



P-ISSN 2355-2794  
E-ISSN 2461-0275

## Key Predictors of Proficiency in L2 Writing: Syntactic, Lexical Complexity, and Accuracy

Shinjae Park\*

College of General Education, Kookmin University, Seoul 02707, KOREA

### Abstract

*While many studies have examined individual aspects of complexity, accuracy, and fluency (CAF) as indicators of language performance, only few have explored how syntactic complexity, lexical complexity, and accuracy function together to measure writing proficiency. This study aims to fill that gap by identifying reliable methods to measure 14 indices of syntactic complexity, five of lexical complexity, and nine of accuracy. It additionally investigates how these indices contribute to holistic writing assessments. It further examines which specific indices best predict writing proficiency and determine writing quality. A total of 138 essays by Korean learners of English were analyzed using the L2 Syntactic Complexity Analyzer and Text Inspector. Multiple regression analyses were conducted to interpret the data. The results show that proficiency is best predicted by the number of coordinate phrases per clause and dependent clauses per T-unit (syntactic complexity), sophistication (lexical complexity), and article usage (accuracy). Lexical sophistication emerged as the strongest predictor of proficiency. The paper also discusses the Common European Framework of Reference for Languages (CEFR) rubrics and offers insights into selecting and categorizing writing proficiency measures. Implications highlight the need to prioritize lexical sophistication and syntactic complexity in L2 writing assessment and pedagogy.*

**Keywords:** Accuracy, assessment, lexical complexity, sophistication, syntactic complexity.

## 1. INTRODUCTION

The importance of Complexity, Accuracy and Fluency (CAF) in second language (L2) speaking and writing has elicited interest from L2 linguists during the last decade (e.g. Barrot & Agdeppa, 2021; Khushik & Huhta, 2020; Tabari et al., 2023). From a theoretical perspective,

---

\* Corresponding author, email: [muhando@kookmin.ac.kr](mailto:muhando@kookmin.ac.kr)

**Citation in APA style:** Park, S. (2025). Key predictors of proficiency in L2 writing: Syntactic, lexical complexity, and accuracy. *Studies in English Language and Education*, 12(3), 1119-1132.

Received November 4, 2024; Revised January 20, 2025; Accepted August 26, 2025; Published Online September 30, 2025

<https://doi.org/10.24815/siele.v12i3.42039>

Copyright © 2025 by Authors, published by *Studies in English Language and Education*. This is an open-access article distributed under the Creative Commons Attribution 4.0 International License (<https://creativecommons.org/licenses/by/4.0/>)

CAF measures have been considered key constructs in research on L2 speaking and writing and have been adopted to find valid and reliable indices of L2 learners' language development and global proficiency (Skehan, 2009). From a practical standpoint, such measures can be used as frameworks to better understand L2 performance in the context of English as a Foreign Language (EFL) (Kuiken et al., 2010) and implement appropriate pedagogical interventions. Under a CAF framework, therefore, learners are expected to produce more complex, accurate and fluent spoken and written output. If this expectation challenges learners, calibrated pedagogical interventions can be initiated.

Complexity has been examined extensively as an index of language performance, writing quality and writing proficiency in L2 (Barrot & Agdeppa, 2021; Park, 2022a). For example, some studies have distinguished writing quality based on complexity measures and revealed their relationship with proficiency (Casal & Lee, 2019; Martínez, 2018; Park, 2022a). Although these studies have dealt with one or more CAF measures as indices of proficiency, many have focused primarily on syntactic complexity measures. The insights that they derived point to the need for a more comprehensive scrutiny of the three dimensions of L2 writing performance in relation to language proficiency.

Accuracy refers to the degree of deviation from a certain norm, with such deviations generally characterised as errors (Michel, 2017, pp. 8-11). However, defining the scope of errors as a measure of accuracy remains a challenging issue because standards can vary among researchers or evaluators. This problem raises the question of whether to set the linguistic level demonstrated by an ideal native speaker of a target language as the norm or to expand criteria to include non-native usage patterns that are acceptable in certain social contexts or regional communities (Housen & Kuiken, 2009, p. 462). As can be seen, measuring accuracy using valid and reliable methods is a crucial requirement.

In contrast to complexity and accuracy, which may pertain to oral and written L2 performance, fluency is first and foremost a measure of spoken language (Michel, 2017). Tavakoli and Skehan (2005) formulated a regularly cited speech-oriented definition, which indicates fluency as consisting of three subdimensions: (1) speed or rate (e.g., the number of words uttered per minute), (2) silence or breakdown (e.g., the number, location and duration of [filled] pauses), and (3) repair (e.g., false starts, repetitions and self-corrections). A limited number of studies have suggested the potential for extending this work into the writing domain. These studies employed keystroke logging software to record online writing features, such as the number of characters typed between pauses and the ratio of characters produced during writing to that in the final text (Michel, 2017). While fluency is a multifaceted construct with several subcomponents that make it a useful metric for assessing L2 production quality, it remains a challenging dimension to measure and conceptualize in L2 writing.

Skehan (2009) argued that one of the major omissions in performance measures is that of lexis, indicating the need for CAF competencies to be supplemented by the characteristics of lexical use (Park, 2023b). This is important because empirical evidence suggests that the latter is a separate aspect of overall performance, and it shows that lexical access and retrieval figure prominently in all models of language production (Park, 2023b). Despite these requirements, however, few studies have examined integrated measures as indices of writing proficiency. This gap has limited the understanding of L2 writing proficiency in relation to necessary integrated measures, such as those revolving around CAF and lexis. To fill this gap, the three dimensions useful for assessing L2 writing—syntactic complexity, lexical complexity, and accuracy—were thoroughly examined. Specifically, the three dimensions and the corresponding measures suitable for this analysis were selected, calculated using reliable methods, and subjected to appropriate statistical analysis. Additionally, the relationship between these measures and writing proficiency, as categorized by the Common European Framework of Reference for Languages (CEFR) standards, was analysed. This examination was guided by the following research questions:

1. What variables of syntactic complexity affect L2 proficiency levels?
2. What variables of lexical complexity affect L2 proficiency levels?
3. What variables of accuracy affect L2 proficiency levels?

4. Which of the variables identified from the analyses have the greatest impact on L2 proficiency levels?

## 2. LITERATURE REVIEW

Researchers in applied linguistics and instructed L2 acquisition have long been interested in measuring L2 performance. For example, [Brumfit \(1984\)](#) highlighted accuracy and fluency as two important aspects of language use, and [Skehan \(1998\)](#) later added complexity, introducing the CAF triad as the three fundamental dimensions characterising L2 performance. These three dimensions have since proven to be essential and useful measures for assessing various aspects of L2 proficiency ([Skehan, 2009](#)).

The three dimensions were defined provisionally by [Ellis \(2003\)](#) as follows. Complexity refers to the extent to which the language produced by learners is elaborate and varied. Accuracy is defined as the degree to which the language produced by a learner conforms to target language norms. Fluency pertains to the extent to which the language produced by a learner shows evidence of pausing, hesitation or reformulation. Meanwhile, fluency will only be briefly mentioned in the current section, as previously stated, this dimension is primarily a measure of spoken language and is challenging to conceptualise and measure in L2 writing ([Michel, 2017](#)). Additionally, writing fluency is less reasonable to measure than oral fluency, as writing inherently allows for planning, monitoring and editing ([Michel, 2017](#)).

Given the importance of the three dimensions described above, the literature on syntactic complexity, lexical complexity, and accuracy in L2 writing was explored. These are key areas that have been extensively studied in relation to L2 proficiency. Accordingly, this study examines the extent to which these areas account for differences in L2 writing proficiency.

### 2.1 Syntactic Complexity

Syntactic complexity has been a focal point in L2 research, as it provides insights into learners' ability to use a wide range of grammatical forms and structures, reflecting their language sophistication ([Barrot & Agdeppa, 2021](#); [Skehan, 2009](#)). Such complexity was elaborated by [Lu \(2010\)](#) as encompassing 14 indices, including the mean length of a sentence (MLS), the mean length of a T-unit (MLT), the mean length of a clause (MLC), the number of clauses per T-unit (CT) and the complex T-unit ratio (CTT). [Martínez \(2018\)](#) compared the syntactic complexity of argumentative essays written by lower-intermediate and intermediate secondary education students in Spain. She used measures such as MLS, the simple sentence ratio, the compound sentence ratio, the complex sentence ratio, the compound-complex sentence ratio, the number of clauses per sentence (CS), the number of dependent clauses per clause (DCC) and the number of coordinate phrases per clause (CPC). Her findings revealed that intermediate students score higher on most measures, although the limited scope of the study may not have captured the full range of syntactic complexity. [Kuiken and Vedder \(2019\)](#) investigated the syntactic complexity of argumentative texts using both overall ratio measures and fine-grained measures of subordination, coordination and phrasal complexity. Their study, which involved learners with Dutch, Italian and Spanish as their first languages (L1), indicated that higher-proficiency learners produce more complex syntactic structures, albeit this trend varies across students of different L1 backgrounds. Similar tendencies were found by [Khushik and Huhta \(2020\)](#), who examined the English argumentative essays produced by Pakistani and Finnish teenagers and found that proficiency levels influence subordination, phrasal density and the length of production units. In contrast, [Park \(2022b\)](#) identified complex nominals per T-unit (CNT) as the best predictor of proficiency among Korean L2 learners. Despite variations across studies, there is general agreement that more proficient L2 learners tend to produce longer and more complex syntactic structures ([Casal & Lee, 2019](#); [Martínez, 2018](#)).

## 2.2 Lexical Complexity

Lexical complexity, which encompasses diversity, sophistication and density, has received less attention than syntactic complexity despite it being another critical dimension of language use (Read, 2007). This assertion is supported by Skehan (2009), who emphasised the need to incorporate lexical measures into L2 research, given their importance in understanding task performance. Lexical diversity is typically measured using D, a measure that calculates the extent to which a speaker avoids recycling the same set of words (Malvern & Richards, 2002). A low D suggests considerable reliance on a limited vocabulary set, which implies minimal lexical diversity (Skehan, 2009). Lexical sophistication refers to the use of less frequent, more advanced words, as measured by tools such as Laufer and Nation's (1999) Lexical Frequency Profile. This tool is used to assess the frequency with which words are used in a given text, allowing researchers to gauge the extent to which a writer relies on more frequent versus less frequent terms. Finally, lexical density is employed to measure the ratio of lexical (content) words to the total number of words in a text (Ure, 1971). Lexical words, such as nouns, adjectives, verbs and adverbs, carry the core meanings of a text, while function words (e.g. articles, prepositions) serve grammatical purposes without adding significant meaning (Halliday, 1994). Notwithstanding the importance of lexical complexity in L2 writing, research on this dimension has been relatively sparse. Incorporating lexical measures is essential for a comprehensive understanding of L2 proficiency, as it offers insights into the richness and depth of a learner's vocabulary.

## 2.3 Accuracy

Accuracy in L2 writing refers to the extent to which a learner's language use aligns with the norms of a target language, encompassing both form and structure (Michel, 2017; Skehan, 2009). It was defined by Ellis (2003) as a key dimension of language performance, highlighting its importance in evaluating L2 proficiency. Accuracy has been measured using various developed indices, ranging from holistic measures, such as error-free clauses and the number of errors per 100 words, to specific indices that target particular aspects of language use, such as errors in noun–adjective–gender agreement (Michel, 2017). The selection and categorisation of errors allow for comparisons across studies but require careful interpretation and categorisation based on a study's scope (Park, 2023a). To address these challenges, the Center for English Corpus Linguistics at Université Catholique de Louvain (Dagneaux et al., 1996) proposed an alternative error categorisation scheme, which comprises up to four levels of classification depending on error category. This scheme was adopted by S. Yoon et al. (2022) through the identification of 40 distinct error categories, covering aspects such as form, grammar, lexico-grammar, lexis, word use, register and style. Although much of the existing research has focused on syntactic complexity, accuracy remains a crucial component of L2 proficiency, as it reflects a learner's ability to produce language that conforms to target language norms. However, more explorations are needed to fully understand how accuracy interacts with other dimensions of L2 performance.

In summary, syntactic and lexical complexity and accuracy are all critical for understanding L2 proficiency. While there has been substantial research on syntactic complexity, less attention has been given to lexical complexity and accuracy in relation to predicting proficiency levels. This review highlighted the necessity of more comprehensive investigations that incorporate all three dimensions to provide a more complete picture of L2 writing proficiency. This gap was addressed in the current study through an examination of the effects of syntactic complexity, lexical complexity and accuracy on L2 proficiency levels.

## 3. METHOD

This section explains how the study was conducted, starting with a description of the corpus data, including details about the participants and their writing tasks. It also introduces the tools and methods used to measure key linguistic features, such as syntactic complexity, lexical

richness, and accuracy. Together, these elements form the foundation for analyzing the factors that influence L2 writing proficiency.

### 3.1 Samples and Data Collection

The data analysed in this study were derived from 138 essays included in the Multi-language Learner Corpus (MULC) (S. Yoon et al., 2022). The participants who produced part of the corpus were recruited through an open call issued at a university in South Korea between 2018 and 2022. The majors pursued by the participants varied, and they had an average age of 20.9 years. Among them, 47.1% were male, and 52.9% were female (Table 1). They were asked to choose one of four topics on which to write an essay, which was later electronically stored in the MULC database. The topics of the essays are listed below:

1. Should everyone get married?
2. Is it essential to wear uniforms in middle and high schools?
3. Should elementary, middle and high school students be allowed to carry mobile phones in class?
4. Should any college student join a club?

Each of the participant's writing proficiency was assessed by a native English speaker who was specifically trained for this purpose, with the rubrics of the CEFR used as bases (Appendix). The evaluator holds a PhD in linguistics and is currently a professor at a university in Korea. The CEFR has been recognised as a standard for L2 language progression throughout Europe since 2001 and has gradually expanded in use worldwide (Glover, 2011). It is divided into six levels of proficiency: A1, A2, B1, B2, C1 and C2 (with A1 being the lowest and C2 being the highest). However, since the data used in this study were generated mostly by first- and second-year university students who may not have been accustomed to writing in English, none of the writings reached the C1 or C2 level. Therefore, their writing levels were categorised into four groups: A1, A2, B1 and B2 (Table 1). More specifically, the data were grouped as follows: A1 ( $n = 12$ , or 8.7%), A2 ( $n = 77$ , 55.8%), B1 ( $n = 40$ , 29.0%) and B2 ( $n = 9$ , 6.5%).

**Table 1.** Description of participants.

Gender		Proficiency			
Male	Female	A1	A2	B1	B2
65	73	12	77	40	9
47.1%	52.9%	8.7%	55.8%	29.0%	6.5%

Table 2 shows the descriptive statistics of the data. The total number of tokens was 32,128, with an average of 232.81 words per essay ( $SD = 101.467$ ) across the 138 pieces. The total number of sentences was 2,320, with an average of 16.81 sentences per essay ( $SD = 6.833$ ). Some of the essays had as many as 597 words and 41 sentences.

**Table 2.** Description of corpus.

	Tokens	Mean	SD	Min.	Max.
Word	32,128	232.81	101.467	55	597
Sentence	2,320	16.81	6.833	3	41

### 3.2 Instruments and Data Analysis

This section describes the tools and methods used to analyse the linguistic features of the corpus data, focusing on syntactic complexity, lexical complexity, and accuracy. Each dimension was measured using established computational tools, complemented by manual verification to ensure accuracy. The following subsections detail the specific measures and procedures employed for each dimension.



### 3.2.1 Syntactic complexity

The L2 syntactic complexity analyser (L2SCA) was used to measure the syntactic complexity indices adopted in this study. More details on each index can be found in [Lu's \(2010\)](#) work. The five major categories and the number of indices that they subsumed were as follows: the length of production units ( $n = 3$  indices), overall sentence complexity ( $n = 1$  index), the number of subordination clauses ( $n = 4$  indices), the number of coordination clauses ( $n = 3$  indices) and phrasal sophistication ( $n = 3$  indices). The specific linguistic features and indices gauged using the L2SCA are summarised in Table 3, which also presents the descriptive statistics of the categories. A comparative analysis based solely on the mean values of each measure is not meaningful, and only relative comparisons through the analysis of their relationship with proficiency are important. This issue is further addressed in the results section.

**Table 3.** Syntactic complexity measures ([Lu, 2010](#)).

	Categories	<i>M(SD)</i>
Length of production unit	Mean length of clause (MLC: words/clause)	7.863(1.234)
	Mean length of sentence (MLS: words/sentence)	14.110(3.927)
	Mean length of T-unit (MLT: words/T-unit)	12.861(3.145)
Sentence complexity	Clauses per sentence (CS: clauses/sentence)	1.807(0.487)
Subordination	Clauses per T-unit (CT: clauses/T-unit)	1.647(0.366)
	Complex T-unit per T-unit (CTT: complex T-units/T-unit)	0.459(0.176)
	Dependent clauses per clause (DCC: dependent clauses/clause)	0.360(0.101)
	Dependent clauses per T-unit (DCT: dependent clauses/T-unit)	0.621(0.288)
Coordination	Coordinate phrases per clause (CPC: coordinate phrases/clause)	0.164(0.094)
	Coordinate phrases per T-unit (CPT: coordinate phrases/T-unit)	0.264(0.152)
	T-units per sentence (TS: T-units/sentence)	1.096(0.133)
Particular structures	Complex nominals per clause (CNC: complex nominals/clause)	0.823(0.239)
	Complex nominals per T-unit (CNT: complex nominals/T-unit)	1.351(0.469)
	Verb phrases per T-unit (VPT: verb phrases/T-unit)	2.242(0.538)

### 3.2.2 Lexical Complexity

In L2 writing research, lexical complexity is a multidimensional construct consisting of various subconstructs that can be measured using specific metrics. A popular approach is to conceptualise lexical complexity as encompassing three main subconstructs: lexical sophistication/rarity, lexical diversity/variability and lexical density ([Bui, 2021](#); [Tabari et al., 2023](#)). This multidimensional framework was adopted in the present work to analyse lexical complexity (Table 4).

Lexical sophistication was estimated using the Academic Word List and Beyond-2000 scores based on the British National Corpus (BNC) ([Park, 2023b](#)). To enquire into lexical sophistication, researchers employ frequency-based metrics to compare the lexical information in L2 writing samples against frequency data extracted from language corpora ([Johnson, 2017](#)). The use of low-frequency lexis is considered indicative of L2 writing development ([Laufer & Nation, 1999](#)). Therefore, sophisticated lexical usage reflects a low overall lexical frequency ([Crossley & McNamara, 2012](#); [H. J. Yoon & Polio, 2017](#)).

Lexical diversity refers to the variety of vocabulary used in a text—the more varied the vocabulary, the higher the lexical diversity. For a text to be highly lexically diverse, a speaker or writer must use many different words with minimal repetition ([Johansson, 2008](#)). Lexical diversity indicates the range of a learner's vocabulary as demonstrated in their language use ([Lu, 2012](#); [McCarthy & Jarvis, 2007](#)). To measure lexical diversity, this study used metrics such as Textual Lexical Diversity (MTLD) and Vocab-D ([McCarthy & Jarvis, 2007](#)). These metrics were calculated using the Text Inspector, a tool prototyped by Stephen Bax and further developed by the software team at Versantus ([Owen et al., 2021](#)).

Lexical density is defined as the proportion of content words in a written text (Bui, 2021) and reflects how densely lexical items are used within grammatical structures (Halliday & Martin, 1993). It was determined in the present research through a series of manual operations. Specifically, before an automated analysis, several processes were added. First, a cleaned text file was tagged for parts of speech (POS) using the Stanford tagger, which assigns each token in a language sample a label that indicates the POS category to which it belongs (e.g. noun, verb, adjective, adverb). This tagging was necessary for the calculation of lexical density, which is crucial in distinguishing between content and function words. To reduce errors that might occur due to the automatic tagging process, a manual review by linguists, including the present author, was conducted. This approach was expected to considerably increase the accuracy of measurements compared with those performed in studies that rely solely on results obtained using a single automatic tagger. AntConc 3.5.9 was used to advance the calculation of content words using the POS-tagged sample. Table 4 presents the measures, definitions and descriptive statistics of lexical complexity. The analysis of its relationship with proficiency is presented in the results section.

**Table 4.** Lexical complexity measures.

	Measures	Definitions	<i>M(SD)</i>
Diversity	VOCD	A mathematical transformation of the standard type-token ratio (TTR), which reduces the intervening impacts of text length and indicates the degree of word repetition in a text	75.120(19.895)
	MTLD	The average length of sequential word strings in a text which maintain a given TTR value	65.519(17.769)
Density	Verbal E./S.	Verbal word ratio	2.149(1.210)
	Noun E./S.	Noun word ratio	2.069(1.339)
Sophistication	Beyond BNC (2K)	The Beyond-2000 values calculated by subtracting K1 and K2 ratios from 100%	0.287(0.103)

### 3.2.3 Accuracy

In this study, the error category scheme proposed by the Center for English Corpus Linguistics at Université Catholique de Louvain (Dagneaux et al., 1996) was used to divide errors into nine categories: morphology, articles, nouns, pronouns, verbs, spelling, lexis, style and word omissions and redundancies. These categories were selected because they represent the top nine in which the learners committed the most errors, as determined based on the frequency analysis of the corpus used in this study (Park, 2023a). Other categories either had very few errors or were part of the same grammatical category but in different subcategories of errors. These were combined into a single category. For example, verb tense and verb voice were merged under the verb category. An excessively small number of tokens per category may lead to unreliable results. Thus, I divided the categories according to the frequency of errors in writing.

For error annotation, the Universidad Autónoma de Madrid Corpus Tool 3.0 (UAM) (<http://www.wagsoft.com/CorpusTool/>) was used, which is designed specifically to support typical user workflows, with everything needed to perform annotation tasks included (Wang & Yu, 2022; S. Yoon et al., 2022). After automated annotation using the UAM, four linguists, including the present author, manually double-checked the categorisation of grammatical errors and held regular weekly meetings to reach a consensus on a common grammatical standard. Following the annotation, a native English speaker on the corpus team, who is also a qualified linguist, tagged the proficiency levels reflected in the essays according to the CEFR scheme (S. Yoon et al., 2022).

The descriptions of each error category were excerpted from Dagneaux et al. (1996) and Granger et al. (2022), as presented in Table 5. The analysis of the relationship between accuracy and proficiency is provided in the results section. This approach ensured reliable categorisation and interpretation of errors prior to statistical analysis.

**Table 5.** Accuracy measures (Dagneaux et al., 1996; Granger et al., 2022).

Categories	Descriptions	M(SD)
Morphology	1. Inflectional errors: Misuse of grammatical morphemes (e.g. plurals, possessives, verb forms or adjective degrees) 2. Derivational errors: Incorrect affix added to an existing word	1.18(1.581)
Articles	Issues with definite, indefinite or zero articles	5.99(4.405)
Nouns	1. Misuse of the Saxon genitive 2. Addition/omission of plural morphemes	2.80(2.334)
Pronouns	Errors with all pronoun categories: Demonstrative, interrogative, personal, possessive, quantifying, reciprocal, reflexive, relative, indefinite	1.11(1.527)
Verbs	1. Verb number: All errors of concord between a subject and its verb 2. Verb morphology: Erroneous uses of existing verb forms (e.g. a simple past form instead of a past participle form, an infinitive instead of a past participle) 3. Non-finite/finite verb forms 4. Verb voice: Confusion of the passive and active voice 5. Verb tense: Misuse of verb tense or aspect	3.19(2.582)
Spellings	1. Misuse or omission of capital letters 2. Non-standard word coinages 3. Incorrect use of borrowings 4. Confusion between homophones 5. Doubling of consonants/vowels 6. Variations in American spelling 7. Misuse/omission of hyphens or spaces in compound words	2.60(2.743)
Lexis	Issues with errors involving the semantic (conceptual or collocational) properties of words or phrases	5.97(4.178)
Style	Long, foreign-sounding or clumsy text chunks that a native speaker would want to reformulate	4.64(3.308)
Word	1. Word redundant: Unnecessary repetitions of words 2. Word missing: Omission of words 3. Word order: Problems of word order	4.36(2.952)

#### 4. RESULTS

This study examined the extent to which measures in various domains can predict performance in L2 writing. The goal was to identify which factors among syntactic complexity, lexical sophistication, and accuracy have the most significant impact on L2 writing proficiency. To achieve this aim, multiple regression analyses were conducted, with proficiency as the dependent variable and these linguistic domains as the independent variables. Table 6 summarizes the results, showing the contributions of these measures, with certain factors reaching levels of statistical significance.

**Table 6.** Results of multiple regression Analysis (backward method).

Models		B	SE	$\beta$	<i>t</i>	<i>p</i>	<i>R</i> <sup>2</sup>	Adjusted <i>R</i> <sup>2</sup>	Durbin–Watson
Accuracy	(Constant)	2.129	0.102		20.838	<0.001			
	Article	0.27	0.014	0.165	1.957	0.052	0.027	0.020	2.292
Lexis	(Constant)	3.409	0.150		22.702	<0.001			
	Sophistication	-3.898	0.492	-0.562	-7.919	<0.001	0.316	0.311	2.174
Complexity	(Constant)	1.502	0.182		8.235	<0.001			
	CPC	0.814	0.202	0.327	4.027	<0.001	0.135	0.122	2.110
	DCT	1.727	0.616	0.227	2.804	0.006			

Note. 1) Criterion for removing independent variables: Probability of  $F \geq 0.100$

2) Durbin–Watson statistic: Ranges from 0 to 4, with values closer to 2 indicating more independent values



#### 4.1 Syntactic Complexity

With regard to syntactic complexity, when CPC and DCT served as independent variables, the model exhibited statistical significance ( $p < 0.05$ ) and explained 13.5% of the variances in the L2 essays (adjusted  $R^2 = 0.135$ ). These results demonstrate that structural features such as coordination and subordination are predictive of proficiency. This finding suggests that the ability to construct more complex sentence structures is a meaningful indicator of learners' writing development.

#### 4.2 Lexical Complexity

Regarding lexical complexity, the model with sophistication (use of advanced or less frequent words) as an independent variable was statistically significant ( $p < 0.001$ ) and explained 31.6% of the variances in the L2 writing pieces (adjusted  $R^2 = 0.316$ ). This indicates that lexical sophistication is the strongest predictor among the three domains. This result underscores the importance of advanced vocabulary uses in distinguishing higher-proficiency writers from their peers.

#### 4.3 Accuracy

In terms of accuracy, however, with article usage as the independent variable, the model was statistically non-significant ( $p = 0.052$ ), but it accounted for 2.7% of the variances in proficiency (adjusted  $R^2 = 0.027$ ). While accuracy measures, such as article usage, showed limited predictive power in this study, they remain relevant for understanding language conformity to norms, as discussed in prior research (Ellis, 2003; Michel, 2017). These findings indicate that while accuracy alone explains only a small portion of proficiency differences, it still contributes useful information when considered with other dimensions.

#### 4.4 Comparative Impact

To assess the impact of the most significant independent variables subsumed under the three major measures for proficiency, their standardised beta coefficients from the multiple regression analyses were compared. Article use under accuracy had a standardised beta coefficient of 0.165, suggesting a modest effect on proficiency. Sophistication under lexis had a coefficient of -0.562, indicating a substantial impact on proficiency, while CPC and DCT under complexity had a coefficient of 0.327, showing a moderate effect. The largest absolute beta value was that of sophistication, suggesting that it significantly influences proficiency.

Sophistication, represented by the use of Beyond-2000 words, is reflected by the occurrence of less frequently used words outside the 2000 most commonly occurring terms in the BNC. Such sophistication may lead to more complex writing, potentially impacting proficiency ratings. The relationship between lexical sophistication and proficiency is nuanced, suggesting that while sophisticated lexical use indicates advanced language skills, it also complicates writing and affects overall proficiency assessment.

### 5. DISCUSSION

This study explored the extent to which measures of syntactic complexity, lexical complexity, and accuracy can predict L2 writing proficiency. The multiple regression analyses generated key findings that addressed the research questions. The following subsections discuss these results in relation to each dimension and their implications for understanding L2 writing proficiency.

### **5.1 Syntactic Complexity in Relation to L2 Proficiency Levels**

The results revealed that among the various measures of syntactic complexity, the combination of CPC and DC was a significant predictor of L2 writing proficiency. The model incorporating these variables was statistically significant and explained 13.5% of the variances in proficiency scores. This finding suggests that complex sentence structures, as evidenced by the presence of numerous clauses per sentence and dependent clauses, contributes positively to perceived proficiency in L2 writing. This aligns with previous research indicating that the ability to construct syntactically complex sentences is often associated with high proficiency levels (Casal & Lee, 2019; Martínez, 2018).

### **5.2 Lexical Complexity in Relation to L2 Proficiency Levels**

Lexical sophistication emerged as a particularly strong predictor of L2 writing proficiency, with the sophistication measure alone accounting for 31.6% of the variances in proficiency. The use of advanced and low-frequency vocabulary was strongly associated with higher proficiency, indicating that lexical choices serve as a key marker of advanced language skills. Writers who utilise sophisticated vocabulary demonstrate not only a broader lexical repertoire but also an ability to adapt their language to contextually appropriate expressions. These findings are consistent with prior research, such as H. J. Yoon and Polio (2017), which underscores the importance of lexical richness in differentiating proficiency levels.

### **5.3 Accuracy in Relation to L2 Proficiency Levels**

In contrast to syntactic and lexical complexity, accuracy—specifically the use of articles—did not emerge as a statistically significant predictor of L2 proficiency. Its influence was limited but close enough to the threshold to warrant consideration. These results suggest that while accuracy in article usage might play a role in proficiency, its impact is less pronounced than that of syntactic and lexical factors. They are consistent with the idea that although accuracy is important, it may not be as critical in determining overall proficiency levels as the ability to produce complex and sophisticated language (Housen & Kuiken, 2009).

### **5.4 Key Determinants of L2 Proficiency Levels**

Among the variables analysed, lexical sophistication stood out as the most impactful predictor of L2 proficiency. This finding not only underscores the importance of lexical choices but also highlights that among the various measures of complexity and accuracy, lexical sophistication is the key determinant of writing proficiency. The prominence of this variable suggests that in the context of L2 writing, enhancing vocabulary sophistication should be a primary focus for improving overall proficiency. It also implies that teaching strategies and assessment practices should prioritise the development of advanced vocabulary to elevate writing quality.

Syntactic complexity, as represented by CPC and DCT, had a significant impact on proficiency, though to a lesser extent, while the influence of accuracy, particularly article usage, was even less pronounced. This reinforces the idea that while grammatical complexity and accuracy are important, the richness and variety of vocabulary play a more decisive role in determining proficiency. These findings echo the observations of Crossley and McNamara (2012), who noted that proficient writers consistently demonstrate advanced vocabulary and diverse syntactic structures. Consequently, L2 writing instruction should focus on both grammatical complexity and the enhancement of learners' ability to use advanced vocabulary.

## 6. CONCLUSION

Overall, the findings contribute to our understanding of how different aspects of linguistic complexity and accuracy influence L2 writing proficiency. The strong predictive power of lexical sophistication suggests that teaching practices should prioritise enhancing learners' vocabulary, particularly by encouraging the use of more sophisticated lexical items. While syntactic complexity also plays a significant role, it is slightly less influential than lexical sophistication but still underscores the importance of developing students' ability to construct complex sentence structures. However, the relatively minor impact of accuracy measures, such as article usage, implies that while accuracy should not be neglected, greater emphasis should be placed on fostering syntactic complexity, particularly lexical sophistication, to improve overall proficiency.

Moreover, this study underscores the importance of integrating multiple measures—syntactic complexity, lexical complexity, and accuracy—into the assessment of L2 writing proficiency. Through the use of tools such as the L2SCA and Text Inspector, 138 essays from Korean learners of English were analysed with precision, employing multiple regression analyses to determine the most significant predictors of writing proficiency. The successful application of automated text analysers in this study highlights their potential as reliable tools for evaluating writing proficiency in both research and educational contexts. These tools can provide objective, scalable assessments that complement traditional evaluation methods.

This study is limited by its focus on Korean learners of English and a specific set of linguistic features, which may not fully capture the range of factors influencing L2 proficiency across diverse contexts. Finally, the study points to several directions for future research. Expanding the range of linguistic features analysed and exploring different learner populations and writing contexts will help validate and refine the current findings. Additionally, further exploration of how automated tools can be integrated into classroom assessment practices can enhance the effectiveness and efficiency of L2 writing instruction.

## ACKNOWLEDGMENTS

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5A8054734).

## REFERENCES

- Barrot, J. S., & Agdeppa, J. Y. (2021). Complexity, accuracy, and fluency as indices of college-level L2 writers' proficiency. *Assessing Writing*, 47, Article 100510. <https://doi.org/10.1016/j.asw.2020.100510>
- Brumfit, C. (1984). *Communicative methodology in language teaching: The roles of fluency and accuracy*. Cambridge University Press.
- Bui, G. (2021). Influence of learners' prior knowledge, L2 proficiency and pre-task planning on L2 lexical complexity. *International Review of Applied Linguistics in Language Teaching*, 59(4), 543-567. <https://doi.org/10.1515/iral-2018-0244>
- Casal, J. E., & Lee, J. J. (2019). Syntactic complexity and writing quality in assessed first-year L2 writing. *Journal of Second Language Writing*, 44, 51-62. <https://doi.org/10.1016/j.jslw.2019.03.005>
- Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: The roles of cohesion and linguistic sophistication. *Journal of Research in Reading*, 35(2), 115-135. <https://doi.org/10.1111/j.1467-9817.2010.01449.x>
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.

- Dagneaux, E., Denness, S., Granger, S., Meunier, F., Neff, J., & Thewissen, J. (1996). *Error Tagging Manual Version 1.1*. Centre for English Corpus Linguistics, Université Catholique de Louvain.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Glover, P. (2011). Using CEFR level descriptors to raise university students' awareness of their speaking skills. *Language Awareness*, 20(2), 121-133. <https://doi.org/10.1080/09658416.2011.555556>
- Granger, S., Swallow, H., & Thewissen, J. (2022). *The Louvain Error Tagging Manual Version 2.0*. Centre for English Corpus Linguistics, Université Catholique de Louvain
- Halliday, M. A. K. (1994). Spoken and written modes of meaning. In O. Boyd-Barrett & D. Graddol (Eds.), *Media texts: Authors and readers* (pp. 51-73). Multilingual Matters in association with The Open University.
- Halliday, M. A. K., & Martin, J. R. (1993). *Writing science: Literacy and discursive power*. Pittsburgh University Press.
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461-473. <https://doi.org/10.1093/applin/amp048>
- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Working papers/Lund University, Department of Linguistics and Phonetics*, 53, 61-79. <https://journals.lub.lu.se/LWPL/article/view/2273>
- Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing*, 37, 13-38. <https://doi.org/10.1016/j.jslw.2017.06.001>
- Khushik, G. A., & Huhta, A. (2020). Investigating syntactic complexity in EFL learners' writing across common European framework of reference levels A1, A2, and B1. *Applied Linguistics*, 41(4), 506-532. <https://doi.org/10.1093/applin/amy064>
- Kuiken, F., & Vedder, I. (2019). Syntactic complexity across proficiency and languages: L2 and L1 writing in Dutch, Italian and Spanish. *International Journal of Applied Linguistics*, 29(2), 192-210. <https://doi.org/10.1111/ijal.12256>
- Kuiken, F., Vedder, I., & Gilabert, R. (2010). Communicative adequacy and linguistic complexity in L2 writing. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 81-100). EuroSLA Publisher.
- Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, 16(1), 33-51. <https://doi.org/10.1177/026553229901600103>
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474-496. <https://doi.org/10.1075/ijcl.15.4.02lu>
- Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190-208. [https://doi.org/10.1111/j.1540-4781.2011.01232\\_1.x](https://doi.org/10.1111/j.1540-4781.2011.01232_1.x)
- Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19(1), 85-104. <https://doi.org/10.1191/0265532202lt221oa>
- Martínez, A. C. L. (2018). Analysis of syntactic complexity in secondary education EFL writers at different proficiency levels. *Assessing Writing*, 35, 1-11. <https://doi.org/10.1016/j.asw.2017.11.002>
- McCarthy, P. M., & Jarvis, S. (2007). vocd: A theoretical and empirical evaluation. *Language Testing*, 24(4), 459-488. <https://doi.org/10.1177/0265532207080767>
- Michel, M. (2017). Complexity, accuracy, and fluency in L2 production. In S. Loewen & M. Sato (Eds.), *Routledge handbook of instructed second language acquisition* (pp. 50-68). Routledge. <https://doi.org/10.4324/9781315676968-4>
- Owen, N., Shrestha, P., & Bax, S. (2021). *Researching lexical thresholds and lexical profiles across the Common European Framework of Reference for Languages (CEFR) levels*

- assessed in the APTIS test (AR-G/2021/1) [Research report]. British Council. [https://www.britishcouncil.org/sites/default/files/owen\\_shrestha\\_and\\_bax\\_layout.pdf](https://www.britishcouncil.org/sites/default/files/owen_shrestha_and_bax_layout.pdf)
- Park, S. (2022a). Syntactic complexity in a learner written corpus and L2 speaking quality: Suggestions for distinguishing L2 speaking proficiency. *Journal of Language and Linguistic Studies*, 18(1), 361-371. <https://www.jlls.org/index.php/jlls/article/view/3062>
- Park, S. (2022b). A corpus-based comparison of syntactic complexity in spoken and written learner language. *Canadian Journal of Applied Linguistics*, 25(2), 47-70. <https://doi.org/10.37213/cjal.2022.32477>
- Park, S. (2023a). Grammatical accuracy among Korean EFL learners by proficiency level. *The Journal of Mirae English Language and Literature*, 28(4), 139-160. <https://doi.org/10.46449/MJELL.2023.11.28.4.139>
- Park, S. (2023b). Comparison of lexical complexity in L2 speaking and writing and factors predicting English speaking proficiency. *3L: Southeast Asian Journal of English Language Studies*, 29(1), 125-138. <http://doi.org/10.17576/3L-2023-2901-09>
- Read, J. (2007). Second language vocabulary assessment: Current practices and new directions. *International Journal of English Studies*, 7(2), 105-126. <https://revistas.um.es/ijes/article/view/49021>
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510-532. <https://doi.org/10.1093/applin/amp047>
- Tabari, M. A., Lu, X., & Wang, Y. (2023). The effects of task complexity on lexical complexity in L2 writing: An exploratory study. *System*, 114, Article 103021. <https://doi.org/10.1016/j.system.2023.103021>
- Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239-277). John Benjamins. <https://doi.org/10.1075/llt.11.15tav>
- Ure, J. (1971). Lexical density and register differentiation. In G. E. Perrem & J. L. M. Trim (Eds.), *Applications of linguistics: Selected papers of the Second International Congress of Applied Linguistics* (pp. 443-452). Cambridge University Press.
- Wang, L., & Yu, D. (2022). UAM Corpus Tool: A convenient and comprehensive software. *Scholars International Journal of Linguistics and Literature*, 5(10), 352-356. <https://saudijournals.com/articles/8085/>
- Yoon, H. J., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, 51(2), 275-301. <https://doi.org/10.1002/tesq.296>
- Yoon, S., Park, S., & Yoo, H. (2022). Learner corpus: Collection and application of the English part of Incheon National University Multi-Language Learner Corpus (INU-MULC). *Korean Journal of Applied Linguistics*, 38(4), 33-58. <https://doi.org/10.17154/kjal.2022.12.38.4.33>



## APPENDIX

### Overall written production levels (Council of Europe, 2001).

C2	Can write clear, smoothly flowing, complex texts in an appropriate and effective style and a logical structure which helps the reader to find significant points
C1	Can write clear, well-structured texts of complex subjects, underlining the relevant salient issues, expanding and supporting points of view at some length with subsidiary points, reasons and relevant examples and rounding off with an appropriate conclusion
B2	Can write clear, detailed texts on a variety of subjects related to his field of interest, synthesizing and evaluating information and arguments from a number of sources
B1	Can write straightforward connected texts on a range of familiar subjects within his field of interest, by linking a series of shorter discrete elements into a linear sequence
A2	Can write a series of simple phrases and sentences linked with simple connectors like ‘and’, ‘but’ and ‘because’
A1	Can write simple isolated phrases and sentences