

## Influences of Writing Genres and Proficiency on Lexical Diversity<sup>1)</sup>

Lee, Eun-Hee\* · Oh, Hee Jeong\*\*,<sup>†</sup>

\* Associate Professor, School of General Education, Kyung Hee Cyber University (First author, E-mail: ehlee1@khcu.ac.kr)

\*\* Associate Professor, General English Department, Seoul Women's University (Corresponding author, E-mail: hjoh@swu.ac.kr)

(Received March 13, 2018; Revised April 18, 2018; Accepted April 25, 2018)

**ABSTRACT.** Providing ESL/EFL writers with freedom to choose a writing topic has been regarded as important for promoting their learning motivation and improving their quality of writing. Lexical diversity is one of the critical elements in evaluating writing quality, and several factors such as major/content knowledge, age, gender, or language proficiency influence the overall quality of writing even when L2 learners freely choose their writing topics. This study explored the relationships between overall language proficiency and lexical diversity, and between writing genres and lexical diversity when Korean EFL writers had freedom to choose their topics. Seventy-two university students participated in this study, and they submitted two essays of different writing genres during one semester. Their writings were analyzed in terms of the number of types, tokens, TTR, and VocD using a computer program, CLAN. It was found that VocD values differed depending on writing genres, but overall proficiency did not contribute to the differences. The length of essays was not related to the level of VocD either. The findings suggest that in order to improve the lexical diversity of EFL writers, educators need to provide them with diverse writing activities that include different types of writing genres.

**Key words:** Writing genres, Proficiency, Lexical diversity, TTR, VocD

### I. Introduction

Writing is a difficult skill to acquire for most ESL/EFL learners. To ease their difficulties, researchers (Durán, Malvern, Richards, & Chipere, 2004; Zamel, 1983) analyzed learners' writings and realized that analyzing the writing process rather than the final product provided more educational insights. Among several issues related to the writing process, it was found that providing ESL/EFL writers with thought-provoking topics could stimulate their writing activities and motivation (Allwright & Bailey, 1994; Lee & Suh, 1998; Savignon, 1991; Wang, 2010).

Non-native writers choose their topics on the basis of their content knowledge; when they are familiar with a topic, they feel more comfortable to write about it (Gradwohl & Schumacher, 1989; Wang, 2010). This psychological stability positively influences writing performance, so the quality of writings is better when they have freedom to choose a topic. Although language learners in most studies have been given freedom of topic, their writings demonstrated proficiency differences even when every student was given the same prompt (Gradwohl & Schumacher, 1989; Lee, 2004; Lee & Anderson, 2007; Wang,

2010).

Providing freedom to choose a writing topic may not be good enough to facilitate writing activities and to improve writing ability, but there is a lack of empirical research about issues after providing the freedom. This study investigated what really happened after language learners freely chose a topic, and what the roles of writing genres and overall language proficiency were in writing performance. Lexical diversity was measured to evaluate writing proficiency, and the relationships between overall proficiency and lexical diversity, and between writing genre and lexical diversity were explored. The research questions were as follows:

- 1) Is there a relationship between Korean EFL learners' overall language proficiency and lexical diversity?
- 2) Is there a relationship between writing genre and lexical diversity?

### II. Literature Review

#### 1. Importance of Writing Topics

ESL/EFL writers write better when they have freedom to choose their topic (Gradwohl & Schumacher 1989; Wang, 2010), and the generality of the topic should be guaranteed across all test-takers on a large-scale test (Lee & Anderson, 2007). A certain topic which favors a certain

1) A part of this study was presented at the 37<sup>th</sup> Thailand TESOL International Conference, January, 2017, Bangkok, Thailand.

group of test-takers will damage the reliability of the test. Gradwohl and Schumacher explored the relationship between content knowledge and topic choice in writing and found that participants showed more knowledge about the self-selected topic than teacher-selected or don't-want topics. The study demonstrated that content knowledge influenced not only the writing product but also the writing process. Having an option to choose their own writing topic can also promote writers' motivation and engagement in learning. Lo and Hyland (2007) claimed that topic choice empowered low-achieving young ESL writers in Hong Kong and enhanced their level of confidence, motivation, and engagement.

In order to examine the influencing factors to the writing performance in large-scale writing tests, Lee and Anderson (2007) examined the effect of writing topics after controlling their participants' English proficiency. They examined the influences of academic topics, participants' majors, background knowledge, and general English proficiency on writing performance. The results indicated that writers' majors did not affect writing performance, while academic topics did. When non-native writers had freedom to choose their topics, the quality of writing was better. When writers have prior knowledge related to a topic, they also write better (Barry, Nielsen, Glasnapp, Poggio, & Sundbye, 1997; Lee & Anderson, 2007; Stapa, 2001, cited in Lee & Anderson, 2007; Tedick, 1988, 1990).

On the other hand, Lee (2004) found that the effect of background knowledge was different depending on participants' majors. She divided her participants into four groups depending on their majors: business, humanities, technology, and life sciences. The results indicated that only participants whose majors were business and life sciences showed a field-specific topic effect; their writings about topics related to their majors were significantly better than those about general topics. The other two groups showed no differences in different types of writing topics. As seen, many factors influence language learners' writing performance, so language learners need more guidelines and instruction to improve their writing. Writing genres have been known to contribute to language learners' performance (Berman & Verhoeven, 2002; Bhatia, 2002; Sadeghi & Dilmaghani, 2013; Watanabe, 2016; Woerfel & Yilmaz, 2011), and overall language proficiency has also been shown to be associated with writing performance (Cumming, 1989; Zamel, 1983). Therefore, this study sought to explore the effect of writing genres and overall proficiency on writing performance.

## 2. Lexical Diversity in Writing

To evaluate proficiency in writing, lexical diversity was measured. Durán et al. (2004) argued that higher lexical diversity generally signified more advanced proficiency, and lexical diversity is often considered to represent a level of overall text quality. In order to measure lexical diversity, researchers have often used Type-Token Ratio (TTR), a ratio between type and token, but the value of TTR is influenced by the number of total words in a sample, so nowadays researchers (Durán et al., 2004; Ghu, 2017; Nam, 2015; Sadeghi & Dilmaghani, 2013; Yu, 2009) use vocabulary diversity (VocD) calculated by CLAN, a computer program, in order to measure lexical diversity.

The number of types indicates the total number of different words in a writing sample, and that of tokens means the total number of all used words. If TTR is close to 0, it means that the same words are repeated. In contrast, if TTR is close to 1, it means that there is little repetition among the written words. CLAN (MacWhinney, 2000) automatically calculates the VocD value, which is more reliable than TTR for measuring lexical diversity because VocD is less influenced by total words in a sample. According to Durán et al. (2004), a larger value of VocD stands for more diversity while a smaller value of VocD indicates less diversity. Adult authors' highest VocD value is 119.20, adult authors' mean VocD value is 90.59, and five year olds' average VocD value is 64.02.

Yu (2009) tested the effect of different writing topics on lexical diversity and found that there was a relationship between them even after controlling participants' proficiency levels. When the participants wrote about familiar topics, their lexical diversity increased. Their proficiency and their major were highly correlated to their lexical diversity. For example, the lexical diversity of participants majoring in nursing was higher than other major participants when they wrote about plastic surgery. Yu examined participants' written and spoken data and found that their lexical diversity was related with each other. Yu's participants did not choose a topic, but they wrote better when they were familiar with the topic.

The question arises of whether ESL/EFL writers show differences in terms of lexical diversity when every student has freedom to choose a writing topic and what the factors are that influence the differences. Mahmoudi and Mahmoudi (2017) investigated the effect of topic familiarity on reading comprehension. They divided participants into two groups, low and high, depending on their proficiency, and the results indicated that the high profi-

ciency group performed better than the low proficiency group regardless of topic familiarity. When Zamel (1983) investigated the writing process, she found that participants spent more time writing the first draft than writing other drafts and that participants' proficiency level influenced their final product. Mahmoudi and Mahmoudi's study examined the relationship between topic familiarity and language proficiency in reading comprehension, and Zamel's participants did not have freedom to choose the topic for their writing. Therefore, it is speculated that overall language proficiency might be the influencing factor that contributes to the differences in writing performance when language learners are given the same freedom for choosing their writing topic.

### 3. Writing Genres

Previous studies (Berman & Verhoeven, 2002; Sadeghi & Dilmaghani, 2013; Woerfel & Yilmaz, 2011) reported that lexical diversity differs depending on writing genres, and Martin (2009) and Watanabe (2016) have emphasized the importance of genre-based pedagogy in ESL/EFL writing instruction. Harman (2013), Bunch and Willett (2013) and de Oliveira and Lan (2014) reported the effectiveness of genre-based instruction in an ESL writing classroom, and Emilia and Hamied (2015) and Watanabe (2016) investigated the effectiveness of genre-based pedagogy in EFL classrooms.

Not every researcher has agreed with the categories of writing genres, but generally writing genres can be divided into four groups: narrative, argumentative, comparative, and expository. Watanabe (2016) claimed that L2 writing genre-based pedagogy played a significant role in improving learners' performance as well as their awareness of language use in different genres. Therefore, it was suggested that L2 writers need to be aware of specific needs of each genre and learn about them (Kang, 2005). McCabe and Bliss (2003) emphasized that each genre has specific characteristics and requires a different learning process. For example, language learners naturally learn features in narrative writing outside of the classroom, while discourse formats in expository writing need to be learned at school.

Sadeghi and Dilmaghani (2013) investigated the relationship between lexical diversity and writing genre among Iranian intermediate EFL learners. They asked 30 participants to write three essays of different genres: argumentative, narrative and comparative, so it is assumed that the participants did not choose their topic. The researchers measured lexical diversity in the essays using VocD

and found differences between narrative and comparative, and between narrative and argumentative genres. However, there was no statistically significant difference between comparative and argumentative. Also, overall writing quality was correlated with lexical diversity of comparative writing, while the relationship with writing quality was not statistically significant in narrative and argumentative essays when they were scored analytically. When the essays were scored holistically, overall writing quality was not correlated with lexical diversity in all three genres.

## III. Methods

### 1. Participants

A total of 72 female university freshmen in Seoul, South Korea, participated in this study. The university divided freshmen into three proficiency levels after the placement test at the time of the entrance. Forty-six out of 72 participants belonged to the low level, and 26 participants were from the intermediate level. The mean TOEIC score of the all participants was 533 (Min: 295, Max: 725, *SD*: 118.38); the average of the low level was 458.91 (Min: 295, Max: 580, *SD*: 75.74), and that of the intermediate level was 664.23 (Min: 600, Max: 725, *SD*: 38.23) (Table 1).

The TOEIC scores were used to estimate the participants' overall language proficiency when they were compared according to language proficiency and lexical diversity in writing.

### 2. Course Description

As a graduation requirement, students at the university were required to attend two mandatory English courses which focus on all four language skills: speaking, listening, reading, and writing. The participants attended two classes per week, and the maximum number of students per class was 25. The data were collected from two low-level classes and one intermediate class. Since the course was mandatory, the participants' majors and ages were varied, but the majority of them were freshmen. They submitted two essays during the semester, and each level had different writing topics.

The low-level participants wrote about an important

Table 1. TOEIC Scores (*N* = 72)

	<i>N</i>	Min.	Max.	<i>M</i>	<i>SD</i>
Low Level	46	295	580	458.91	75.74
Intermediate Level	26	600	725	664.23	38.23
All Participants	72	295	725	533	118.38

event in their lives and a cover letter for a resume. The first topic of the intermediate-level participants was the same as that for the low level. They wrote about a memorable event in their lives, but for the second essay they described six pictures of their choosing. The participants of both levels were allowed to choose any event for the first essay, but they had to comply with the basic guidelines for the assignment. All participants wrote about a memorable, interesting or unusual event in their lives after learning some useful expressions and completing speaking and written activities in the textbook. For the second topic of the intermediate level, participants chose a movie or a TV drama and chose an interesting scene. They captured six pictures from the scene and described them in the second essay. They also discussed their topics during the class and participated in brainstorming activities.

### 3. Data Collection and Analysis

This study compared essays from different genres, but the writing topics varied. One of the researchers taught all the participants, and the researcher explained the purpose of the study at the end of the semester after collecting all essays. The participation was all voluntary; no monetary award or course credit was provided. Students who did not want to participate in the study or did not submit one of the essays were excluded in the data analysis.

The CHILDES (Child Language Data Exchange System) Project provides tools in order to analyze transcription, and CLAN is one of the tools. CLAN is a free computer program that analyzes transcription (MacWhinney, 2000). It can be downloaded from the CHILDES Project website at <http://alpha.talkbank.org/clan/>, and the website also provides user manuals for the programs. Many studies (Ghu, 2017; MacWhinney, 2000; Nam, 2015; Yu, 2009) have used this program in order to analyze lexical diversity in English data, and the program is useful to analyze not only English but also other languages (Ghu, 2017; Nam, 2015).

In order to use CLAN, written data should be saved in CHAT format. The participants submitted two versions of essays: electronic version via email and a hard copy in face-to-face class. The electronic version was written using either the MS-Word program (.doc) or HWP program (.hwp), so the essays were converted into plain text (.txt) for the CLAN analysis. The plain text files were uploaded into the CLAN program and converted into CHAT format. The assignment guidelines asked the participants to write their name, department, student ID number, and the submission date in the header of the essays,

but the information was deleted in the text (.txt) format because it could cause an error in the CLAN analysis.

CLAN automatically produces an EXCEL spreadsheet (.xls) when 'eval' and 'kideval' commands are executed. Among the assorted information in the spreadsheet, the following information was selected in order to analyze the lexical diversity of the essays. Studies about lexical diversity have often used TTR (the type-token ratio), but that is usually influenced by the sample size, so this study analyzed VocD as well.

- Types: The number of different words used in the sample
- Tokens: The total number of all words in the sample
- TTR: The Type-Token ratio, a measure of lexical diversity. This number ranges from .001 to 1.0; a low value indicates a lot of repetition; a high value means each word in the sample was different.
- VocD: Vocabulary diversity. This is a twist on TTR, but supposedly more reliable when looking at files that vary substantially in length (Ratner & Brundage, 2018, pp. 26-27).

Linear regression analyses and independent samples t-tests were performed for the data analysis to analyze the relationship between proficiency level and lexical diversity. Participants wrote two essays of different genres, so dependent samples t-tests were performed in order to compare differences between different writing genres.

## IV. Results and Discussions

### 1. Factors Influencing VocD

When Lee and Suh (1998) surveyed 67 ESL learners in the States, the 11 topics that students were most interested in were health, business, computers, art, food, the United States, dreams, family, the human mind, language learning, and work/jobs. Wang (2010) also categorized essay topics into 13 areas: "family/friends, job, recreation/sports, health, education, finance, communication, society, frustrations, beauty, travel, science, and arts" (p. 173). The participants' favorite topics were education, frustrations, society and communication.

In the first essay, participants wrote about their most memorable moment in their lives, so they wrote about their happy, sad, mysterious, embarrassing or horrible experiences. The stories were usually about happy moments, ghosts they encountered, family illness, family matters, friendship, accidents, traveling, future plans, etc.

Because they were young, being freshmen at university, family illness or death was unexpected and shocking. Their language proficiency was not a factor in choosing a topic and did not correlate to total words in an essay. No specific topic was prevalent in a certain proficiency level. Linear regression analyses were performed between total words in an essay and language proficiency, but the results were not statistically significant.

In the second essay for the intermediate level, after participants chose a movie or a drama, they captured six scenes and explained each picture step by step. The stories were connected with each other because the six scenes were from the same movie or drama, but word repetition occurred less than in the first topic because there were six different pictures. The second essay topic for the low level was a cover letter for their resume, so they proclaimed their strengths and explained how qualified they were for the position for which they were applying. As they described several different areas such as educational background or previous job experiences, repetition occurred less than with the first topic as well.

The mean number of total words in the first essay was 204.49 ( $SD = 74.666$ ), and there were 108.58 types ( $SD = 27.686$ ) and 206.94 tokens ( $SD = 75.785$ ). The mean of VocD was 75.8765, which ranked slightly lower compared to average adult academic writings. According to Durán et al. (2004), the average VocD of native speaking adults' academic writings was 90.59, ranging from 69.74 to 119.20, and adult ESL learners' VocD was about 55, ranging from 40 to 70. Therefore, the current participants' VocD was below the native speakers' but above the average ESL learners' (Table 2).

As seen in Table 3, a total of 46 participants belonged to the low level and 26 participants were from the interme-

**Table 2.** Overview of the CLAN Results (First Essay,  $N = 72$ )

	Min.	Max.	<i>M</i>	<i>SD</i>
Total Words l	109	520	204.49	74.666
FREQ Types	74	213	108.58	27.686
FREQ Tokens	109	522	206.94	75.785
FREQ TTR	.41	.68	.5417	.06735
VocD	34.60	133.22	75.8765	19.45118

mediate level. The intermediate-level participants' total number of words in the first essay was higher than that of the low level, but their TTR and VocD values were lower than the low-level participants'. On the other hand, the total words, types, and tokens of the intermediate level were higher than those of the low level, which indicated that intermediate-level participants repeated same words more frequently than low-level participants did. In order to compare the group differences, independent t-tests were administered.

Table 3 illustrates that the differences in TOEIC scores between intermediate and low levels were statistically significant at .01, but the group differences in total words, types, tokens, TTR, and VocD were not statistically significant. Yu (2009) reported that higher-level writers' TTR and VocD were higher than lower-level writers', but the lower-level participants' lexical diversity was higher than that of the higher-level participants in this study. Yu did not specify what topics participants wrote about, and she only mentioned that the essay topics were three impersonal and two personal topics. Therefore, the topic variable may cause the results to differ, and writing genres could influence the differences as well.

The t-test results were not anticipated, so linear regression analyses were performed in order to find influencing factors on lexical diversity. They were not grouped by two

**Table 3.** Results of t-test between Low- and Intermediate-Levels

Factor	Group	<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>
Total Words l	Low	46	196.91	58.437	1.005	.322
	Intermediate	26	217.88	96.922		
FREQ Types	Low	46	106.50	21.864	.744	.462
	Intermediate	26	112.27	35.977		
FREQ Tokens	Low	46	199.04	59.328	1.034	.308
	Intermediate	26	220.92	98.271		
FREQ TTR	Low	46	.5480	.06563	-1.060	.293
	Intermediate	26	.5305	.07018		
VocD	Low	46	76.6941	20.79565	-4.72	.639
	Intermediate	26	74.4300	17.10780		
TOEIC	Low	46	458.91	75.740	15.265	.000
	Intermediate	26	664.23	38.228		

**Table 4.** Regression Analysis Results

Factor	<i>B</i>	Beta	<i>t</i>	<i>p</i>
Total Words 1	.044	.169	1.437	.155
FREQ Types	.289	.412	3.780	.000
FREQ Tokens	.043	.166	1.412	.162
FREQ TTR	129.829	.450	4.211	.000
TOEIC	-.020	-.121	-1.018	.312

levels, intermediate and low, but their TOEIC scores were used in order to represent their language proficiency. Five different linear regression analyses were administered. The value of Durbin-Watson was not close to either 0 or 4, and normality and homoskedasticity of lexical diversity were assumed. *Table 4* presents the regression analysis results between all factors and VocD.

Both VocD and TTR can measure the level of lexical diversity (Ratner & Brundage, 2018), so the close relationship between TTR and VocD was expected. Type stands for different words in an essay, while token indicates total words, so the close relationship between type and VocD could be expected as well. However, token and total words were not regressed with the lexical diversity indicator, VocD, which indicates that a longer essay does not necessarily signify greater lexical diversity. The TOEIC scores, overall language proficiency, and VocD were not regressed either, which meant that the writings of the more proficient learners were not automatically diverse in lexicality.

The final drafts of the writing assignments were not written in class. Before submitting the final version, during classes, the participants had some brainstorming activities as a group, wrote a first draft, went through a peer feedback process, and had a writing conference with the instructor. Because they wrote the final version of the essays at home or outside of the classroom, they could

freely use a dictionary or other supplementary writing tools including a translator between Korean and English regardless of their proficiency. Therefore, the current results could be different from those of the onsite language composition examination in Lee and Anderson (2007) and Yu (2009).

## 2. VocD in Writings of Different Genres

Previous studies (Berman & Verhoeven, 2002; Sadeghi & Dilmaghani, 2013; Woerfel & Yılmaz, 2011) have reported that writing genres influence ESL/EFL writers' performance. In order to examine the influence of writing genres, paired samples *t*-tests were performed. All participants submitted two essays of two different writing genres, so the participants' first and second essays were compared. The intermediate-level participants' narrative and descriptive writings were compared, and the low-level participants' narrative and expository writings were compared (*Table 5*).

The first topic of the intermediate level was the most memorable event in their lives, which was narrative, and the second topic was picture description, which was descriptive. Similar to the previous studies, differences of VocD between the two genres were statistically significant at .01 level even though differences in total words, types, tokens, and TTR were nonsignificant at .05 level. Intermediate-level participants wrote more diverse words in descriptive writing than in narrative writing, and length of writing was not a predictor of lexical diversity. *Table 6* illustrates the results of paired samples *t*-tests of the low level. A total of 46 participants wrote the first essay, but 44 participants' first and second essays were compared because two of the participants did not submit the second essay.

The first topic of the low level was the same as the intermediate level, the most memorable event in their lives,

**Table 5.** Paired-samples *t*-tests (Intermediate Level, *N* = 26)

Factor	Essay	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>
Total Words	1 <sup>st</sup>	217.88	96.922	11.46	70.35	.831	.414
	2 <sup>nd</sup>	206.42	72.641				
FREQ Types	1 <sup>st</sup>	112.27	35.977	-1.73	30.34	-.291	.774
	2 <sup>nd</sup>	114.00	28.814				
FREQ Tokens	1 <sup>st</sup>	220.92	98.271	10.42	70.95	.749	.461
	2 <sup>nd</sup>	210.50	72.799				
FREQ TTR	1 <sup>st</sup>	.5305	.07018	-.02527	.07007	-1.839	.078
	2 <sup>nd</sup>	.5558	.06281				
VocD	1 <sup>st</sup>	74.43	17.11	-17.81	23.76	-3.824	.001
	2 <sup>nd</sup>	92.24	24.03				

**Table 6.** Paired-samples *t*-tests (Low Level, *N* = 44)

Factor	Essay	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>
Total Words	1 <sup>st</sup>	195.14	59.152	-13.93	52.01	-1.777	.083
	2 <sup>nd</sup>	209.07	31.366				
FREQ Types	1 <sup>st</sup>	106.07	22.268	-8.795	21.26	-2.745	.009
	2 <sup>nd</sup>	114.86	15.081				
FREQ Tokens	1 <sup>st</sup>	197.20	60.025	-3.68	52.73	-4.463	.646
	2 <sup>nd</sup>	200.89	31.074				
FREQ TTR	1 <sup>st</sup>	.5509	.06563	-.0242	.0659	-2.439	.019
	2 <sup>nd</sup>	.5752	.03827				
VocD	1 <sup>st</sup>	76.93	21.131	-14.11	24.44	-3.829	.000
	2 <sup>nd</sup>	91.04	17.699				

and the second topic was a cover letter for their resume. The first was categorized as narrative, and the second was expository. The differences in types, TTR, and VocD between the first and second essays were statistically significant at .05 level, but those of total words and tokens were statistically nonsignificant.

The TTR of the narrative was .5509, and VocD was 76.93. Intermediate-level participants' TTR was .5305, and VocD was 74.43. The TTR of the expository writing was .5752, and VocD was 91.04. Intermediate-level participants' TTR of the descriptive writing was .5558, and VocD was 92.24. Even though it is hard to compare expository and descriptive writing because the proficiency levels of the expository and descriptive writings were different, different writing genres may influence the differences in lexical diversity, and their overall proficiency did not predict the level of lexical diversity.

In the narrative essays of both levels, participants often used the first-person pronoun *I* as a subject and *be* verb as a verb. For example, a participant wrote about her experience training her pet. She wrote a total of 17 sentences, and the first-person pronoun *I* was used 10 times, and the possessive pronoun *my* was used seven times. The past tense of the *be* verb *was* was used eight times. Even when participants did not use pronouns, they often repeated the same subjects consecutively in order to tell their stories. For example, a participant wrote a story about seeing a pigeon in a park. She wrote about when she saw the pigeon and what the pigeon did. Naturally, she repeated the word "pigeon" several times as a subject or an object. The repetition of pronouns or nouns may make the essays more monotonous, and the values of type, TTR, and VocD decreased. On the other hand, in the descriptive writing, writers' word usage was more varied. They explained six different pictures, so the characteristics of the assignment compelled participants to use more diverse words, sub-

jects and verbs.

In the expository writing, participants needed to explain their strengths and experiences in the cover letter for their resume, so participants often used the first-person pronoun *I* as well. In the cover letter, their address, names, dates, and final greeting were deleted because the information was not necessary and could bias the value of VocD. Even though participants often used the first-person pronoun as a subject, their stories, educational qualifications and job experiences were not connected with each other like the stories in the narrative writing. In the narrative, they wrote about one story that was the most memorable to them, so its lexical diversity was lower than that in the expository or descriptive writing.

## V. Conclusions

Freedom to choose a writing topic can increase writers' motivation, enhance overall writing quality, and increase lexical diversity (Gradwohl & Schumacher, 1989; Wang, 2010; Yu, 2009), and writing genres can influence lexical diversity as well (Berman & Verhoeven, 2002; Sadeghi & Dilmaghani, 2013; Woerfel & Yilmaz, 2011). Because few empirical studies have been conducted which integrated these factors, writing topic choice, lexical diversity, and writing genre, this study investigated the influence of writing topics, writing genres, and writers' overall proficiency on lexical diversity.

The results indicated that providing topic choice did not automatically promote lexical diversity in writing even though the freedom could enhance writers' motivation and enthusiasm to write further and more. Lee and Anderson (2007) emphasized the influence of language proficiency on lexical diversity, and the length of essays was assumed to increase the level of lexical diversity. More proficient writers were expected to write more varied

words, but the present study revealed that longer essays and higher language proficiency were not necessarily correlated with the level of lexical diversity.

Prior knowledge (Barry, Nielsen, Glasnapp, Poggio, & Sundbye, 1997; Lee & Anderson, 2007) and writing genre (Sadeghi & Dilmaghani, 2013) have been shown to influence lexical diversity. When Sadeghi and Dilmaghani measured lexical diversity between narrative and comparative, and between narrative and argumentative genres, differences were found. However, there was no statistically significant difference between comparative and argumentative genres. Similar to Sadeghi and Dilmaghani's study, participants' levels of lexical diversity in narrative writing were lower than that in descriptive and expository writing, and differences between narrative and descriptive, and between narrative and expository writing were statistically significant. Even though the difference between expository and descriptive writing could not be compared, their values of VocD indicated that the differences between them were not as significant as those between narrative and the other two genres.

Wang (2010) claimed that her participants showed differences between self-selected topics and large-scale test prompts because of their lack of practice with adopting more effective writing strategies. She concluded that language learners need to learn how to choose a topic for their writing in strategy training. The importance of freedom to choose a topic was well noted, but the results of the current study indicate that writing teachers need to develop diverse activities to meet language learners' needs and to be careful to provide appropriate instruction for each writing genre. Language learners need to be prepared to write all different types of writing genres for their jobs or their majors, and need to learn genre specific characteristics, vocabulary, and writing structures.

Unlike the results of Lee and Anderson (2007) and Yu (2009), lexical diversity was not regressed with proficiency. The different writing environments, regardless of whether language learners could use a dictionary or not when writing, might be the reason for the different results. The analyzed writing assignments were written outside of the classroom, but other studies analyzed essays written for a test. For further study, a comparison of lexical diversity between essays from a test and from an assignment could provide more insight to compare the differences in research results. Also, this study did not compare writings of individuals across four different genres, so having learners write essays of different genres could provide a bigger picture about the influence of writing genres on

lexical diversity.

## References

- Barry, A., Nielsen, D., Glasnapp, D., Poggio, J., & Sundbye, N. (1997). Large scale performance assessment in writing: Effects of student and teacher choice variables. *Contemporary Education*, 69(1), 20-26.
- Berman, R. A., & Verhoeven, L. (2002). Cross-linguistic perspectives on the development of text-production abilities: Speech and writing. *Written Language and Literacy*, 5(1), 1-43.
- Bhatia, V. K. (2002). Applied genre analysis: A multi-perspective model. *IBÉRICA*, 4, 3-19.
- Bunch, G. C., & Willett, K. (2013). Writing to mean in middle school: Understanding how second language writers negotiate textually-rich content-area instruction. *Journal of Second Language Writing*, 22(2), 141-160.
- Cumming, A. (1989). Writing expertise and second-language proficiency. *Language Learning*, 39(1), 81-135.
- de Oliveira, L. C., & Lan, S. (2014). Writing science in an upper elementary classroom: A genre-based approach to teaching English language learners. *Journal of Second Language Writing*, 25, 23-39.
- Durán, P., Malvern, D., Richards, B., & Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics*, 25(2), 220-42.
- Emilia, E., & Hamied, F. A. (2015). Systemic functional linguistic genre pedagogy (SFL GP) in a tertiary EFL writing context in Indonesia. *TEFLIN Journal*, 26(2), 155-182.
- Gradwohl, J., & Schumacher, G. (1989). The relationship between content knowledge and topic choice in writing. *Written Communication*, 6(2), 181-195.
- Ghu, D. Y. (2017). *Lexical diversity in the writings of intermediate level multi-cultural learners and foreign learners*. Unpublished master's thesis, University, Seoul.
- Harman, R. (2013). Literary intertextuality in genre-based pedagogies: Building lexical cohesion in fifth-grade L2 writing. *Journal of Second Language Writing*, 22(2), 125-140.
- Kang, J. Y. (2005). Written narratives as an index of L2 competence in Korean EFL learners. *Journal of Second Language Writing*, 14(4), 259-279.
- Lee, H. K. (2004). Impacts of topics and writers' department background on writing performance. *English Teaching*, 59(3), 271-292.
- Lee, H. K., & Anderson, C. (2007). Validity and topic generality of a writing performance test. *Language Testing*, 24(3), 307-330.
- Lee, D., & Suh, C. S. (1998). What do ESL students want to read and write about? Comparing students' interests with topics found in theme-based ESL reading/writing textbooks. *The Journal of English Language Teaching*, 8, 47-60.

- Lo, J., & Hyland, F. (2007). Enhancing students' engagement and motivation in writing: The case of primary students in Hong Kong. *Journal of Second Language Writing*, 16(4), 219-237.
- MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. 3<sup>rd</sup> Edition. Mahwah, NJ: Lawrence Erlbaum Associates.
- Mahmoudi, S., & Mahmoudi, A. (2017). Can topic familiarity override language proficiency in reading comprehension? *Journal of Language Teaching and Research*, 8(3), 496-506.
- Martin, J. R. (2009). Genre and language learning: A social semiotic perspective. *Linguistics and Education*, 20(1), 10-21.
- Nam, J. Y. (2015). *A study of oral complexity of Korean learner: Syntactic complexity and lexical complexity*. Unpublished doctoral dissertation, University, Seoul.
- Ratner, N. B., & Brundage, S. B. (2018). *A clinician's complete guide to CLAN and PRAAT*. Last update: February 2018. Retrieved from <https://talkbank.org/manuals/Clin-CLAN.pdf>
- Sadeghi, K., & Dilmaghani, S. K. (2013). The relationship between lexical diversity and genre in Iranian EFL learners' writings. *Journal of Language Teaching and Research*, 4(2), 328-334.
- Simmons, J. (2007). Dealing with troubled writers: A literacy teacher's dilemma. *Journal of Adolescent & Adult Literacy*, 51(1), 4-8.
- Stapa, M. (2001). Assessing ESL writing performance: The influence of background knowledge on writing performance. Paper presented in annual meeting of Midwest Association of Language Testers (MwALT), Ann Arbor, MI.
- Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing*, 32, 17-35.
- Tedick, D. J. (1990). ESL writing assessment: Subject-matter knowledge and its impact on performance. *English for Specific Purposes*, 9(2), 123-143.
- Tedick, D. J. (1988). *The effects of topic familiarity on the writing performance of non-native writers of English at the graduate level*. Unpublished doctoral dissertation, The Ohio State University.
- Wang, D. (2010). Chinese students' choice of writing topics: A comparison between their self-selected topics and writing prompts in large-scale tests. *The Journal of Asia TEFL*, 7(3), 165-187.
- Watanabe, H. (2016). Genre analysis of writing tasks in Japanese university entrance examinations. *Language Testing in Asia*, 6(4). Retrieved from <https://doi.org/10.1186/s40468-016-0026-8>
- Woerfel, T., & Yilmaz, S. (2011). Lexical development of German-Turkish bilinguals: A comparative study in written discourse. In Chris Cummins et al. (eds.), *Proceedings of the 6<sup>th</sup> Cambridge Postgraduate Conference in Language Research* (pp. 240-251). Cambridge: Cambridge Institute of Language Research.
- Yu, G. (2009). Lexical diversity in writing and speaking task performances. *Applied Linguistics*, 31(2), 236-259.
- Zamel, V. (1983). The composing processes of advanced ESL students: Six case studies. *TESOL Quarterly*, 17(2), 165-187.