

# Korean EFL Learners' Use of *I* in English Argumentative Writing: Focusing on Genre-specific and Proficiency-specific Characteristics

Eun Kyoung Lee & Sun-Young Oh<sup>†</sup>

Seoul National University

---

## ABSTRACT

The present study explored Korean EFL learners' use of the first person singular pronoun, *I*, in argumentative essays. To examine whether learners' *I* differs according to genre and proficiency, the use of *I* in argumentative writing was compared with personal writing and between learners of low and high proficiency. The frequency, variability, and collocates analyses of *I* showed that learners used less *Is* with little variability in argumentative essays, mostly using it to deliver opinions in formulaic expressions or to strengthen arguments with boosters. In terms of proficiency, high-level learners in general used *I* less often and with more hedges compared to low-level learners, more closely resembling native speakers' use of *I*. The study suggests that learners depend on *I* to express their opinions and achieve the goal of the argumentative genre and that the knowledge of using *I* effectively in writing is gradually acquired as learners' proficiency increases.

**Keywords:** first person pronoun, second language writing, genre, proficiency, corpus linguistics

---

## 1. Introduction

First person pronouns serve multiple functions in academic prose including the argumentative genre. One major effect of using first person pronouns in argumentative essays is that they can represent writers' "identities that are created within the genre" (Tang & John 1999: S25). That is, writers use the pronouns in order to present themselves in their essays and achieve genre-specific goals of argumentative writing such as to deliver the writer's position or to guide readers through the essay (Tang & John 1999; Hyland 2002). Yet, it has also been suggested that the use of first person pronouns, especially *I*, is considered as undesirable features in academic prose. This is because it may hinder the formal register and objective tone typically

---

<sup>†</sup> Corresponding author: sunoh@snu.ac.kr



associated with delivering opinions and persuading readers (Hinkel 1999; Kuo 1999). In this respect, it is generally recommended that personal pronouns should be used with planned intentions so that they can more effectively serve its functions while avoiding the risk of harming the message being delivered.

While native speakers who are aware of such conventions use language and register according to genre, learners often experience difficulty in doing so (Deng et al. 2014). As a result, one of the major characteristics observed in the learners' argumentative writing is the excessive use of first personal pronouns (McCrostie 2008; Petch-Tyson 1998; Y Choung & S-Y Oh 2017). When compared against native speakers' productions, learners often used personal pronouns with high frequencies and in contexts where the use of the pronouns increased the subjectivity of learners' arguments (Akahori 2007). In particular, previous literature reported overuses of the first person singular pronoun, *I* (Recski 2004; Paquot et al. 2013; Natsukari 2012). Learners used *I* to deliver personal opinions and often used it in formulaic phrases such as *I think* (Aijmer 2001, 2002; Gilquin & Paquot 2008; Granger 1998; Hasselgård 2009; S-Y Oh 2007; Paquot et al. 2013).

While abundant research has been conducted on learners' use of first person pronouns in argumentative essays as well as published research articles individually (Martinez 2005; Henderson & Barr 2010), not much attention has been given to genre-specific characteristics of learners' productions of *I* through comparisons between genres (Paquot et al. 2013). Also, while some studies observed *I* according to learners' grade levels in university (McCrostie 2008; Neff et al. 2004), only limited research has been conducted on the difference according to learners' proficiency (Y Choung & S-Y Oh 2017). Considering the view that appropriate use of personal pronouns plays genre-specific roles in argumentative essays (Kuo 1999; Hinkel 1999) and the possibility of acquisition of such knowledge as proficiency increases (Y Choung & S-Y Oh 2017), the present study aims to conduct a detailed examination of learners' *I* in the argumentative writing, focusing on its genre- and proficiency-specific characteristics.

In terms of analysis, it is important to note that most studies examined the use of *I* based on frequency analyses (e.g. Neff et al. 2004) or by inspecting a limited pool of common contexts in which *I* appears (e.g. Natsukari 2012). So far, not many have elaborately analyzed the collocates associated with *I* as well as the variability of *I* across learners to see whether the high frequency of *I* was a general or individualistic pattern of learners (Recski 2004; Lee & Deakin 2016). By conducting a more systematic analysis of how learners use *I* in terms of its frequency, variability,

and collocates, the current study aims to carry out a more well-rounded investigation of learners' *I* in argumentative writing.

## 2. Literature Review

### 2.1. First person pronouns in learners' argumentative writing

Learners' excessive use of personal pronouns was found in the argumentative genre in numerous studies, although the use of *I* was often not observed separately. Petch-Tyson (1998) reported that non-native speakers showed greater writer/reader visibility than native speakers in argumentative essays. Learners used first person pronouns (*I, me, my, mine, we, us, our, ours*) in their writing two to four times the rate in native speakers' writing. This study was picked up by McCrostie (2008), in which Japanese EFL learners overused the personal pronouns that were observed in Petch-Tyson (1998). Both studies showed that learners tended to use the pronouns to deliver personal thoughts in a more direct manner. This differed from observations of native speakers' productions found in Tang and John (1999), where native speakers most often used the pronouns to guide readers through their essays rather than to explicitly state arguments. Such frequent use of first person pronouns (*I, we*) was found in Korean EFL learners' argumentative writing as well in Y Choung and S-Y Oh's (2017) study.

Meanwhile, a contrasting result was discovered in a study conducted by Lee and Deakin (2016). The researchers observed the frequency of self-mentions in Chinese EFL learners' and native speakers' argumentative writing by looking at first person pronouns (*I, me, my, mine, our, us, and we*) as well as other self-mention words like *the author, the researcher, and the writer*. The results showed that learners refrained from presenting themselves in their writing and used significantly less words that increases writer visibility compared to native speakers. However, in this study, words other than first person pronouns were also included in analysis and the findings were discussed with reference to previous studies that examined journal articles instead of the same genre, argumentative writing. Therefore, it is unclear whether the results of this study indicates learners' underuse of personal pronouns.

The current study suggests that the mixed results above can be explained by the fact that most studies examined personal pronouns altogether without separating the use of the first person singular pronoun, *I*, from other pronouns. Personal pronouns

were often all grouped together in analysis, being categorized into an overarching term, such as “self-mention” words or words that show writer visibility, with other personal pronouns (e.g. Hyland 2002; Lee & Deakin 2016). However, there are several reasons why there is a need to observe the use of *I* separately from other first person pronouns. First, some studies reveal numerical differences between the frequency of *I* and other personal pronouns, showing that compared to other first person pronouns, *I* was used significantly more often by learners (Recski 2004; Paquot et al. 2013). This indicates that a closer inspection of *I* could provide better explanations as to why learners use the pronoun so frequently. Moreover, when compared with the first person plural *we*, the singular *I* is considered as a representative pronoun of the subjective voice in writing. It is often suggested that an excessive use of the first person singular pronoun could damage the objectivity or power of arguments that is required especially in the argumentative genre (Akahori 2007; Natsukari 2012). Therefore, in order to examine whether learners are sensitive to genre conventions and use *I* in the most effective ways to deliver their arguments, it is critical to observe learners’ use of *I* in detail, apart from other personal pronouns.

## 2.2. The first person singular pronoun *I* in learners’ argumentative writing

There are only a few studies that paid specific attention to learners’ and native speakers’ use of *I* among other personal pronouns. These studies overall disclosed a particular overuse of *I* in learners’ argumentative writing. Neff et al.’s (2004) study examined argumentative essays written by Spanish learners of English and native speakers and found that the learners used *I* more often than native speakers. Natsukari (2012) also revealed an overuse of *I* in advanced-level Japanese EFL learners’ argumentative essays. It was noticed that learners used *I* especially to “write about personal matters and to express their opinions” by excessively using the phrase, *I think* (Natsukari 2012: 74). Gilquin and Paquot (2008) reported a similar result, pointing out a greater frequency of the phrase, *I think* in argumentative and literary essays of learners with various L1s. Yet, all three studies did not verify whether the frequency difference between native and non-native groups was statistically significant and also did not inspect the actual uses of *I* in specific writing contexts, such as collocates associated with *I*, in detail.

In this regard, the study that most closely examined learners’ collocates of *I* was Recski’s (2004) research. The study demonstrated that learners of various L1s used

*I* frequently in argumentative essays in order to express personal opinions as shown in Petch-Tyson's (1998) and McCrostie's (2008) works. In contrast, native speakers instead often used *I* to recount past experiences as shown in Tang and John (1999). Yet, the researcher made an additional remark about the distribution of *I* across learners' essays, pointing out that "there are a large number of writers who are prone to overusing *I* instead of all the users employing the pronoun too extensively" (Recski 2004: 12). Such finding regarding the uneven distribution of *I* across learners was also discovered in Lee and Deakin (2016). Although this presents an interesting point about the variability across individuals in learners' use of *I*, it is somewhat limited in that the distribution was not compared with native speakers. Thus, a comparison of the variability in learners' and native speakers' use of *I* could lead to a better understanding of learners' overall tendency in using the first person pronoun.

### 2.3. Genre effects in learners' use of *I*

While abundant research has been conducted on learners' use of personal pronouns, including *I*, in the published journal articles and argumentative prose, there has not yet been much consideration about the use of the pronoun across different genres. There have been studies that examined non-native speakers' use of first person pronouns in published research papers. Hyland (2002) found an underuse of personal pronouns (*I, me, my, we, us, our*) in Hong Kong students' academic reports compared to native speakers' published research papers in various disciplines. Martinez (2005) also disclosed learners' less frequent use of first person pronouns (*we, our, us*) in published papers in the biology discipline. While such previous literature that investigated the use of personal pronouns in research articles showed that there was a tendency of non-native speakers to avoid the use of first person pronouns (but also see Henderson & Barr 2010), those that observed the use in argumentative writing mostly revealed opposite results (e.g. Natsukari 2012) (but also see Lee & Deakin 2016). Such different tendencies of learners' overuse or underuse of personal pronouns according to distinct genres suggest that learners may be altering their use of the pronouns according to genre. A direct comparison between genres could reveal such genre effect in learners' use of personal pronouns.

In this regard, Paquot et al.'s (2013) study observed writer/reader visibility in learners' and native speakers' argumentative and research writing. The researchers discovered that both learners and native speakers used *I* more often in argumentative writing compared to research articles, although the difference was not statistically

significant. Despite the fact that further contextual analyses of *I* in the two genres were not conducted, the results of Paquot et al. (2013) present a possibility that learners can differentiate between genres and use *I* in different ways. Further research on the genre differences in learners' use of personal pronouns is required to justify this claim. In this respect, the present study aims to examine the genre-specific use of *I* in learners' argumentative writing by comparing it with the use in personal writing, a genre that significantly differs from argumentative prose in terms of writing goal and register.

#### 2.4. Proficiency effects in learners' use of *I*

In terms of a proficiency effect on learners' use of *I*, there exist only a handful of studies that observed differences in the use of first person pronouns within the learner group, according to years of L2 writing experience. In McCrostie (2008), 1<sup>st</sup> year Japanese university students used first person pronouns more frequently than 2<sup>nd</sup> year students and in Neff et al. (2004), 1<sup>st</sup> year Spanish college students used *I* more often than 4<sup>th</sup> year students. Thus, both research showed that learners in higher grade levels, with potentially more experience in English writing used less *Is* in their written productions compared to those in lower grade levels. Y Choung and S-Y Oh's (2017) study added on to this observation with the finding that Korean EFL learners' proficiency, which was measured by the ratings of texts based on the TOEFL iBT writing section scoring rubric, showed a negative correlation with their use of *I* in argumentative writing: as proficiency increased, the frequency of *I* decreased. Meanwhile, Storch (2009) conducted a learning study with non-native speakers who received one semester of instruction on English writing conventions in an English-speaking setting. The study revealed that learners significantly reduced their use of first person pronouns in academic writing after the instruction, showing that with exposure to English writing conventions, learners' use of first person pronouns became more similar to that of native speakers. While previous works did not separate learner groups based on a specific criteria, such as a standardized proficiency test, they presented an overall trend in learners' use of personal pronouns becoming more native-like as they gain more experience in English writing.

Considering the fact that the overly frequent use of *I* can indicate an inappropriate register in written discourse, as it more closely resembles patterns in spoken language (Akahori 2007; Thompson 2013; Y Choung & S-Y Oh 2017), it is necessary to observe whether such ability to attend to the conventions of argumentative writing

improves as learners' proficiency increases. Therefore, in order to more effectively observe the direct relationship between this ability and learners' proficiency, the current study compares native speakers' and learners' use of *I* and also examines differences between low- and high-level learners by dividing the learner group based on the widely-used Common European Framework of Reference for Languages (CEFR).

Altogether, the present study seeks to investigate learners' and native speakers' use of *I* in argumentative writing by focusing on its genre-specific and proficiency-specific characteristics. While most corpus studies observed the frequency of *I* in learners' and native speakers' writing, only a few have made limited attempts to analyze the variability across individuals and to look into specific collocates associated with *I* (e.g. Recski 2004). Therefore, this study contributes to the research on learners' use of *I* by employing a more systematic approach of analysis and examining the frequency, variability, and collocates of *I* in argumentative writing.

The current study aims to answer the following research questions:

- 1) Does Korean EFL learners' use of *I* in argumentative writing differ from that of personal writing?
- 2) Does Korean EFL learners' use of *I* in argumentative writing differ according to proficiency level?

### 3. Method

#### 3.1. Data

In order to examine Korean EFL learners' use of *I* in English argumentative essays, the current study used the Yonsei English Learner Corpus (YELC) as the main corpus of analysis and the Louvain Corpus of Native English Essays (LOCNESS) as a reference corpus to which learners' texts were compared (Table 1). YELC is a learner corpus that consists of two types of writing collected from the same group of learners, who produced the texts as part of an English exam at Yonsei University in Korea (S-C Rhee & C K Jung 2014). One portion of YELC contains learners' responses to essay prompts to which learners wrote a complete, argumentative essay expressing their position on the given topic. An example essay prompt is: "Why should people receive a college education? State your opinion." (S-C Rhee & C K Jung 2014). The other section of YELC is a collection of more personal pieces of

writing by the same learners, who wrote responses to essay questions that asked for descriptions of their personal experiences such as “What was your favorite extracurricular activity in high school? What made you join the activity?” (S-C Rhee & C K Jung 2014). The fact that the argumentative and personal essays were collected from the same learners makes YELC a suitable corpus to compare learners’ productions according to genre. The texts in YELC are divided into nine proficiency levels based on the Common European Framework of Reference for Languages (CEFR). For the purpose of comparing low- and high-level learners’ use of *I*, the present study divided the nine levels into two proficiency groups, low and high, so that the number of texts in each group was comparable: the low proficiency group (A1, A1+, A2) and high proficiency group (B1+, B2, B2+, C1, C2). The B1 level was excluded to make a clearer distinction between the two proficiency groups.

In addition to learners’ data, native speakers’ argumentative essays from LOCNESS were analyzed in order to explain potential differences found in learners’ *I* according to proficiency.<sup>1)</sup> LOCNESS is a corpus that consists of American and British college students’ argumentative and literary essays, and it is widely used in learner corpus studies as a reference corpus (Granger 2015; Altenberg & Granger 2001). In order to control the type of writing to the argumentative genre, the current study selected the subcorpus of argumentative essays written by American students, which was most comparable to the argumentative writing subcorpus of YELC; both subcorpora consisted of college students’ answers to prompts that asked them to write about their opinions on similar topics including controversial social issues (Granger 2015). The detailed descriptions of all the texts analyzed in the study are shown in Table 1.

**Table 1.** Examined subcorpora in YELC and LOCNESS

	Texts	Tokens	Types	STTR
Argumentative writing				
Low-level learners (YELC)	910	157,314	7,391	33.35
High-level learners (YELC)	1,203	336,767	10,328	34.95
Native speakers (LOCNESS)	175	149,559	10,379	40.02
Personal writing				
Low-level learners (YELC)	910	77,207	5,328	36.11
High-level learners (YELC)	1,203	123,036	6,597	37.59

1) A native corpus of personal essays was not analyzed since it would go beyond the scope of this paper, which focuses on learners’ use of *I* in the argumentative genre; also, a comparable reference corpus of personal essays was unavailable.



### 3.2. Procedure

The present study investigates the use of *I* in Korean EFL learners' and native speakers' argumentative writing in three different aspects: frequency, variability, and collocates. In order to determine the genre-specific and proficiency-specific characteristics of learners' use of *I* in argumentative writing, for each analysis, three types of comparisons were conducted: the use of *I* in argumentative writing was compared against personal writing, and also compared between groups in two different ways: learners versus native speakers and low-level learners versus high-level learners.

First, the frequency of *I* in learners' and native speakers' both argumentative and personal essays was calculated. A comparison between learners' frequencies of *I* in the two genres was conducted. Subsequent analyses compared learners' and natives' frequencies as well as low- and high-level learners' frequencies of *I* in argumentative essays.

Second, the variability of individuals' frequency of *I* in learners' argumentative and personal essays was examined by taking individual frequencies of *I* and calculating figures required to plot a boxplot. Then, the boxplots were used for analyzing learner variability. The same process was applied to native speakers' argumentative essays. The variability in learners' argumentative writing was compared against personal writing and then further analyzed by comparing between groups (learners versus natives, low-level versus high-level learners).

Finally, a closer examination of the actual usages of *I* was carried out by manually inspecting patterns of learners' and native speakers' use of *I*. Collocates of *I* in learners' argumentative and personal writing were analyzed by looking at the top 10 most frequent words that immediately preceded and followed *I* in both genres. For the proficiency comparison, the top 25 most frequent words that preceded and followed *I* in both learners' and native speakers' argumentative writing were observed. Further discussions of the results were made by referring to the actual contexts in which the collocates were used and analyzing sample concordances of the specific pairs of collocates and *I* observed in the top 25 collocates list.

The tools employed for analyses included WordSmith Tools 7.0 (Scott 2017), RStats Chi-square Calculator (Daniel & Kostic 2017), and RStudio (RStudio Team 2015). WordSmith was used for frequency calculations and pattern inspections of the selected corpora; the built-in "Concord" and "Wordlist" functions were applied. Statistical analyses and graphical representations of data were conducted using the RStats Chi-square Calculator and RStudio.

4. Results and Discussion

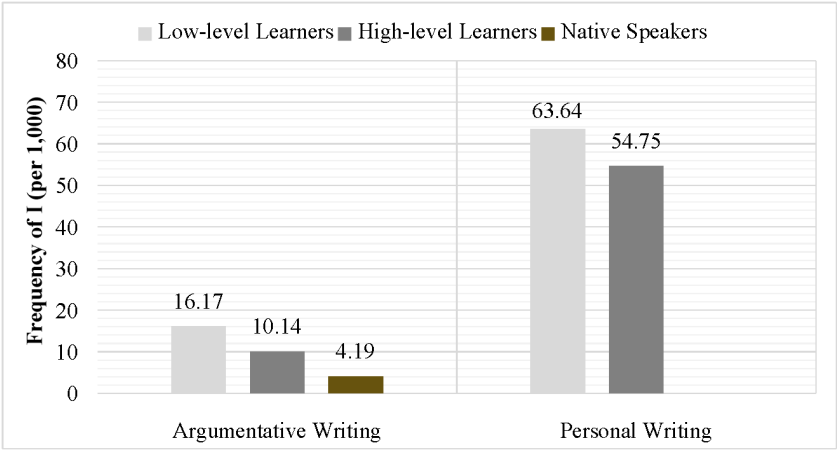
This section is organized into two parts, divided according to the type of analysis conducted: frequency and variability of *I* (4.1.) and collocates of *I* (4.2.). For each part, the results are presented in two subsections according to the research questions of the present study: genre-specific characteristics of *I* (4.1.1., 4.2.1.) and proficiency-related characteristics of *I* (4.2.1., 4.2.2).

4.1. Frequency and variability of *I*

The overall frequency results are presented in Table 2 and Figure 1, and the variability results are shown in Table 3 and Figure 2.

**Table 2.** Learners’ and native speakers’ frequency of *I* according to genre and proficiency

	Argumentative writing		Personal writing	
	Raw	Per 1,000	Raw	Per 1,000
Low-level learners	2,552	16.17	4,905	63.64
High-level learners	3,406	10.14	6,715	54.75
Native speakers	627	4.19	N/A	N/A

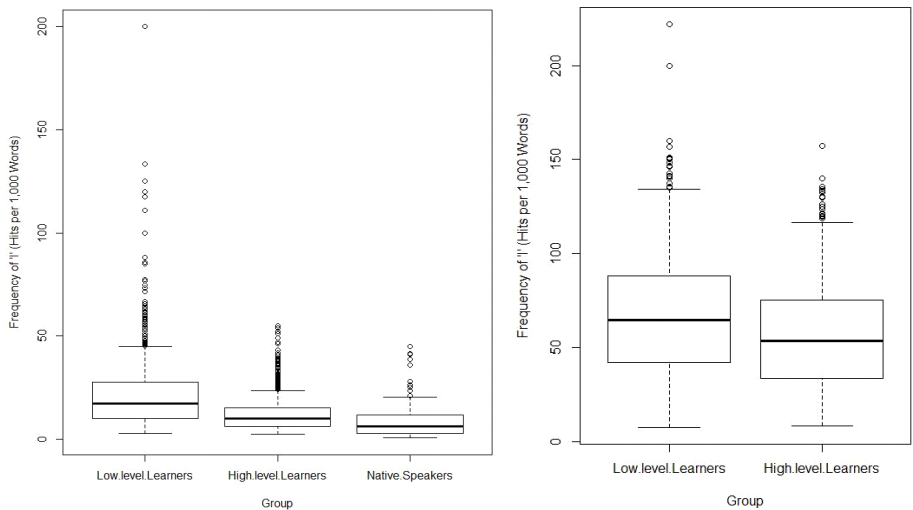


**Figure 1.** Learners’ and native speakers’ frequency of *I* according to genre and proficiency.

The frequency values of *I* for each group are recorded in Table 2 and plotted in Figure 1. The results showed that there are significant differences in the frequency of *I* among groups for both argumentative and personal writing. More detailed analyses are covered in the subsections below.

**Table 3.** Descriptive statistics on learners’ and native speakers’ frequency of *I* (per 1,000) according to genre and proficiency

	Argumentative writing			Personal writing	
	Low-level learners	High-level learners	Native speakers	Low-level learners	High-level learners
Minimum	2.99	2.61	0.59	7.30	8.33
25 <sup>th</sup> Percentile	10.13	6.43	2.71	42.17	37.76
Median	17.24	9.97	6.13	64.52	53.44
Mean	22.22	11.88	9.00	66.75	56.28
75 <sup>th</sup> Percentile	27.78	15.23	11.81	88.24	75.27
Maximum	200.00	55.05	44.81	222.22	157.14



**Figure 2.** Variability of leaners’ and native speakers’ frequency of *I* in argumentative writing (left) and personal writing (right).

The variability of individuals’ data is shown in the boxplot (Figure 2) representing each group (native speakers, low- and high-level learners) based on six significant

values (Table 3) required to draw a boxplot. In the boxplot, the middle line inside the box represents the median, and the length of the box symbolizes the interquartile range, from the 25<sup>th</sup> to the 75<sup>th</sup> percentile value. In other words, the middle 50 percent of data is captured within the box. The whiskers show the extent to which is usually considered acceptable extremes of data, calculated as plus and minus the interquartile range distance from the median. Finally, the individual dots represent values considered as outliers, which deviate far from the middle 50 percent of the data. Detailed analyses of the variability results are presented with the frequency results in the subsections below.

#### 4.1.1. Genre comparison of frequency and variability of *I*

The results of the frequency analysis of learners' use of *I* in argumentative writing showed that for both low and high proficiency groups, learners used *I* less often in argumentative writing compared to personal writing (Table 3; Figure 1). Low-level learners used *I* less frequently in argumentative writing (16.17 per 1,000) compared to personal writing (63.64 per 1,000) and the difference was statistically significant ( $\chi^2=3,765.007$ ,  $p<.001$ ). High-level learners also showed a similar pattern, using *I* less often in argumentative writing (10.14 per 1,000) compared to personal writing (54.75 per 1,000) ( $\chi^2=8,275.912$ ,  $p<.001$ ). The variability results revealed that there was less individual variation in argumentative writing compared to personal writing for both low-level and high-level learners. In other words, there was a general, large tendency for learners overall to use *I* less often in argumentative writing, while there was more variability across learners in using *I* in personal writing.<sup>2)</sup>

In contrast to studies that revealed learners' lack of sensitivity to genre differences (Deng et al. 2014), the results of the current study support the findings of previous works that have recognized learners' ability to differentiate language use according to genre conventions in terms of the use of *I* (Paquot et al. 2013) and other aspects (Way et al. 2000; Lu 2011; H Yoon 2006). Since the data of the two genres were collected from the same group of learners and compared against each other, the results indicate that learners overall refrained from using *I* often in argumentative writing compared to the personal writing genre. The fact that both low- and

---

2) It is interesting to note that the frequency difference found in learners' argumentative and personal writing resembled the comparable distinctions between genres found in native speakers' productions. For example, an analysis of native speakers' frequency of *I* in the British National Corpus revealed a significantly higher frequency in letters compared to school and university essays ( $\chi^2=5.368$ ,  $p<.001$ ).

high-level learners altered their use of *I* according to genre, generally using it less frequently in argumentative writing, suggests that learners are to some extent aware of genre differences. In addition, the more evident genre difference in high-level learners compared to learners with lower proficiency demonstrates the fact that as proficiency increases, learners acquire the ability to adjust their language use according to specific genres.

#### 4.1.2. Proficiency comparison of frequency and variability of *I*

The direct comparison of the frequency of *I* in learners' and native speakers' argumentative writing revealed that learners used *I* more frequently than native speakers (Table 2). Low-level learners used *I* significantly more often (16.17 per 1,000) than native speakers (4.19 per 1,000) ( $\chi^2=1,082.290$ ,  $p=.000$ ,  $p<.001$ ) and high-level learners also used *I* significantly more frequently (10.14 per 1,000) than native speakers (4.19 per 1,000) ( $\chi^2=441.563$ ,  $p<.001$ ). A meaningful difference was observed between learner groups as well, with low-level learners using *I* more often (16.17 per 1,000) compared to high-level learners (10.14 per 1,000) ( $\chi^2=335.843$ ,  $p<.001$ ).

The comparison of variability between the three groups (Table 3, Figure 2) demonstrated that native speakers showed less variability compared to both low-level and high-level learners. That is, native speakers as a group showed an overall tendency to use *I* less frequently compared to learners in argumentative writing. There was also a difference of variability between the low-level and high-level learner groups (Figure 2). The larger length of the box and greater number of outliers (the individual dots in the boxplot) in the low-level learner group compared to the high-level learner group indicated that there was greater variability in low-level learners' frequency of *I* compared to high-level learners. This means that while learners with higher proficiency had a more overall tendency to use less *Is*, low-level learners showed a greater variety in the choice to use *I*. There were some low-level learners who used *I* very frequently while other low-level learners used it less often.

Overall, in argumentative writing, learners used *I* more frequently than native speakers, which has also been reported by multiple studies (Y Choung & S-Y Oh 2017; Neff et al. 2004; Natsukari 2012; Recski 2004). Unlike learners, native speakers used less *Is* with low variability among individuals, implying that they overall refrained from using *I* in the argumentative genre. Yet, it is still significant to note that some learners did actually use *I* with similar frequencies to those of native speakers. This kind of pattern suggests that there may be developmental charac-

teristics in learners’ use of *I* in argumentative writing; some learners grasp how to use *I* effectively and thus use the pronoun with similar frequencies to native speakers while others lack the knowledge to do so and use it too often with less variation in their language. The fact that high-level learners used *I* less frequently, with less variability compared to low-level learners also supports the idea that knowledge of *I* is acquired through learners’ L2 development. Previous studies have reported that learners of higher grade levels in college showed lower frequencies of *I* compared to those in lower grade levels, but it was not clear whether it was specifically due to learners’ proficiency (Neff et al. 2004; McCrostie 2008). The differences found in the current study between low- and high-level learners present strong evidence that there is a proficiency effect in learners’ use of *I*. As proficiency increases, learners’ use of *I* in argumentative writing begins to approximate that of native speakers; learners gradually show less variability in their choices to use *I* as more and more learners begin to use *I* less often along with the increase in proficiency (Y Choung & S-Y Oh 2017). This finding counters the remark made by Recski (2004) that the frequent use of *I* is only observed in some learners instead of it being present in most learners. The current study showed that the distribution in fact differs according to learners’ proficiency, as learners with higher proficiency begin to show similar patterns to native speakers, using less *Is* in their writing.

#### 4.2. Collocates of *I*

Table 4 shows the top 25 most frequent collocates of *I* in learners’ and native speakers’ argumentative writing, and Table 5 presents the top 10 most frequent collocates of *I* in the personal essays of low-level and high-level learners.<sup>3)</sup>

**Table 4.** Top 25 most frequent collocates of *I* in learners’ and native speakers’ argumentative writing

N	Low-level learners		High-level learners		Native speakers	
	L1	R1	L1	R1	L1	R1
1	SO	THINK	BUT	THINK	AND	HAVE
2	BUT	AGREE	SO	AGREE	THAT	WAS
3	AND	DON'T	HOWEVER	DON'T	BUT	FEEL

3) Since the main focus of this paper is on the argumentative genre, the collocates of *I* in argumentative essays were examined in detail while those in personal writing were analyzed for the purpose of comparison with argumentative writing.

**Table 4.** Continued

Low-level learners			High-level learners		Native speakers	
N	L1	R1	L1	R1	L1	R1
4	WHEN	WAS	REASONS	DISAGREE	IF	THINK
5	YES	DISAGREE	AND	WAS	WHEN	WOULD
6	THAT	HAVE	WHEN	HAVE	BECAUSE	AM
7	BECAUSE	AM	WHY	STRONGLY	THOUGH	KNOW
8	NO	KNOW	THAT	BELIEVE	AS	BELIEVE
9	SCHOOLS	WANT	AS	AM	WHICH	HAD
10	DRIVING	CAN	THEREFORE	DO	TIME	DO
11	EXPERIMENTS	DO	INTERNET	KNOW	IT	COULD
12	SERVICE	CAN'T	WHAT	CAN	BEFORE	WILL
13	PUNISHMENT	WILL	CONCLUSION	SAW	MONEY	DON'T
14	REASONS	SAW	SERVICE	WANT	LIFE	CAN
15	INTERNET	HEARD	PUNISHMENT	ALSO	ALL	AGREE
16	HOWEVER	HAD	BECAUSE	SAID	WORLD	THOUGHT
17	THEREFORE	ALSO	DRIVING	WOULD	UNTIL	GOT
18	PEOPLE	REALLY	SCHOOLS	HEARD	DAY	GUESS
19	WHY	HATE	BUILDINGS	WILL	EVIL	DID
20	IF	FEEL	EXPERIMENTS	MENTIONED	SAID	DIDN'T
21	LIFE	SHOULD	SCHOOL	SUGGEST	PEOPLE	SAY
22	COURSE	INSIST	YES	COULD	KNOW	UNDERSTAND
23	TIME	ALWAYS	REASON	MEAN	WELL	REMEMBER
24	OPINION	LOVE	FIRST	FIRMLY	HOUSE	WERE
25	NOW	SAY	COURSE	FEEL	HOWEVER	CANNOT

\*L1 = one word left of *I* / R1 = one word right of *I*

**Table 5.** Top 10 most frequent collocates of *I* in leaners' personal writing

Low-level learners			High-level learners	
N	L1	R1	L1	R1
1	SO	LIKE	WHEN	WAS
2	WHEN	WAS	THAT	HAVE
3	AND	HAVE	AND	LIKE
4	THAT	WANT	SO	WANT
5	BECAUSE	THINK	IF	CAN
6	BUT	CAN	BECAUSE	COULD
7	IF	AM	WHY	THINK
8	TIME	USUALLY	BUT	USUALLY
9	SCHOOL	WENT	SCHOOL	HAD
10	WHY	GO	TIME	AM

\*L1 = one word left of *I* / R1 = one word right of *I*

#### 4.2.1. Genre comparison of collocates of *I*

The inspection of collocates of *I* in learners' argumentative writing showed various unique genre-specific characteristics when compared with those in personal writing. Learners showed similar patterns in the choice of words that preceded *I* in argumentative and personal writing. In both genres, *I* was generally used after conjunctions including *so*, *but*, *and*, *when*, and *that*. No striking genre-related contrasts were observed in the words that preceded *I*.

On the other hand, words that followed *I* in argumentative writing differed from those in personal writing. In the argumentative genre, learners frequently used *I* with words that are often used to deliver opinions such as *think*, *agree*, and *disagree*, while in personal writing, they often used *I* with *like*, *was*, *have*, *want*, and *can*, which are used to express personal state. This kind of difference indicates that compared to a more personal type of writing, learners tend to rely on words that can explicitly deliver their arguments, which can be seen as an attempt to achieve the goal of making clear claims according to the conventions of the argumentative genre. This suggests that learners are to some extent aware of the genre-specific characteristics of argumentative writing, and therefore overuse particular collocates of *I* which they would not use as often in other genres.

In particular, the overuse of *I think* in learners' written productions was more evident in the argumentative genre compared to personal writing: *think* was the most frequent collocate of *I* in argumentative writing for both low- and high-level learners, whereas it was the top 5 and 7 most frequent collocate in low- and high-level learners' personal writing, respectively. This kind of finding supports the results of previous studies that have reported learners' high dependency on the phrase *I think* when stating personal opinions (Y Choung & S-Y Oh 2017; Gilquin and Paquot 2008; Hyland & Milton 1997; Natsukari 2012; S-Y Oh 2007). The fact that learners used *I think* as well as *I agree/disagree* more frequently in argumentative writing than in personal writing suggests that learners rely on particular formulaic expressions when delivering their personal opinions, especially in the argumentative genre (H-J Yoon & Polio 2016).

#### 4.2.2. Proficiency comparison of collocates of *I*

According to the inspection of the top 25 collocates of *I* in argumentative writing (Table 4), learners and native speakers shared the most frequent words that preceded *I*. They included conjunctions such as *but*, *and*, *when*, *that*, and *because*, which commonly



serve the function of joining clauses. Yet, learners frequently used *so I* while native speakers rarely used such phrase. The frequent use of *so I* can be explained by learners' lack of sensitivity to the relatively casual register of the word *so* in rather formal type of writing compared to native speakers who refrain from using the conversational conjunction (E-J Lee 2004; S-Y Oh 2009; Šimčikaitė 2012; H-J Yoon 2006).

**Table 6.** Learners' sample concordances of frequent collocates (stating opinion)

lose small self-esteem than adults <b>so</b>	<b>I</b>	think the appropriate physical	<i>Low-level</i>
to repect like other humen. That's <b>why</b>	<b>I</b>	don't think animals must be used in	<i>Low-level</i>
but also social problems. <b>Therefore</b>	<b>I</b>	strongly urge that we should strictly	<i>High-level</i>
we could cost lesser money. <b>I</b>	<b>think</b>	that we must use their real name	<i>High-level</i>
cause more accidents. Therefore <b>I</b>	<b>disagree</b>	with the idea that drivers of	<i>High-level</i>

**Table 7.** Native speakers' sample concordances of frequent collocates (stating experience)

ever visited. <i>I</i> can admit that <b>before</b>	<b>I</b>	came to spain i was very ignorant
before, <i>I</i> had no choice but to wait <b>until</b>	<b>I</b>	was twenty years old to even begin
to follow because <i>I</i> am living proof. <b>I</b>	<b>have</b>	been working at an amusement
off to new york to have an abortion. <b>I</b>	<b>was</b>	surprised that she got pregnant but
to the profits of the county's transit. <b>I</b>	<b>feel</b>	that the city might lose more

The observation of the collocates that preceded *I* showed that learners also used *I* to restate their position such as in phrases, *so I*, *why I*, *therefore I*, *(in) conclusion I*, and *reason(s) I* (Table 6). In contrast, these collocates were not often present in native speakers' writing. For example, the raw frequency of the phrase *so I* was 286 for low-level and 204 for high-level learners while only one instance of the use was found in native speakers' data. In contrast to learners, native speakers' unique use of *I* was shown by phrases such as *before I*, *until I*, and *if I* (Table 7), which are involved in delivering personal anecdotes or suggesting hypothetical situations as part of examples that support arguments. Such difference in the words that precede *I* represents learners' unique tendency to clarify their arguments and reasons by explicitly presenting themselves in the writing and guiding the readers through the essay.

A critical difference was observed in learners' and native speakers' choices of words that followed *I* as well. Learners most frequently used words like *think*, *agree*, and *disagree* after *I* (Table 6), directly expressing their opinion or answering the given essay prompt in a straightforward way. The frequency of *I think* vastly differed between learners' (1,023 and 1,163 for low-level and high-level, respectively) and native speaker's (41) data, and the same tendency was observed for the phrase *I*

*agree/disagree* (low-level (241/86), high-level (226/169), native (9/2)) as well. Unlike learners, native speakers' most frequent words that came after *I* were *have*, *was*, and *feel* (Table 7), which are generally used to describe personal state or experience. The phrase *I feel*, for example, had a raw frequency of 44 in native speakers' corpus while only 17 in low-level and 20 in high-level learners'.

Learners' high frequency of the phrase *I think* and *I agree/disagree*, which serves a similar function, found in the current study resemble the patterns found in previous studies with learners of various L1s (Aijmer 2001, 2002; Granger 1998; Hasselgård 2009; S-Y Oh 2007; Paquot et al. 2013). Native speakers' frequent use of *I* to introduce personal experiences or hypothetical situations as examples to support their arguments also supports the findings of relevant previous works (Petch-Tyson 1998; Recski 2004; McCrostie 2008). As the native speaker and learner corpora were comparable in terms of the argumentative essay prompt topics, it can be understood that the particular overuse of phrases like *I think* in learners' productions and a relative low frequency in native speakers' data overall represent learners' heavy reliance on formulaic expressions as efforts to achieve the goal of the argumentative genre. This can be seen as learners having awareness of genre-related characteristics of the argumentative prose to some extent (Paquot et al. 2013; Way et al. 2000; Lu 2011; H Yoon 2006). The interpretation is further elaborated at a later part in the paper with reference to the different results found between low- and high-level learners.

Another interesting distinction found between learners and native speakers in their use of *I* was that learners used *I* with collocates that expressed certainty in epistemic modality compared to native speakers who used *I* with rather words denoting doubt or uncertainty (compare Tables 8 and 9). Learners used the phrase *(of) course I* while native speakers used the phrase *well I* relatively more often compared to each other. Also, for words that followed *I*, learners used verbs that contain heavy emotional meaning such as *hate*, *should*, *insist*, *love*, or intensifiers that strengthen following verbs, including *really*, *strongly*, and *firmly*. On the other hand, native speakers used *I* with hedges or words that reveal hesitation or softness in argument such as *feel*, *guess*, and *understand*.

**Table 8.** Learners' sample concordances of frequent collocates (strong words & boosters)

than banning smoking. <b>Of course,</b>	<b>I</b> agree that smoking indoors might	<i>High-level</i>
at buildings. Because, most of all,	<b>I</b> <b>hate</b> smoking. And also I think	<i>Low-level</i>
environment. But certainly, what	<b>I</b> <b>really</b> want to speak is that smokers	<i>Low-level</i>
their opinions freely. However,	<b>I</b> <b>strongly</b> believe that people have to	<i>High-level</i>

**Table 9.** Native speakers’ sample concordances of frequent collocates (hesitation markers & hedges)

, the Jeffrey Dahmer case). <b>Well,</b>	<b>I</b> believe that no matter what the
have the same negative attitude. <b>I</b>	<b>feel</b> that there are both values and
to be more dedicated to teaching. <b>I</b>	<b>guess</b> they have to be that way because
and would fit into a perfect size 1. <b>I</b>	<b>understand</b> that every woman may

In particular, learners’ use of the phrase *I feel* differed from that of native speakers. As Table 10 shows, while native speakers used *I feel* as a hesitation marker, learners used *feel* with its dominant meaning, to express personal emotion.

**Table 10.** Sample concordances of *I feel*

<i>Learners</i>		
the smoking man on the street and	<b>I feel</b> dizzy when I inhale the smoke they	<i>Low-level</i>
of course I was punished sometimes,	<b>I feel</b> very frightened and depressed. And	<i>Low-level</i>
Firstly, I must say that	<b>I feel</b> sad when I think about the situation	<i>High-level</i>
<i>Native speakers</i>		
may have AIDS and not even know it.	<b>I feel</b> this has really made people think	
promoting, I disagree with much of it.	<b>I feel</b> that chivalry is nearly dead because	
American people who lost their jobs.	<b>I feel</b> the American people have been	

It has been demonstrated in previous studies that learners tend to use more boosters while native speakers use more hedges compared to each other (Hyland & Milton 1997; Lee & Deakin 2016; Mirzapour & Mahand 2012; S-Y Oh 2007). The findings of the current study provide new evidence that supports such research, in that learners do exhibit a more assertive voice, and that they often do so by using *I* with markers of epistemic certainty while native speakers use *I* with softer words to express hesitation. Thus, learners, compared with native speakers, have the tendency to use *I* in contexts where they want to strengthen their arguments, using *I* as a strategy to emphasize the power of their claims.

Yet, when the two learner groups were observed separately according to proficiency level, high-level learners showed some uses of epistemic devices expressing probability or possibility with *I* in their essays, including phrases like *I would*, *I could*, and *I feel*, which were frequent in native speakers’ productions, while low-level learners had almost an absence of such phrases. Such difference between low- and high-level learners’ use of epistemic devices with *I* can be explained by the gradual development of epistemic modality observed in learners’ essays with the increase of proficiency

(S-Y Oh & Kang 2013). In other words, it is suggested that as proficiency increases, learners gradually acquire the ability to balance epistemic modality, refraining from constantly putting forth their opinions in their argumentative essays, and thereby leaving some room for negotiation with assumed readers.

One unique characteristic found in high-level learners' use of *I* was their frequent use of the phrase *what I*. High-level learners often used *what I* to refer back to a previously made statement and to clarify their arguments in different wordings (Table 11). This kind of collocate was rarely found in low-level learners' writing.

**Table 11.** High-level learners' sample concordances of *what I*

on the freedom of speech. <b>What</b>	<b>I</b> mean by this is that, we do not have to
might be afraid of the authorities. <b>What</b>	<b>I</b> mean is that, for example, one high
do a safe driving. To summarize <b>what</b>	<b>I</b> mentioned above, I am in the steadfast
However smokers should respect <b>what</b>	<b>I</b> said in advance. They can smoke in
more often than these time. So <b>what</b>	<b>I</b> want to say is understand smokers'

Low-level learners also showed a unique pattern in their use of *I*; they frequently used the phrases, *yes I* and *no I*. While high-level learners also showed some frequencies of *yes/no I*, low-level learners mostly used the phrases in the first sentence of the essay, as an attempt to directly answer the essay prompt, while high-level learners used them in the middle of the essay (Table 12). In detail, the low-level learner group used *yes I* in the beginning of sentences 53 times out of the total of 58 times they used the phrase, while high-level learners used it only five out of 26 times in the beginning of sentences. This indicates low-level learners' tendency to start their essay by answering the given prompt in a casual, conversational manner instead of shaping their response into a complete, independent essay.

Overall, high-level learners showed a general tendency to repeat previously stated

**Table 12.** Sample concordances of *yes I*

<i>Low-level learners</i>	
	<b>yes, I</b> agree that people must use their real
	<b>Yes. I</b> absolutely agree that smoking should be
	<b>Yes, I</b> agree with topic. smoking must be
<i>High-level learners</i>	
is safe if they use hands-free phone. <b>Yes</b>	<b>I</b> admit that it is a lot better than using
the owner of their IDs are them. <b>Yes,</b>	<b>I</b> agree that people must use their real
be used in medical experiments. <b>Yes,</b>	<b>I</b> think it is okay to use animals in medical

arguments with easier language using *what I (mean)* to make sure their arguments are conveyed to the reader, while low-learners focused on directly answering the writing prompt by beginning their essay with *yes/no I*. These patterns show distinct uses of *I* according to learners' proficiency and different strategies learners employ in achieving the goal of the argumentative genre, as they gradually build implicit knowledge on using the pronoun most effectively through experience in genre conventions and writing (Tardy 2006).

## 5. Conclusion

The present study has examined Korean EFL learners' use of *I* in argumentative writing, with a specific focus on how it differs from that of personal writing and according to learner proficiency. The results of the frequency, variability, and collocates analyses overall showed significant differences according to genre and proficiency. Compared with a more personal type of writing, learners used *I* less often and with more opinion-providing collocates in argumentative prose. The overall decrease in the frequency of *I* across most learners when writing the argumentative essay suggests that learners are to some extent aware of genre-specific characteristics, not as often using *I* as they did in a different genre. This resembled the distinct frequency of *I* observed in native speakers' productions. Yet, the overuse of identical phrases such as *I think* and *I agree/disagree* demonstrated that learners tend to rely on a limited number of phrases they feel comfortable using in order to most clearly achieve the purpose of delivering their arguments. The comparison between learners of low and high proficiency disclosed different patterns of such attentiveness to the argumentative genre. High-proficiency learners used *I* less frequently and with more appropriate epistemic modality compared to low-level learners, overall suggesting that as proficiency in the second language increases, the knowledge of using *I* with respect to the characteristics of argumentative writing is gradually acquired in the process. In other words, with more experience in the second language, learners gain a sense of how to use *I* appropriate to the register and tone associated with the argumentative genre. Thus, in terms of pedagogical implications, raising awareness of the particular functions *I* can serve in argumentative writing as well as providing opportunities to learn and practice pragmatic writing strategies that meet the expectations of specific genres can help learners acquire the knowledge of using *I* most effectively in their writing. In particular, improvements can be made in learners' use of *I* in phrases

other than formulaic expressions that learners often overuse to explicitly deliver their claim. Abundant exposure to and practice with various ways of incorporating personal presence in argumentative writing could help learners in writing more persuasive and powerful argumentative prose.

Altogether, the results of the current study contribute to the research on learners' use of the first person singular pronoun in multiple aspects. While most research has been conducted without respect to the genre-specific characteristics of argumentative writing, this study confirms the relatively less frequent use of *I* in the argumentative genre and also suggests that learner's more frequent use of *I* compared to native speakers' can be explained as learners' unique attempts to deliver their arguments with attentiveness to the purpose of the genre. Also, the current study demonstrates that learners' knowledge of using *I* in argumentative writing most effectively develops as learners' proficiency improves. The proficiency differences provide evidence to suggest that the knowledge of *I* is implicitly acquired as proficiency increases. Finally, the variability analysis and observations of *I* within specific contexts conducted in the present study reveal significant patterns regarding the use of *I* across individual learners as well as the general tendencies shown in learners' actual usages of *I*.

The limitations of the current study are as follows: i) only Korean EFL learners' data were analyzed, and therefore results may differ for learners with other L1s, ii) the essay prompts of the corpora used in the study were not exactly the same and thus results can show differences according to topics of the prompts, iii) personal essays of native speakers were not observed in detail to offer results on how the genre differences of learners differed from the distinctions in native speakers' productions. Further research on how learners use *I* according to their L1s, how learners differ in their use of *I* according to genres other than the ones observed in previous and current studies and systematically comparing learners' use of *I* with other personal pronouns could all lead to meaningful discoveries regarding the unique characteristics of learner's use of *I* in writing.

## References

- Abbuhl, Rebekha. (2012). Using self-referential pronouns in writing: The effect of explicit instruction on L2 writers at two levels of proficiency. *Language Teaching Research*, 16(4), 501-518.
- Akatori, Naoko. (2007). The excessive use of the first-person pronoun "I" in English

- compositions by Japanese students. *Nihonjin eigo gakushusha no hanashikotoba kakikotoba no corpus sakusei to sono goyoron teki taishobunseki* [Building learner corpora of Japanese learners of English and the contrastive analyses]. (Report for Grant-in-Aid for Scientific Research (B) No. 15320059). Tokyo: Brainsnetwork.
- Aijmer, Karin. (2001). I think as a marker of discourse style in argumentative student writing. *Gothenburg Studies in English*, 81, 247-257.
- Aijmer, Karin. (2002). Modality in advanced Swedish learners' written interlanguage. *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, 55-76.
- Altenberg, Bengt and Sylviane Granger. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics*, 22(2), 173-195.
- Choung, Yeram and Sun-Young Oh. (2017). A systemic functional study of thematic organization in the English writing of Korean college students. *English Teaching*, 72(3), 119-144.
- Daniel, T. and B. Kostic. (2017). RStats chi-square calculator. Retrieved from: <http://www.missouristate.edu/rstats/Tables-and-Calculators.htm>
- Dueñas, Pilar M. (2007). 'I/we focus on...': A cross-cultural analysis of self-mentions in business management research articles. *Journal of English for Academic Purposes*, 6(2), 143-162.
- Thompson, Geoff. (2013). *Introducing Functional Grammar*. Routledge.
- Gilquin, Gaëtanelle and Magali Paquot. (2008). Too chatty: Learner academic writing and register variation. *English Text Construction*, 1(1), 41-61.
- Granger, Sylviane. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. *Phraseology: Theory, Analysis, and Applications*, 145, 160.
- Granger, Sylviane. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research*, 1(1), 7-24.
- Hasselgård, Hilde. (2009). Thematic choice and expressions of stance in English argumentative texts by Norwegian learners. *Corpora and Language Teaching*, 33, 121-139.
- Henderson, Alice and Robert Barr. (2010). Comparing indicators of authorial stance in psychology students' writing and published research articles. *Journal of Writing Research*, 2(2), 245-264.
- Hinkel, Eli (1999). Objectivity and credibility in L1 and L2 academic writing. *Culture in Second Language Teaching and Learning*, 90108.
- Hyland, Ken. (1990). A genre description of the argumentative essay. *RELJ Journal*, 21(1), 66-78.
- Hyland, Ken. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 34(8), 1091-1112.
- Hyland, Ken. (2008). Persuasion, interaction and the construction of knowledge: Representing self and others in research writing. *International Journal of English Studies*, 8(2), 1-23.
- Hyland, Ken and John Milton. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*, 6(2), 183-205.
- Kuo, Chih-Hua. (1999). The use of personal pronouns: Role relationships in scientific journal articles. *English for Specific Purposes*, 18(2), 121-138.

- Lee, Eun-Joo. (2004). A corpus-based analysis of the Korean EFL learners' use of conjunctive adverbials. *English Teaching*, 59(4), 283-301.
- Lee, Joseph J. and Lydia Deakin. (2016). Interactions in L1 and L2 undergraduate student writing: Interactional metadiscourse in successful and less-successful argumentative essays. *Journal of Second Language Writing*, 33, 21-34.
- Lu, Xiaofei. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, 45, 36-62.
- Martínez, Illiana A. (2005). Native and non-native writers' use of first person pronouns in the different sections of biology research articles in English. *Journal of Second Language Writing*, 14(3), 174-190.
- McCrostie, James. (2008). Writer visibility in EFL learner academic writing: A corpus-based study. *ICAME Journal*, 32(April), 97-114.
- Mirzapour, Fatemeh and Mohammad R. Mahand. (2012). Hedges and boosters in native and non-native library and information and computer science research articles. *3L: Language, Linguistics, Literature*, 18(2).
- Natsukari, Sayo. (2012). Use of *I* in essays by Japanese EFL learners. *JALT Journal*, 34(1), 61-78.
- Neff, JoAnne, Francisco Ballesteros, Emma Dafouz, Francisco Martínez, and Juan-Pedro Rica. (2004). Formulating writer stance: A contrastive study of EFL learner corpora. *Language and Computers*, 52, 73-90.
- Oh, Sun-Young. (2007). A corpus-based study of epistemic modality in Korean college students' writings in English. *English Teaching*, 62(2), 147-175.
- Oh, Sun-Young. (2009). How to express causality: A comparison of native and learner corpora. *Korean Journal of the Applied Linguistics*, 25(1), 321-347.
- Oh, Sun-Young and Suk-Jin Kang. (2013). The effect of English proficiency on Korean undergraduates' expression of epistemic modality in English argumentative writing. *The Journal of Asia TEFL*, 10(4), 97-132.
- Paquot, Magali, Hilde Hasselgård and Signe O. Ebeling. (2013). Writer/reader visibility in learner writing across genres. A comparison of the French and Norwegian components of the ICLE and VESPA learner corpora. *Twenty Years of Learner Corpus Research: Looking Back, Moving Ahead*, 377-387.
- Petch-Tyson, Stephanie. (1998). Writer/reader visibility in EFL written discourse. *Learner English on Computer*, 107-118.
- Recski, Leonardo J. (2004). Expressing standpoints in EFL written discourse. *Revista Virtual de Estudos da Linguagem*, 3(2).
- Rhee, Seok-Chae and Chae Kwan Jung. (2014). Compilation of the Yonsei English Learner Corpus (YELC) 2011 and its use for understanding current usage of English by Korean pre-university students. *The Korea Contents Society*, 14(11), 1019-1029.
- RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Retrieved from: <http://www.rstudio.com/>



- Sanz, Rosa L. (2006). 'I will argue that': First person pronouns as metadiscoursal devices in research article abstracts in English and Spanish. *ESP across Cultures*, 3, 23-40.
- Scott, Mike. (2017). WordSmith Tools version 7, Stroud: Lexical Analysis Software.
- Šimčikaitė, Alė. (2012). Spoken discourse markers in learner academic writing. *Studies about Languages*, 20, 24-48.
- Storch, Neomy. (2009). The impact of studying in a second language (L2) medium university on the development of L2 writing. *Journal of Second Language Writing*, 18(2), 103-118.
- Tang, Ramona and Suganthi John. (1999). The 'I' in identity: Exploring writer identity in student academic writing through the first person pronoun. *English for Specific Purposes*, 18, S23-S39.
- Tardy, Christine M. (2006). Researching first and second language genre learning: A comparative review and a look ahead. *Journal of Second Language Writing*, 15, 79-101.
- Way, Denise P., Elizabeth G. Joiner and Michael A. Seaman. (2000). Writing in the secondary foreign language classroom: The effects of prompts and tasks on novice learners of French. *Modern Language Journal*, 84, 171-184.
- Yoon, Hyunsook. (2006). A corpus-based analysis of connectors in Korean students' essay writing. *Korean Journal of Applied Linguistics*, 22(2), 159-178.
- Yoon, Hyung-Jo and Charlene Polio. (2016). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, 51(2), 275-301.

Eun Kyoung Lee

MA Student

The Interdisciplinary Program in Cognitive Science

Seoul National University

1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea

E-mail: rosa303@snu.ac.kr

Sun-Young Oh

Professor

Dept. of English Language Education

Seoul National University

1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea

E-mail: sunoh@snu.ac.kr

Received: July 8, 2018

Accepted: July 23, 2018