L1 texts showed a significant declination, just as the L2 texts rarely contained the words beyond the most frequent 2,000 words. This is rather surprising in a sense that the L1 texts analyzed in this study were mostly the news articles and essays from *Time* magazine, which were assumed to be refined texts written in English. This finding suggests that even in L1, productive vocabulary use in the written texts may remain at a basic level, and that either in L1 or L2, writers might not be able to fully utilize words that they have learned before. Also drawn from this finding, it is suggested that good writing may not necessarily require exceptionally difficult or sophisticated words. The fact that the L1 texts analyzed in this study infrequently use the words beyond the most frequent 2,000 implies that words that convey the ideas properly could be selected in the range of the most frequent 2,000 words, and still maintain standards of good writing.

## 2. Lexical Richness in L2 Texts and Its Relation to Holistic Ratings

In comparisons of the three different levels of L2 texts, it has been shown that there was no significant difference between the L2 Intermediate-High proficiency texts written by L2 juniors in the university and the L2 High proficiency texts that earned the perfect score in the TWE test, in terms of lexical diversity, and lexical sophistication. The texts written by juniors and TWE texts included a similar number of diverse words, and similar number of words at the most frequent 2,000 word level. This finding suggests that L2 university students who have been exposed to intensive English education for two or three years in Korea might have sufficient vocabulary size necessary for obtaining good scores in standardized language writing tests. Most of the junior students who participated in this study received formal education in English

reading, grammar, speaking and writing as core subjects during their first two years in the university. The fact that these students produced words at similar levels of lexical diversity and sophistication to those of TWE essays indicates that for the purpose of test-taking only, these learners might have an adequate range of vocabulary size.

Despite similar levels of lexical diversity and lexical sophistication between the L2 Intermediate-High texts and L2 High texts, these two texts show very different qualities when it comes to holistic ratings of the texts. The L2 Intermediate-High texts received 3.19 out of 6 points on average, and this is considerably lower than the scores of the L2 High texts. Possible reasons for such low holistic ratings might be first, due to the TWE evaluation rubrics including consideration of other categories such as the organization of ideas and syntactic variety. This would result in low ratings of L2 junior-texts since they are being judged on skills other than vocabulary use. That is, although lexical variables may function as a strong indicator of writing quality, lexical variables alone might not be sufficient to discriminate between good and poor writing. Second, L2 students with low ratings were able to utilize words in terms of vocabulary size, to the same degree as L2 learners with high ratings in TWE, however, those with low ratings might use words ineffectively or unsuccessfully. In other words, L2 students with low ratings produced a relatively large number of words, but those words might be presented in the wrong contexts, or used with essential information missing.

In terms of the question of how different lexical measures correlate with ratings of L2 texts, lexical diversity turns out to be closely correlated to the holistic rating of writing. This means that the use of diverse words affects the quality of writing positively at a significant level. This result provides support for the findings from Engber (1995) that indicated the positive correlations between lexical diversity and holistic ratings of writing. However, it is inconsistent with the result from Daller and Phelan (2007) showing that lexical diversity yields only moderate and non-significant correlations with ratings of writing. In this study, D value, a measure that was developed to measure lexical diversity, was the only factor that distinguished between various L2 texts, as well as between L2 texts and L1 texts. For the lexical sophistication, LFP value does not yield any significant correlation with ratings. This implies that the occurrence of rare and difficult words did not affect ratings of writing in this study. In addition, lexical density, as another measure of lexical richness, appeared to be a non-significant variable for holistic ratings of writing. In particular, the fact that the L2 texts with higher lexical diversity had lower lexical density supports the argument that lexical density can be insensitive to the levels of written texts when the texts include less words (Hyldenstam, 1988).



### 3. Pedagogical Implications and Limitations

The implications from the current study's findings are that first, L2 learners need to be aware that they may not have to necessarily employ a large proportion of rare or low-frequency words to become proficient in writing. There is no doubt that the use of sophisticated words is desirable for good writing, but it also should be pointed out that sophisticated words alone do not guarantee the quality of writing. The results of the current study have clearly shown that even L1 speakers or L2 learners who have high proficiency in writing, primarily use vocabulary at the 2,000 most frequent word level. Thus, it is suggested that if L2 learners master how to use the most 2,000 frequent words fully, they may possess a crucial tool for having a good command of their written outcomes in L2.

Next, L2 learners may benefit from knowing that using diverse words can enhance their quality of writing to a great extent. The current study has shown that there are clear-cut differences between L1 and L2 texts, and between L2 Intermediate and L2 High proficiency texts, in terms of lexical diversity. This result suggests that it could be of great help for L2 learners to practice synonyms and expressions that have the same or similar meaning to target words, and classroom instruction can also provide instructional chances to improve L2 learners' lexical diversity.

Lastly, from the results of the present study, we find support for the argument that knowing a word involves more than just word recognition or production; knowing a word may ultimately be concerned with how to use the word (Nation, 2001). In other words, to achieve high proficiency in writing, learners may need to have the ability to know in what patterns the words are used together, and how to arrange the words to make ideas clear and organized, rather than just recognizing or producing words in isolated contexts. This ability can be connected to so-called knowing 'structures' and 'sentences patterns' in the conventional sense, but it is also considered as lexical competence in some other approaches, such as the lexical approach. Thus, it is essential to be aware that lexical competence can be enhanced not by just knowing separate words, but by knowing how, and where those words are used together.

There are a number of limitations in this study. First, the fact that measuring of lexical diversity, sophistication, and density was carried out with only one method must be cautioned. Measures of these lexical features can be attempted with different methods, which might not necessarily generate the same results. Thus, if different methods had been employed to cross-check with one another, measuring lexical richness might have resulted in a greater reliability. Second, the results of this study were based on relatively limited number of texts, which may not be large enough to generalize the results. To get a more comprehensive and reliable picture of the lexical richness of L1 and L2 texts, as well as the relationship between L2 texts and the quality of writing, additional lexical

measures should be examined along with multiple measures of a larger number of texts.

## REFERENCES

- Aliakbari, M. (2002). Writing in a foreign language: A writing problem or a language problem? *Journal of Pan-Pacific Association of Applied Linguistics*, 6, 157-168.
- Arnaud, L. (1984). The lexical richness of L2 written productions and the validity of vocabulary tests. In T. Culhane, C. Klein-Braley & K. Stevenson (Eds.), *Practice and problems in language testing* (pp. 14-28). Colchester: University of Essex, Department of Language and Linguistics.
- Astika, G. (1993). Analytical assessments of foreign language students' writing. *RELC Journal*, 24, 61-71.
- Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition*, 11, 17-24.
- Cobb, T. (2002). *Web Vocabprofile*. Retrieved Oct 11, 2008, from the World Wide Web: <http://www.lextutor.ca/vp/>.
- Cumming, A., & Riazi, A. (2000). Building models of adult second-language writing instruction. *Learning and Instruction*, 10, 55-71.
- Daller, H., & Xue, H. (2007). Lexical richness and the oral proficiency of Chinese EFL students. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 150-164). Cambridge: Cambridge University Press.
- Daller, H., & Phlean, D. (2007). What is in a teacher's mind?: Teaching ratings of EFL essays and different aspect of lexical richness. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 234-244). Cambridge: Cambridge University Press.
- Engber, C. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing*, 4, 139-155.
- Grobe, G. (1981). Syntactic maturity, mechanics, and vocabulary as predictors of quality ratings. *Research in the Teaching of English*, 15, 75-85.
- Hayes, J., & Chenoweth, N. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication*, 18, 80-98.
- Hyltenstam, K. (1988). Lexical characteristics of near-native second-language learners of Swedish. *Journal of Multilingual and Multicultural Development*, 9, 67-84.
- Laufer, B. (1994). The lexical profile of second language writing: Does it change over time? *RELC Journal*, 25, 21-33.

- Laufer, B. (1995). Beyond 2000: A measure of productive lexicon in a second language. In L. Eubank, L. Selinker & M. S. Smith (Eds.), *The current state of interlanguage* (pp. 265-272). Amsterdam: John Benjamins.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16, 307-322.
- Lee, S. H. (2003). ESL learners' vocabulary use in writing and the effects of explicit vocabulary instruction, *System*, 31, 537-561.
- Leki, L., & Carson, J. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL Quarterly*, 28, 81-101.
- Linnarud, M. (1986). *Lexis in composition: A performance analysis of Swedish learners' written English*. Sweden: Gleerup.
- MacWhinney, B. (2000a). *The CHILDES project: Tools for analyzing talk, Vol.1: Transcription format and programs* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- MacWhinney, B. (2000b). *The CHILDES project: Tools for analyzing talk, Vol.2: The database* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19, 85-104.
- Muncie, J. (2002). Process writing and vocabulary development: Comparing lexical frequency profiles across drafts. *System*, 30, 225-235.
- Nation, P. (Ed.) (1984). *Vocabulary lists: Words, affixes and stems*. New Zealand: English University of Wellington, English Language Institute.
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Raimes, A. (1985). What unskilled ESL students do as they write: A classroom study of composing. *TESOL Quarterly*, 19, 229-258.
- Read, J. (2000). *Assessing vocabulary*. New York: Cambridge University Press
- Santos, T. (1988). Professor's reaction to the academic writing of nonnative-speaking students. *TESOL Quarterly*, 22, 69-90.
- Saville-Troike, M. (1984). What really matters in second language learning for academic achievement? *TESOL Quarterly*, 18, 199-219.
- Silva, T. (1992). L1 vs. L2 writing: ESL graduate students' perception. *TESL Canada Journal*, 10, 27-48.
- Tidball, F., & Treffers-Daller, J. (2007). Exploring measures of vocabulary richness in semi-spontaneous French speech. In H. Daller, J. Milton, & J. Treffers-Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 133-149). Cambridge: Cambridge University Press.
- Van Gelderen, A., Snellings, P., & De Glopper, K. (2004). The effect of enhanced lexical retrieval on second language writing: A classroom experiment. *Applied*

*Psycholinguistics*, 25, 175-200.

Waller, T. (1993). Characteristics of near-native proficiency in writing. In H. Ringbom (Ed.), *Near-native proficiency in English* (pp. 183-293). Åbo, Finland: Akademi Univerisity, English Department.

Zimmermann, R. (2000). L2 writing: Subprocesses, a model of formulation and empirical findings. *Learning and Instruction*, 10, 73-99.

Applicable levels: secondary and tertiary L2/ESL/EFL education

Key words: L2 vocabulary richness, vocabulary measure, L2 writing

Sun-hee Kwon

The College of English, Pusan University of Foreign Studies

15 SeokPo-Ro, Nam-Gu, Busan 608-738, Korea

Tel: (051) 640-3050

Fax: (051) 640-3009

Email:skwon@pufs.ac.kr

Received in June 2009

Reviewed in July 2009

Revised version received in August 2009