English Language & Literature Teaching, *Vol. 30*, No. 4, Winter 2024
https://doi.org/10.35828/etak.2024.30.4.27

The Relationships between Lexical Sophistication and Second Language Writing and Speaking

Minkyung Kim
(Korea National University of Education)

Kim, Minkyung. (2024). The relationships between lexical sophistication and second language writing and speaking. *English Language & Literature Teaching, 30*(4), 27-48.

This study examines the relationship between lexical sophistication and the scores of second language (L2) persuasive writing and speaking. The dataset included 120 essays and 120 speeches produced by university-level learners of English as a foreign language. The essays were written as persuasive texts, while the speeches involved dialogues where learners were required to persuade their listeners. Lexical sophistication features were analyzed based on word frequency, psycholinguistic properties (i.e., concreteness, imageability, and age of acquisition), and word associations. Regression analyses revealed that word association scores were the strongest predictor for both L2 writing and speaking score, emphasizing the importance of the word association dimension of lexical sophistication across modalities. The model predicting L2 writing scores showed that higher-rated persuasive essays contained more low-frequency content words and words with restricted contextual usage, explaining 15.8% of the variance in writing scores. The model predicting L2 speaking scores found that higher-rated persuasive speeches included more words with restricted contextual usage, explaining 26.4% in speaking scores. Overall, these findings suggest that while word associations are broadly important, specific lexical features such as frequency and contextual restrictions play varying roles in different modalities of L2 persuasive tasks.

[L2 writing/L2 speaking/lexical sophistication/word frequency/word association]

I. INTRODUCTION

An essential objective of second language (L2)[1] learning is to engage effectively in both written and spoken communication. Producing successful L2 output in these modalities typically requires advanced lexical use and more sophisticated sentence structures (e.g., Bachman & Palmer, 1996). In particular, examining learners' use of sophisticated lexical items is of interest because it reflects their underlying lexical systems (Norris & Manchón, 2012). For instance, the presence of more sophisticated lexical items in either L2 written or spoken samples may indicate a learner's advanced lexical knowledge (Nation, 2001).

While lexical sophistication is widely regarded as essential for both written and spoken communication (e.g., Crossley et al., 2011a, 2011b; Hsieh & Wang, 2019; Kyle & Crossley, 2016; Laufer & Nation, 1995; Lu, 2012; Uchihara & Clenton, 2020), its specific roles in determining the quality of both L2 writing and speaking, especially using comparable indices, remain underexplored. Such research can enhance our understanding of how lexical sophistication influences L2 performance. Investigating lexical sophistication across both modalities within the same genre (e.g., persuasive discourse) can theoretically reveal whether its influence on the evaluation of L2 performance is consistent across modalities. Practically, if the same lexical sophistication features affect scores in both L2 writing and speaking, they could serve as focal points for targeted L2 instruction. Therefore, the purpose of this study is to examine how the use of lexical sophistication features in L2 writing and speaking is related to performance scores in both modalities.

II. LITERATURE REVIEW

1. Lexical Sophistication

Lexical sophistication refers to the proportion of relatively advanced words found in a written or spoken sample (Read, 2000). Examining the degree of lexical sophistication in learners' production is important because the use of more sophisticated words can reflect L2 learners' greater lexical knowledge (Nation, 2001). Sophisticated words are typically conceptualized as low-frequency words (Laufer & Nation, 1995). Beyond frequency measures, the conceptualization of lexical sophistication has been expanded to include words that are used in less diverse contexts (McDonald & Shillcock, 2001), words that are less concrete and less imageable (Salsbury et al., 2011), and words that are acquired in a later age (Kuperman et al., 2012), and words frequently used in academic contexts

[1] In this study, L2 refers to a language other than the learner's mother tongue.

(Coxhead, 2000).

Lexical sophistication in L2 written and oral production can influence raters' judgments of the quality of L2 speaking and writing performance, with its quality being operationalized by the scores assigned by raters to each production (Witte & Faigley, 1981). Indeed, lexical sophistication has been found to be related to L2 writing scores (e.g., Crossley et al., 2011a; Kyle & Crossley, 2016) and L2 speaking scores (e.g., Crossley & McNamara, 2013; Lu, 2012). In this study, among various lexical sophistication features, three dimensions are chosen to examine their potential different roles on predicting L2 writing and speaking scores: frequency, psycholinguistic properties, and word association scores. These three dimensions were chosen due to the mixed findings regarding their impact on L2 writing and speaking scores. Related studies are presented below.

2. Word Frequency and L2 Writing and Speaking Scores

Frequency is a key indicator of lexical sophistication and has been extensively researched, as words that L2 learners encounter more often are usually acquired earlier than those they encounter less frequently (Laufer & Nation, 1995; Read, 2000). Infrequent words are often considered more advanced and require greater cognitive effort to process due to their rare usage or presence in specialized contexts. These words are more difficult to retrieve from memory because their mental representations are weaker. Frequency measures are typically based on word counts from large reference corpora.

The relationships between word frequency and L2 writing scores have been widely supported, such that L2 written production that includes lower frequency words tend to receive higher quality scores (e.g., Crossley et al., 2011a; Crossley & McNamara, 2012; Kim et al., 2018; Laufer & Nation, 1995). For example, Kim et al. (2018) examined the lexical sophistication features in L2 (English) learners' writing and reported that content word frequency measures served as a negative predictor of L2 writing proficiency.

However, the relationships between word frequency and human judgment of L2 spoken production remain inconclusive. Some studies reported their negative correlations (e.g., Crossley et al., 2011b, Crossley & McNamara, 2013) as L2 writing studies have reported, whereas others reported no relations (e.g., Hsieh & Wang, 2019; Lu, 2012). For example, Crossley and McNamara (2013) analyzed transcribed monologic speech samples from the Test of English as a Foreign Language (TOEFL). They demonstrated that content word frequency was a significant predictor of speaking proficiency, with speeches containing more infrequent words receiving higher scores. In contrast, Uchihara and Clenton (2020) analyzed the speeches of advanced L2 (English) learners based on conversational communication with interviewers. They reported that the use of less frequent words did not correlate with raters' perceptions of L2 speech proficiency. The discrepancy in the impact

of word frequency on L2 speaking scores may be partially attributed to differences in speech types (monologic vs. interview-based), speaking genres (e.g., picture description vs. opinion-giving), and the range of learners' proficiency levels (e.g., learners from various proficiency levels vs. only advanced learners). Further research is needed to explore the influence of less frequent word usage on L2 speech ratings.

3. Psycholinguistic Properties of Words and L2 Writing and Speaking Scores

Another dimension of lexical sophistication involves the psycholinguistic properties of words, which reflect how meaning and experience shape their mental representation. These properties include concreteness, imageability, and age of acquisition (AoA; Kuperman et al., 2012). Such measures are often based on native speakers' subjective intuitions.

Concreteness refers to how tangible or perceptible a word's mental representation is. Concrete words (e.g., *dog*) are generally learned earlier than abstract words (e.g., *freedom*) in both first (L1) and L2 because they have stronger semantic representations and rely less on context (Paivio, 1986; Kaushanskaya & Rechtzigel, 2012). Concreteness is closely related to imageability, which describes how easily a word evokes a mental or sensory image. The processing advantage of concrete and imageable words is explained by dual-coding theory (Paivio, 1971), which suggests these words have both verbal and imagery-based codes, making them easier to retrieve. AoA refers to the age at which a word is first learned. Words acquired earlier (e.g., *spoon*) are processed more efficiently than those learned later (e.g., *grief*) due to their prolonged presence in the mental lexicon (Kuperman et al., 2012). As a result, words that are more concrete, more imageable, or learned earlier are typically viewed as less sophisticated in terms of lexical processing.

Previous studies have examined the relationships between concreteness, imageability, and AoA measures, and L2 writing production, reporting that L2 written production with higher quality scores tends to feature the greater use of less concrete words, less imageable, and words acquired at a later age tend to receive (e.g., Crossley et al., 2011a; Kim et al., 2018; Zhang et al. 2022). For example, Crossley et al. (2011a) analyzed written journals from 180 beginning to intermediate L2 (English) learners and 60 native speakers. They reported a negative correlation between word concreteness and L2 writing scores, as well as between word imageability and L2 writing scores. In addition, Zhang et al. (2022) examined 260 argumentative essays produced by English as a foreign language (EFL) learners, showing that AoA scores were positively correlated with L2 essays scores. These findings indicate that raters' evaluation of L2 written production is influenced by the sensory experience (concreteness and imageability) and the cumulative exposure (AoA).

Prior research has explored the relationships between concreteness and imageability

measures and L2 spoken production (e.g., Crossley et al., 2011b; Hsieh & Wang, 2019), but the findings are mixed. Crossley et al. (2011b) and Hsieh and Wang (2019) found negative relationships, while Eguchi and Kyle (2020) reported positive correlations. Specifically, Crossley et al. (2011b) analyzed naturalistic spoken samples from 180 beginning to intermediate L2 (English) learners and 60 native speakers, finding negative correlations between word concreteness and L2 lexical proficiency scores, as well as word imageability and proficiency scores. In contrast, Eguchi and Kyle (2020) examined oral proficiency interviews from 1,281 L2 (English) learners, involving tasks such as self-introductions, descriptions, and role plays. Their study found that a factor combining concreteness and imageability was predictive of speaking scores, with higher-rated spoken samples containing more concrete and more imageable content words.

In comparison, the relationship between AoA and L2 spoken production has received little attention. To my knowledge, only one study (Hsieh & Wang, 2019) has investigated this connection. They examined the monologic spoken production of 179 L2 (English) learners, including tasks such as picture narration and listen/speak exercises, and found that higher-rated responses tended to include words acquired at a later age.

4. Word Association and L2 Writing and Speaking Scores

Although frequency information is derived from word collections, it is criticized for not accounting for how words are stored and retrieved in the mental lexicon, which serves as the mental repository of words and their meanings (De Deyne et al., 2018; Kiss et al., 1973; Nelson et al., 2004). Such mental lexicon can be partly reflected by word association measures, which are collected through individuals' spontaneous word generation (e.g., when presented with a word *home*, individuals might respond with *house* or *family*). That is, word association measures can capture the connections among words as they naturally occur in an individual's mind not only for L1 speakers but also L2 learners (Fitzpatrick & Thwaites, 2020)

Word association measures can also be a proxy for lexical sophistication since words that elicit many associations are likely to be well-known and central in the lexicon, suggesting they are easier to process and retrieve. Conversely, words that produce fewer associations may be more peripheral, less frequently activated, and harder to integrate into an individual's lexicon (Nelson et al., 2004). By extension, learner production that contains more words having fewer lexical associations, which are hard to process, may be better evaluated by human raters than that includes more words occurring in many contexts. However, such assumptions have not been investigated much (c.f., Berger et al., 2017; Zhang, 2022).

Among a few studies that examined the relationships between word associations and L2

production, in examining EFL learners' 686 argumentative essays, Zhang (2022) demonstrated that a word association measure for content words showed a negative correlation with L2 writing scores. On the other hand, Berger et al. (2017) examined 240 naturalistic spoken samples produced by L2 (English) learners and native speakers of English, reporting that an associative measure (Nelson et al., 2004) was negatively correlated with raters' judgment of lexical proficiency. Taken together, these studies indicate that raters' impressions of either spoken production (Berger et al., 2017) or written production (Zhang, 2022) were positively impacted by L2 production containing words that have fewer associated words and are less connected to other concepts in individuals' minds.

5. Current Study

In summary, prior research has identified relationships between lexical sophistication and L2 writing scores, as well as between lexical sophistication and L2 speaking scores. However, the findings, particularly in the area of L2 speaking, remain inconclusive. Furthermore, to the best of my knowledge, no study has concurrently investigated these relationships within the same genre using identical lexical sophistication indices across both modalities. Exploring the relationships between lexical sophistication and both L2 writing and speaking within the same genre (i.e., persuasive genre) could provide deeper insights into the role of lexical sophistication across modalities. If certain lexical sophistication features similarly influence scores in both L2 writing and speaking, the findings could indicate that consistent lexical sophistication constructs are assessed across both modalities. Furthermore, demonstrating that writing and speaking share common lexical predictors suggests that enhancing lexical sophistication in one modality could positively impact the other. In addition, examining the persuasive genre in both modalities could be of particular interest, as it is regarded as a challenging discourse. This genre demands that students use more formal language and articulate their ideas in a logical and persuasive manner (Uccelli et al., 2013). Additionally, from a pedagogical perspective, if certain lexical sophistication domains influence both L2 writing and speaking scores in a similar manner, these domains could be prioritized as key targets for lexical development. Thus, this study is guided by the following two research questions:

1. To what extent can lexical sophistication features predict scores of EFL persuasive writing?
2. To what extent can lexical sophistication features predict scores of EFL persuasive speaking?

III. METHOD

1. Data

The data for this study were drawn from the International Corpus Network of Asian Learners of English Global Rating Archive (ICNALE GRA) (Ishikawa, 2020). The ICNALE GRA includes two datasets created by two separate groups: one containing 140 speeches and another containing 140 essays. Out of the 140 samples in each dataset, four speeches and essays were produced by native English speakers, and 16 were produced by English as a second language (ESL) learners, so these were excluded from the analysis. The study analyzed the remaining 120 samples in each dataset, which were produced solely by EFL students. These EFL students came from six Asian regions, with 20 students each from China, Indonesia, Japan, Korea, Taiwan, and Thailand. The learners' proficiency levels, assessed using either self-reported standardized test scores or a vocabulary size test and aligned with the Common European Framework of Reference (CEFR), ranged from A2 (Waystage) to B2 (Vantage). Therefore, the EFL learners in the ICNALE GRA were categorized as having beginner to intermediate proficiency levels. Details about the learners' gender or age were not provided.

The speaking data in the ICNALE GRA were initially collected as part of the broader ICNALE Spoken Dialogue corpus (Ishikawa, 2019). This corpus featured different activities in a dialogue format between an L2 learner and an interviewer (often the learner's instructor), including activities such as descriptions, role-plays, and reflections. While the entire dialogue lasted an average of 37 minutes, specific durations for each activity were not provided. The speaking data analyzed in this study, drawn from the ICNALE GRA, focused on role-plays where L2 learners had to persuade their college supervisors to allow them to continue working part-time. The transcripts for this data provided word-for-word records of the learners' spoken language. These transcripts were cleaned by removing unnecessary orthographic markings, such as fillers (e.g., *hmm*), interjections (e.g., *ah*), and repeated words from false starts (e.g., *the, the, the job* was simplified to *the job*), as these were not essential for analyzing the learners' lexical and phraseological use.

The writing data in the ICNALE GRA were initially collected from the larger ICNALE Written Essays corpus (Ishikawa, 2013). The essay prompt used in this study (i.e., the writing section of the ICNALE GRA) asked: "Do you agree or disagree with the following statements? Use reasons and specific details to support your opinion. It is important for college students to have a part-time job." Students were instructed to clearly express their opinions, providing reasons and examples to support their views. They were asked to write an essay of 200 to 300 words within a time frame of 20 to 40 minutes, using MS Word or a similar word processing tool.

It is important to note that the speeches and essays in the ICNALE GRA focused on the same topic, which was college students working part-time jobs. In addition, both were designed to persuade their respective audiences, whether through reading or listening. As a result, comparing these two forms was deemed appropriate due to their shared subject matter (part-time jobs) and persuasive genre.

The ICNALE GRA contains ratings, each evaluated by 80 raters who were either native English speakers or highly proficient English learners. Two types of rubrics were used for evaluation: holistic rubrics and analytic rubrics. The analytic rubrics consisted of 10 subcategories: intelligibility, complexity, accuracy, fluency, comprehensibility, logicality, sophistication, purposefulness, willingness, and involvement. Both holistic and analytic assessments had a maximum score of 100 points each. In this study, we used the average overall rating scores, which are the mean values of both holistic and analytic ratings, with a maximum possible score of 100 points. For more details on the rating procedures, refer to Ishikawa (2020).

2. Lexical Sophistication Measures

To measure lexical sophistication, six metrics were employed. Function words were excluded from the analysis because they typically serve a grammatical or syntactic purpose with contributing minimum semantic meaning (Corver & van Riemsdijk, 2001). Semi-lexical categories. Berlin: Mouton de Gruyter). Therefore, only the lexical sophistication of content words (CWs) was assessed. Lexical features for each CW were calculated using the Tool for the Automatic Analysis of Lexical Sophistication 2.0 (TAALES 2.2; Kyle et al., 2018), a natural language processing tool.

Word frequency was assessed using the spoken and academic subsections of the Corpus of Contemporary American English (COCA; Davies, 2009). Recognizing differences in lexical use between speaking and writing (Dang et al., 2017), we included both subsections. We assumed that frequency indices from the spoken and academic sections would align more closely with L2 learners' speaking and writing production, respectively. The frequency scores were expressed as log-transformed values.

Word concreteness and imageability were derived from the MRC Psycholinguistic Database (Coltheart, 1981) concreteness and imageability. In this database, participants rated words for concreteness and imageability using a 7-point interval scale. Both scales were adjusted to range from 100 to 700, with higher values indicating greater concreteness or imageability. In the database, 8,228 words have concreteness scores, and 9,240 words have imageability scores. For instance, the word *banana* has concreteness and imageability scores of 633 and 644, respectively, while *intention* scores 285 and 286 on these dimensions. It is important to note that a word can be concrete without being highly

imageable, and vice versa. For example, *saltpetre* is relatively concrete, with a score of 525, but has low imageability, scoring 283. Conversely, *infinity* is less concrete, with a score of 204, but more imageable, with a score of 469.

Next, age-of-acquisition scores were calculated using data from Kuperman et al. (2012), which reflect the average age in years at which native English speakers believe they first learned a word. Most of these ratings range from ages five to 14. Given that age-of-acquisition effects (i.e., words learned earlier are processed more quickly) in a L1 tend to influence L2 learners to some extent (e.g., Izura et al., 2011), words acquired later by L1 speakers may also be perceived as more difficult for L2 learners.

Lastly, a free association index based on the University of South Florida (USF) stimuli count norm (Nelson et al., 2004) was selected. This norm reflects the number of distinct stimulus words that prompted a specific response word during a free association task. For example, the word *car* was triggered by 259 stimulus words, while *fate* was prompted by only five. Thus, words with higher USF norm scores are typically used in a broader range of contexts.

The score for each index in each sample was calculated as the mean index score, determined by dividing the sum of the index scores for the words in the sample by the number of words in that sample that had corresponding scores. Words that were not found in the list for each index were excluded from the calculation of the mean scores.

3. Statistical Analysis

To address the research questions, two regression analyses were conducted with lexical sophistication features as independent variables to predict L2 writing scores (research question 1) and L2 speaking scores (research question 2). The process began by checking the normality of all variables, ensuring that skewness values fell between -3 and +3, and kurtosis values ranged from -10 to +10 (Kline, 2011). Correlational analyses were then carried out to explore the relationships between lexical sophistication features and both writing and speaking scores. In addition, multicollinearity among the lexical sophistication features was assessed, with correlations exceeding .90 flagged as potential issues. After these preliminary steps, stepwise regression analyses were conducted to determine the predictive power of lexical sophistication features on L2 writing scores (for research question 1) and L2 speaking scores (for research question 2). Only variables that significantly correlated with the dependent variables were included to ensure relevant predictors were used. To validate the models, 10-fold cross-validation was applied to assess their performance on unseen data. The assumptions for regression, including linearity, normality of residuals, absence of multicollinearity, and homoscedasticity, were tested and met. All analyses were performed using the R software (R Core Team, 2020).

IV. RESULTS

1. Descriptive Statistics and Correlations

Table 1 shows descriptive statistics for lexical sophistication features and scores.

TABLE 1
Means (*SD*s) for Lexical Features and Scores and Correlations with Scores

Feature	L2 writing (n = 120) Mean (*SD*)	L2 writing (n = 120) Correlation with scores	L2 speaking (n = 120) Mean (*SD*)	L2 speaking (n = 120) Correlation with scores
Score	51.43 (10.62)	-	49.83 (13.55)	-
COCA academic frequency CW (log)	2.50 (.11)	-.19*	2.41 (.15)	-.28**
COCA spoken frequency CW (log)	2.50 (.15)	-.33**	2.69 (.17)	-.21*
MRC word concreteness CW	353.46 (14.46)	-.11	339.04 (23.31)	-.41**
MRC word Imageability CW	385.73 (14.90)	-.12	370.24 (25.00)	-.45**
Age of acquisition CW	5.76 (.31)	.21*	5.24 (.35)	.04
Word association USF CW	40.18 (7.10)	-.35**	43.47 (13.51)	-.51**

*p < .05; **p < .01

2. Predicting L2 Writing Scores

To answer research question 1, a stepwise multiple regression analysis was conducted, using the four lexical sophistication features with significant correlations as predictors for L2 writing scores: COCA academic frequency CW (log), COCA spoken frequency CW (log), Age of acquisition CW, and Word association USF CW. No multicollinearity was detected among these features, allowing all four to be included in the regression model. This analysis resulted in a significant model (see Table 2), as indicated by F (2, 117) = 10.98, p < .01. Two predictors were identified as statistically significant: Word association USF CW and COCA spoken frequency CW (log). The findings suggest that higher-rated L2 written texts generally used fewer words with broad contextual usage and fewer high-frequency content words from a spoken corpus. Together, these two variables explained 15.80% of the variance in L2 writing scores. In a follow-up 10-fold cross-validation, the model remained consistent, accounting for 17.38% of the variance.

TABLE 2
Regression Model for Predicting L2 Writing Scores

Variable	β	*SE*	*B*	*p*	R^2	R^2 change
Word association USF CW	-.38	.14	-2.67	< .01	.12	-
COCA spoken frequency CW (log)	-15.45	6.64	-2.33	< .05	.16	.04

Notes. Constant = 105.19; β = unstandardized beta; *SE* = standard error; *B* = standardized beta

2. Predicting L2 Speaking Scores

To address research question 2, a stepwise multiple regression analysis was conducted using lexical sophistication features. Five indices demonstrated significant correlations with L2 speaking scores: Word Association USF CW, MRC word concreteness, MRC word imageability, and two frequency scores. Due to multicollinearity between MRC word concreteness and word imageability (r = .95), only word imageability, which had a stronger correlation with speaking scores, was included in the regression model, while word concreteness was excluded. As a result, the remaining four indices were used in the stepwise regression analysis. The analysis yielded a significant model (see Table 3), indicated by by $F(1, 118) = 42.43$, $p < .01$. One predictor was identified as statistically significant: word association USF CW. This finding indicates that higher-rated L2 spoken responses tended to feature fewer content words with broad contextual usage. The variable accounted for 26.45% of the variance in L2 speaking scores. In a follow-up 10-fold cross-validation, the model remained stable, explaining 27.42% of the variance.

TABLE 3
Regression Model for Predicting L2 Speaking Scores

Variable	β	*SE*	*B*	*p*	R^2
Word association USF CW	-.52	.08	-6.51	< .01	.26

Notes. Constant = 72.25; β = unstandardized beta; *SE* = standard error; *B* = standardized beta

V. DISCUSSION

1. Lexical Sophistication and L2 Writing Scores

The first research question addressed how the use of lexical sophistication features in L2 writing predicted L2 writing scores. Results indicated that two indices, word association and word frequency, were retained in the final regression model. This finding indicates that higher-rated L2 written texts tend to include more words with restricted contextual usage, aligning with previous L2 writing research (Zhang, 2022). However, it extends prior research by demonstrating that word association scores emerged as the strongest predictor of EFL persuasive writing scores, explaining approximately 12% of the variance. This suggests that, in the context of persuasive writing, word associations, reflecting the natural lexical connections between words in an individual's mind (Fitzpatrick & Thwaites, 2020), may have been the most influential lexical sophistication feature for raters. If a written sample includes words that occur in more limited contexts and are therefore more challenging to activate, it is possible that the sample may be perceived as more

sophisticated, leading to higher scores being assigned.

In addition, frequency was a significant predictor of L2 writing scores, accounting for approximately 4% of the variance. Higher-rated L2 writing samples tended to contain more low-frequency content words from a spoken corpus. This finding aligns with prior L2 writing research (e.g., Crossley et al., 2011a; Crossley & McNamara, 2012).

To further illustrate, Tables 4 and 5 present excerpts from lower-rated and higher-rated writing samples, respectively. The lower-rated sample features more words with higher frequency scores and higher word association scores (e.g., *important, life, differen*t). In contrast, the higher-rated sample includes more words with lower frequency scores and lower word association scores (e.g., *negative, management, entry*).

TABLE 4
Lower-Rated Written Sample Excerpt

Excerpt	Content word	Frequency score (COCA spoken, log)	Word association score (USF)
[...] **Second**, a part-time job is **kind** of **social life.** Therefore, they are **experience** to **social life** and after their **graduate university**, it is **important** to **experience social life.** [...]	second	2.44	15
	part	2.75	30
	time	3.26	119
	job	2.55	74
	kind	2.94	36
	social	2.12	17
	life	2.84	107
	experience	2.15	8
	social	2.12	17
	life	2.84	107
	graduate	1.23	3
	university	2.19	8
	important	2.60	41
	experience	2.15	8
	social	2.12	17
	life	2.84	107
	M	2.44	44.71

Notes. Excerpt from written sample *GRA_PTJ0_055_ORIG*, whose score is 33.12. Content words are presented in bold.

TABLE 5
Higher-Rated Written Sample Excerpt

Excerpt	Content word	Frequency score (COCA spoken)	Word association score (USF)
[...] Though there are some **negative aspects** of **part time jobs** for **college students** like over **exertions** and some **sort** of **tensions,** these **jobs provide** the **students** with the much **needed experience** which **comes handy** [...]	negative	1.68	12
	aspects	1.21	1
	part	2.75	30
	time	3.26	119
	jobs	2.23	74
	college	2.08	41
	students	2.08	1
	exertions	-.52	n/a
	sort	2.64	3
	tensions	1.02	6
	jobs	2.23	74
	provide	1.91	2
	students	2.08	1
	needed	2.02	3
	experience	2.15	8
	comes	2.42	1
	handy	.44	1
	M	1.86	23.56

Notes. Excerpt from written sample *GRA_PTJ0_126_ORIG,* whose score is 69.21. Content words are presented in bold.

On the other hand, the word frequency index derived from the academic corpus was not included as a significant predictor in the regression model, despite showing a significant correlation with writing scores (r = -.19, p < .05). The fact that frequency scores from the spoken corpus proved to be a better predictor than those from the academic corpus suggests that frequency effects become more prominent when assessing the lexical use in L2 persuasive writing samples using a spoken corpus. This may be because the spoken corpus, primarily consisting of informal language, associates higher frequency scores with less formal words. As a result, the frequent use of such informal language, typical of spoken contexts, may negatively influence the evaluation of L2 persuasive writing, which demands more formal language use.

Next, the age of acquisition index was also not included as a significant predictor, despite showing a significant correlation with writing scores (r = -.21, p < .05). This finding suggests that, when assessing the relationship between lexical sophistication and L2 persuasive writing scores, other metrics, such as word associations and word frequency, may serve as more reliable indicators of lexical sophistication than age-of-acquisition norms.

Finally, the word concreteness index and the word imageability index were not significantly correlated with L2 writing scores, contrasting with previous research that

reported negative relationships (e.g., Crossley et al., 2011a). This discrepancy could be attributed to differences in writing genres. For instance, Crossley et al. (2011a) analyzed L2 written journals, where the use of more abstract or less imageable words may be seen as a mark of advanced language use. In contrast, the present study focused on L2 persuasive writing, which tends to prioritize argumentation over the inclusion of abstract or less imageable content. Consequently, the influence of these psycholinguistic features may have been diminished in this study's writing evaluations.

2. Lexical Sophistication and L2 Speaking Scores

The second research question addressed how L2 lexical sophistication features predicted L2 speaking scores. Among the six indices of lexical sophistication examined in this study, five showed significant correlations with L2 persuasive speaking scores. Ultimately, only one index, word association, was retained in the final regression model. This finding suggests that higher-rated L2 persuasive speaking samples tended to contain more words with restricted contextual usage, consistent with previous L2 speaking research by Berger et al. (2017), which identified word association measures as significant predictors of lexical proficiency. The current study extends prior research by demonstrating that word association scores were the strongest and only significant predictor of EFL persuasive speaking scores, explaining approximately 26% of the variance. This implies that, in the context of L2 persuasive speaking, word associations may have been the most influential lexical sophistication feature. A possible explanation for this finding is that the use of words with restricted or specialized contextual usage may signal advanced lexical proficiency, as it suggests that the speaker employs more precise language to communicate ideas clearly. This precision can make the L2 speech appear more polished and suitable for formal discourse, such as persuasive tasks, which likely contributes to higher scores.

To illustrate, Tables 6 and 7 present excerpts from lower-rated and higher-rated speaking samples, respectively. The lower-rated sample contains more words with higher word association scores (e.g., *poor, money, buy*), whereas the higher-rated sample features more words with lower word association scores (e.g., *graduation, treat, attitude)*.

Of secondary interest are the lexical sophistication features that showed significant correlations with L2 speaking ratings but were not included in the final regression analysis. Although these metrics did not predict L2 speaking scores in the final model, they remain theoretically relevant. First, the two frequency measures, based on both academic and spoken corpora, exhibited significant negative correlations with L2 speaking scores, consistent with previous studies (e.g., Crossley et al., 2011b; Crossley & McNamara, 2013). Additionally, word concreteness and imageability scores also showed significant negative correlations with L2 speaking scores, aligning with prior research (Crossley et al., 2011b;

Hsieh & Wang, 2019).

TABLE 6
Lower-Rated Spoken Sample Excerpt

Excerpt	Content word	Word association score (USF)
[...] I must **work** on my **part time job.** Because I'm **very poor.** and my **parents** cannot **support** me. So I should I must **going** on this **job. Study.** But no **money.** no **study** because, **study** if I **want** to **study,** I **need** more **books** and some **affluence** [...]	work	196
	part	30
	time	119
	job	74
	very	2
	poor	51
	parents	21
	support	19
	going	6
	job	74
	study	60
	money	302
	study	60
	study	60
	want	33
	study	60
	need	32
	books	9
	affluence	n/a
	M	67.11

Notes. Excerpt from spoken sample *SD_KOR_06_PTJ_ROL_xx_012_B1_1*, whose score is 31.34. Content words are presented in bold.

TABLE 7
Higher-Rated Spoken Sample Excerpt

Excerpt	Content word	Word association score (USF)
[...] when it **comes** to **graduation** or before **graduating,** there is no **chance** for me to **have** a **real social experience.** When it **comes** to **part time job,** I **have** to **serve** my **bosses** and **customers.** Through that **case,** I can **learn** how we **make** a **service** [...]	comes	1
	graduation	2
	graduating	3
	chance	16
	have	17
	real	18
	social	17
	experience	8
	comes	1
	part	30
	time	119
	job	74
	have	17
	serve	3
	bosses	29
	customers	3
	case	23

	learn	39
	make	40
	service	24
	M	24.20

Notes. Excerpt from spoken sample *SD_KOR_06_PTJ_ROL_xx_013_B2_0*, whose score is 75.93. Content words are presented in bold.

While further research is needed to clarify the roles of frequency, concreteness, and imageability in explaining L2 speaking scores, existing studies (e.g., Crossley et al., 2011b; Crossley & McNamara, 2013) and this study suggest that the use of less frequent words may be perceived as an indicator of advanced lexical knowledge, making them more likely to contribute to higher scores. This could result in L2 spoken production with many low-frequency words receiving higher ratings. Similarly, in the context of L2 persuasive speech, the use of less concrete and less imageable words may signal higher cognitive and lexical sophistication, as well as more formal language use. In contrast, over-reliance on concrete, easily imageable words might make speech sound too simplistic or less appropriate for formal discourse, potentially leading to lower ratings.

Additionally, it should be noted that AoA scores were not correlated with L2 speaking scores, which is unexpected given that a previous L2 speaking study (Hsieh & Wang, 2019) reported a positive correlation. A possible explanation for the lack of correlation in this study may relate to the topic, which focused on college students working part-time jobs. This topic likely limited the range of words used, as it constrained speech to vocabulary related to college life and employment. These words may tend to be acquired at similar ages among native speakers, resulting in narrower variation in the AoA of the words used in the speech samples. Consequently, the AoA of words may not have played a significant role in influencing raters' evaluations of L2 persuasive speech in this context. However, further research is needed to confirm this hypothesis and explore the relationship between AoA and L2 speaking performance across different task types.

Taken together, it is important to highlight the similarities in the effects of lexical sophistication on L2 writing and speaking scores. First, word association scores emerged as the most significant predictor for both L2 persuasive writing and speaking scores potentially due to its ability to capture contextual specialization. Words with fewer associations may reflect advanced lexical knowledge, which is critical in persuasive tasks. This aligns with raters' expectations for clear and sophisticated language use in both modalities, contributing to higher scores. This finding suggests that word associations, which capture lexical connections and contextual usage, have a stronger influence on raters' evaluations of both modalities compared to other lexical sophistication measures, such as word frequency, concreteness, imageability, and AoA. That is, word association emerged as the strongest predictor of scores in both writing and speaking. To my

knowledge, the discovery that word associations are predictive of both L2 writing and speaking within the same genre (i.e., persuasion) is novel in this field.

Additionally, both frequency measures—based on academic and spoken corpora—showed negative correlations with L2 writing and speaking scores, indicating that frequency effects (i.e., more frequently used words being easier to process and, therefore, perceived as less sophisticated) persist across both modalities. This suggests that raters evaluate lexical sophistication in L2 persuasive discourse in a modality-insensitive manner, favoring the use of less frequent words regardless of whether the discourse is written or spoken.

Differences in the role of lexical sophistication across L2 writing and speaking also merit attention. Specifically, the psycholinguistic properties of words, concreteness, imageability, and AoA, exhibited modality-sensitive effects on performance scores. The significant correlations between word concreteness and imageability scores with L2 speaking, but not writing, highlight the negative influence of sensory experience (Paivio, 1971) on raters' evaluations of L2 persuasive speaking. This suggests that reliance on more concrete and easily imageable words may detract from the perceived sophistication of spoken discourse. In contrast, the significant correlation between AoA scores and L2 writing, but not speaking, indicates that cumulative exposure, reflected by AoA scores, plays a more prominent role in raters' evaluations of L2 persuasive writing. This implies that in writing, the use of words typically learned later in life is perceived as a marker of lexical sophistication, while it has less relevance in the context of spoken discourse.

VI. CONCLUSION

This study investigated the roles of lexical sophistication in both L2 writing and speaking within the same genre of persuasion, using identical indices. The findings revealed that word association scores were the most significant predictor for both L2 persuasive writing and speaking scores, emphasizing the importance of the word association dimension of lexical sophistication across modalities. Additionally, frequency scores were negatively correlated with both L2 writing and speaking scores, reinforcing the frequency effects in the evaluation of both modes. Furthermore, the results demonstrated that the psycholinguistic properties of words, such as word concreteness, imageability, and age of acquisition, operate in a modality-sensitive manner, influencing L2 writing and speaking scores differently. Overall, this study highlights both the similarities and differences in how lexical sophistication features predict performance across the two modalities.

The findings provide valuable pedagogical insights. They suggest that when teaching

lexical sophistication for persuasive writing and speaking, L2 instructors should focus on lexical items related to word associations and frequency. For instance, learners engaged in persuasive tasks can be encouraged to use words that appear in restricted, specialized contexts and words that are less frequently used. In addition, activities that encourage learners to use words with restricted contextual usage, such as context-specific writing and speaking tasks, can enhance lexical sophistication in both modalities. Additionally, L2 instructors should recognize the distinct lexical demands of writing and speaking: less concrete and less imageable words are likely to enhance persuasive speaking performance, while using words acquired at a later age may improve persuasive writing outcomes. From an assessment perspective, identifying modality-insensitive features, such as word association scores, highlights lexical characteristics consistently evaluated across writing and speaking. This finding supports integrated assessment designs with unified criteria, providing a holistic view of learners' lexical proficiency.

This study has several limitations that should be addressed in future research. First, unpaired L2 writing and speaking data were used, meaning that the samples were not produced by the same learners. To better compare the effects of lexical sophistication acro ss modalities, future studies could analyze writing and speaking samples from the same par ticipants. Second, only one task type (persuasion) was examined. Future research could ex plore the role of lexical sophistication in other types of tasks across both modalities. Third, this study's sample was limited to EFL learners from specific regions with beginner-to-intermediate proficiency levels, which may restrict the generalizability of the findings. Fut ure research should consider more diverse populations. Lastly, the scope of this study was confined to a limited set of lexical sophistication features, which might have restricted the extent to which these features accounted for the variance in the regression results. Future re search should consider incorporating a broader range of lexical features and other linguistic features, such as lexical diversity, phraseological sophistication, and syntactic complexity, to offer a more comprehensive understanding of the relationships between language use a nd L2 performance.

REFERENCES

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and de veloping useful language tests*. Oxford University Press.

Berger, C., Crossley, S. A., & Kyle, K. (2017). Using novel word context measures to pr edict human ratings of lexical proficiency. *Educational Technology & Society*, *20*, 201–212.

Coltheart, M. (1981). Disorders of reading and their implications for models of normal reading. *Visible Language*, *15*(3), 245−286.

Corver, N., & van Riemsdijk, H. (Eds.). (2001). *Semi-lexical categories*. Mouton de Gruyter.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, *34*(2), 213−238. https://doi.org/10.2307/3587951

Crossley, S., & McNamara, D. (2013). Applications of text analysis tools for spoken response grading. *Language Learning & Technology*, *17*, 171−192. http://doi.org/10125/44329

Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2011a). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing*, *28*(4), 561−580. https://doi.org/10.1177/0265532210378031

Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2011b). What is lexical proficiency? Some answers from computational models of speech data. *TESOL Quarterly*, *45*(1), 182−193.

Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The Academic Spoken Word List. *Language Learning*, *67*(4), 959−997. https://doi.org/10.1111/lang.12253

Davies, M. (2009). The 385+ million word Corpus of Contemporary American English (1990-2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*, *14*, 159−190.

De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M., & Storms, G. (2019). The "Sma ll World of Words" English word association norms for over 12,000 cue words. *B ehavior Research Methods*, *51*, 987−1006. https://doi.org/10.3758/s13428-018-1115-7

Eguchi, M., & Kyle, K. (2020). Continuing to explore the multidimensional nature of lexical sophistication: The case of oral proficiency interviews. *Modern Language Journal*, *104*(2), 381−400. https://doi.org/10.1111/modl.12637

Fitzpatrick, T., & Thwaites, P. (2020). Word association research and the L2 lexicon. *Language Teaching*, *53*(3), 237−274. https://doi.org/10.1017/S0261444820000105

Hsieh, C.-N., & Wang, Y. (2019). Speaking proficiency of young language students: A aiscourse-analytic study. *Language Testing*, *36*, 27−50. https://doi.org/10.1177/0265532217734240

Ishikawa, S. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. *Learner Corpus Studies in Asia and the World, 1,* 91−118. https://doi.10.24546/81006678

Ishikawa, S. (2019). The ICNALE spoken dialogue: A new dataset for the study of Asian learners' performance in L2 English interviews. *English Teaching*, *74*(4), 153−177. https://doi.org/10.15858/engtea.74.4.201912.153

Ishikawa, S. (2020). Aim of the ICNALE GRA project: Global collaboration to collect ratings of Asian learners' L2 English essays and speeches from an ELF perspective. *Learner Corpus Studies in Asia and the World*, *5*, 121−144.

Izura, C., Pérez, M. A., Agallou, E., Wright, V. C., Marín, J., Stadthagen-González, H., & Ellis, A. W. (2011). Age/order of acquisition effects and the cumulative learning of foreign words: A word training study. *Journal of Memory and Language*, *64*(1), 32−58. https://doi.org/10.1016/j.jml.2010.09.002

Kaushanskaya, M., & Rechtzigel, K. (2012). Concreteness effects in bilingual and monolingual word learning. *Psychonomic Bulletin & Review*, *19*(5), 935−941. https://doi.org/10.3758/s13423-012-0271-5

Kim, M., Crossley, S. A., & Kyle, K. (2018). Lexical sophistication as a multidimensional phenomenon: Relations to second language lexical proficiency, development, and writing quality. *The Modern Language Journal*, *102*, 120−141. https://doi.org/10.1111/modl.12447

Kiss, G. R. (1973). Grammatical word classes: A learning process and its simulation. *Psychology of Learning and Motivation*, *7*, 1−41.

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition r atings for 30,000 English words. *Behavior Research Methods*, *44*, 978−990. https://doi.org/10.3758/s13428-012-0210-4

Kyle, K., & Crossley, S. (2016). The relationship between lexical sophistication and independent and source-based writing. *Journal of Second Language Writing*, *34*, 12−24. https://doi.org/10.1016/j.jslw.2016.10.003

Kyle, K., Crossley, S., & Berger, C. (2018). The tool for the automatic analysis of lexical sophistication (TAALES): version 2.0. *Behavior Research Methods*, *50*(3), 1030−1046. https://doi.org/10.3758/s13428-017-0924-4

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

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *Modern Language Journal*, *96*(2), 190−208. https://doi.org/10.1111/j.1154 0-4781.2011.01232_1.x

McDonald, S. A., & Shillcock, R. C. (2003). Eye movements reveal the on-line computation of lexical probabilities during reading. *Psychological Science*, *14*(6), 648−652. https://doi.org/10.1046/j.0956-7976.2003.psci_1480.x

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge University Pr ess.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computer*, *36*(3), 402−407. https://doi.org/10.3758/BF03195588

Norris, J. M., & Manchón, R. M. (2012). Investigating L2 writing development from multiple perspectives: Issues in theory and research. In Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 219−245). Gruyter.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology*, *45*(3), 255−287.

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

R Core Team. (2018). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org

Salsbury, T., Crossley, S. A., & McNamara, D. S. (2011). Psycholinguistic word information in second language oral discourse. *Second Language Research*, *27*(3), 343−360. https://doi.org/10.1177/0267658310395851

Uccelli, P., Dobbs, C. L., & Scott, J. (2013). Mastering academic language: Organization and stance in the persuasive writing of high school students. *Written Communication*, *30*, 36−62. https://doi.org/10.1177/0741088312469013

Uchihara, T., & Clenton, J. (2020). Investigating the role of vocabulary size in second language speaking ability. *Language Teaching Research*, *24*(4), 540−556. https://doi.org/10.1177/1362168818799371

Witte, S. P., & Faigley, L. (1981). Coherence cohesion, and writing quality. *College Composition and Communication, 32*(2), 189−204.

Zhang, X. (2022). The relationship between lexical use and L2 writing quality: A case of two genres. *International Journal of Applied Linguistics*, *32*(3), 371−396. https://doi.org/10.1111/ijal.12420

Examples in: English

Applicable Languages: English

Applicable Levels: Tertiary

Minkyung Kim (professor)

Department of English Education

Korea National University of Education

28173, Taeseongtabyeon-ro 250, Heungdeok-gu, Cheongju-si, Chungcheongbuk-do, Korea

Email: minkyungkim@knue.ac.kr

Received: November 1, 2024
Reviewed: December 6, 2024
Revised version received: December 23, 2024