Syntactic complexity in a learner written corpus and L2 speaking quality:

Suggestions for distinguishing L2 speaking proficiency

Shinjae Park [a] [1] iD

ªIncheon National University, South Korea

Abstract

Based on the quantitative analysis of L2 English texts from Korean undergraduates, the present paper demonstrates the possibility of distinguishing L2 speaking proficiency with differing measures of syntactic complexity in English writing. To this end, 14 measures of complexity were gauged using an L2 Syntactic Analyzer (Lu, 2010) in 89 EFL essays from a learner corpus to find out the best indicators of L2 speaking proficiency. As a result of the normality assumption test, the logistic regression analysis was applied and the model was significant in classification according to L2 proficiency and the classification was correct in 62.9% of the studied cases. Furthermore, it showed that the best predictor of L2 speaking proficiency was the complex nominal per T-unit (CNT). Specifically, upon the acquisition of CNT, EFL learners are 6.4 times more likely to achieve higher proficiency in speaking English. Taken together, this paper suggests that EFL students' quality of writing can be effectively used to distinguish their speaking ability.

Keywords: learner written corpus; syntactic complexity; EFL; logistic regression; assessment

1. Introduction

Widely acknowledged as a construct to effectively evaluate L2 writing development (Ai & Lu, 2013: 249–64; Lu, 2011:36–62; Martínez, 2018), linguistic complexity has been extensively used in previous research to quantitatively analyze meaningful linguistic characteristics of L2 writing (Ai & Lu, 2013: 249–64; Kim, 2014; Lu, 2011: 36–62; Martínez, 2018). Linguistic complexity embraces syntactic and lexical complexity, and both constructs are used to evaluate L1 and L2 students' linguistic performance, proficiency, and development (Kim, 2014; Kuiken, Vedder & Gilabert, 2010; Szmrecsanyi & Kortmann, 2012). With regard to syntactic complexity, which is the focus of this study, measures of syntactic complexity were used to assess L2 writing to investigate the functions of task complexity, writing ability, and teaching methods in L2 teaching of students of different ages, proficiency levels, and developmental periods (Ahmadian, 2012; Lu, 2011; Wu & Ortega, 2013).

In this study, syntactic complexity is defined as "the range and the sophistication of grammatical resources exhibited in language production" (Ortega, 2015: 82). Following Bulté and Housen (2014),

---

[1] Corresponding author.
  E-mail address: tlswo@naver.com.

this study hypothesizes "the more the components a feature or system comprises and the denser the relationships between its components, the more complex is the feature or system" (2014: 46). As argued by Ortega (2015: 82), "syntactic complexity indices the expansion of the capacity to use the additional language in ever more mature and skilful ways, tapping the full range of linguistic resources offered by the given grammar in order to fulfil various communicative goals successfully." Until recently, in the realm of syntactic complexity, L2 writing researchers are in agreement that the stronger L2 writers are, the longer and more complex syntactic structures they produce (Lu, 2011; Lu & Ai, 2015; Ortega, 2003; Wolfe-Quintero et al., 1998).

Recent studies (Biber, Gray & Poonpon, 2011; Biber, Gray & Staples, 2016; Hwang, Jung, & Kim, 2020) comparing speaking and writing in terms of syntactic complexity make the assertion that the two modes are totally separate, but, questions may be raised about whether there is any commonality between the two modes, and whether there is any method to support assessment of L2 performance in speaking and writing by analyzing them in a convergent manner. To this ends, this study examines that EFL students' syntactic complexity in their L2 writing can be effectively used to predict their speech proficiency. Therefore, the aim of this study is to suggest an alternative assessment method to distinguish L2 speaking quality through writing ability by measuring syntactic complexity indices in L2 writing and making use of holistic speaking proficiency of Korean L2 learners. Accordingly, when no oral data are available, researchers can use EFL students' writings to predict their speaking quality. Finding evidence in support of this prediction would suggest an alternative equation, based on quantitative measures gauging linguistic complexity by using L2 Syntactic Complexity Analyzer (L2SCA: Lu, 2010; 2011) and corresponding holistic ratings of speaking quality, to help educators adequately determine learners' proficiency in L2 speaking.

2. Literature Review

Among other linguistic parameters studied in L2 research into written output, syntactic complexity has been extensively investigated. However, using syntactic complexity as a means to measure writing proficiency and assess of language complexity (e.g., Lu, 2010, 2011; Martínez, 2018; Taguchi, Crawford, & Wetzel, 2013: 420–30; Yang, Lu, & Weigle, 2015: 53–67), most studies have relied on quantifiable complexity indices such as length of production unit, sentence complexity, and frequency certain sentence structures. Of these, T-unit (Hunt, 1965) is an important concept defined as the shortest grammatical chunk of a sentence as a unit of analysis. In other words, it is defined as the "shortest grammatically allowable sentences into which (writing can be split) or minimally terminable unit." More technically, a T-unit is a dominant clause and its dependent clauses: as Hunt said, it is "one main clause with all subordinate clauses attached to it" (Hunt 1965: 20). A T-unit is often used in the analysis of written and spoken discourse, such as in studies on errors in second language writing. In a review of 39 papers on L2 writing that investigated various indices of complexity, fluency, and accuracy, Wolfe-Quintero et al. (1998) found that T-unit length and clause length can be used as a measure of writing fluency. Other previously established indicators for syntactic complexity include the mean length of sentence (MLS) and mean number of T-units per sentence (Ortega, 2003).

In addition, traditional studies on writing have mainly focused on the use of clausal subordination. For instance, Flahive and Snow (1980) reported a strong relationship between clausal subordination (evaluated by the number of dependent clauses in each essay) and writing quality. Likewise, in a study on the relationship between a holistic evaluation of ESL writing proficiency and measures of complexity, Homburg (1984) found a strong link between finite clausal subordination and quality of writing; similarly, there was significant relationship between complexity as measured by length measures (namely, MLS and MLC) and holistic grading of composition quality.

Furthermore, there is substantial evidence documenting that more proficient L2 writers produce longer texts (e.g., Crossley & McNamara, 2012; Espada-Gustilo, 2011; Ferris, 1994; Lu & Ai, 2015; Martínez, 2018; McNamara et al., 2010). Indeed, longer essays tend to be rated higher than their shorter counterparts. For example, Ferris (1994) demonstrated that writing proficiency can be reasonably well predicted by the number of words and the length of production. Similar findings were also reported by Espada-Gustilo (2011) and Martínez (2018). For instance, based on the holistic ratings of the essays from secondary school-aged writers having various levels of proficiency, Martínez (2018) revealed a substantial relationship between syntactic complexity and writing proficiency. Specifically, she reported that the use of longer units was a strong indicator of high-quality output; in contrast, the common usage of simple utterences was found to be associated with lower writing quality. In line with these results, proficient L2 writers were found to typically create longer texts with more versatile diverse vocabulary and a higher number of words per sentence (Kim, 2014; McNamara et al., 2010).

However, there is some evidence that more proficient L2 writers do not necessarily use more clauses. In this respect, as argued by Rimmer (2006), important features of syntactic complexity are also phrasal features such as noun post-modifiers. Indeed, in a study that used Syntactic Analyzer to analyze college-level Chinese learners' essays written in the students' L2, Lu (2011: 36–62) found that proficiency was predicted by some of these measures, including mean length of a clause, complex nominals per clause, complex nominals per T-unit, MLT, and MLS. Furthermore, in a study that compared 600 English compositions by Chinese writers with the Louvain Corpus of Native English Essays, Ai and Lu (2013) found that the two groups differed in terms of the MLS, MLC, MLT, and complex nominals. More specifically, in an analysis of argumentative written samples by EFL learners, Taguchi et al. (2013) found that writing quality was largely positively influenced by noun phrase modification (i.e., preceding attributive adjective and prepositional phrase as post-modifiers of nouns).

In a more recent study, Lan and Sun (2019) made a comparison the argumentative essays of Chinese English learners to academic journal articles and discovered that frequent noun modifiers appeared in students' writings much less than in academic journal articles, and the higher-proficiency students tended to use more modifiers in writing. Similarly, Khushik and Huhta (2020) analyzed English argumentative essays by teenagers from Pakistan and Finland, and the length of production units, phrasal density, and subordination differed according to proficiency level. Casal and Lee (2019) also examined the association among syntactic complexity and writing quality in assessed research papers produced by 280 ESL undergraduates, and revealed that mean length of T-units and phrasal measures differed across levels.

Of note, however, Halliday (2006) argued that the complexity of speech should be focused on different aspects from writing because the underlying linguistic characteristics of them are different. In this regard, through the attempts of several influential studies, while major grammatical complexities of speech were found to be clausal subordination features, those of academic writing were found to be phrasal embedding (Biber et al., 2011; Biber et al., 2016; Kyle & Crossely, 2018). Moreover, considering that conversation is acquired first, clausal subordination features were not reported to be associated with a high degree of production complexity (Biber et al., 2011). Instead, it is argued that, since a large amount of complex phrasal embedding was produced and limited to more specialized circumstances of formal forms of writing, these structures represent a extensively higher degree of complexity (Ibid.). Recently, Hwang et al. (2020) examined syntactic complexity in spoken and written data in English produced by 122 Korean children. The results revealed that the written data involved more subordination (DCT), longer structures (MLT), and more verb phrases (VPT) than the spoken data, while the spoken data included more coordination (CPT) than the written data.

On the other hand, studies are also found that may imply that the underlying mechanisms are shared. Cleland and Pickering (2006) discovered that a group of UK students tended to repeat syntactic form across the two modes to the same extent that they did within either mode. Kormos (2014) found that there was no difference in the percentage of subordinating clauses between the mods of Hungarian English learners, while the writing contained more modifiers per noun phrase. Therefore, if the sharing between the syntax complexity represented in one mode and the specific characteristics of the other mode is found, and further studies that have not been conducted before are attempted, linguistics may be expanded to new and practical areas.

As is clear from the review presented above, most previous studies have investigated syntactic complexity indices as predictors of language development or proficiency in either the writing mode or the speaking mode. However, since speaking materials are acknowledged as unwieldy and messy (Rajadurai, 2007), there have been a number of studies on data of L2 writers. Therefore, few attempts have been made to characterize syntactic complexity in both speaking and writing, especially studies that have explored the complexity in cross-mode methods such as applying syntactic characteristics of one mode to predict proficiency of other modes.

To fill these gaps, I examine whether 14 complexity measures in EFL writing samples can be used to effectively predict English speaking proficiency. From the applied perspective, considering that this study is the first to empirically test the convergence across modes (writing and speaking), these results can provide valuable insights for practitioners in the EFL field. This study addresses the following research questions:

1. Which syntactic features significantly predict L2 speaking proficiency levels?

2. Is it possible to derive a significant logistic regression equation to distinguish L2 speaking proficiency?

3. If so, to what extent can the derived equation predict the classification of the observed data?

3. Experimental/Materials and Methods

3.1. Corpus selection

For the analysis, I used the Incheon National University Multi-Language Learner Corpus (INU-MULC, see Park, 2020a; Park & Yoon, 2021) to select 89 essays written by college-level English learners. The MULC is a corpus of data written and spoken by undergraduate students at a university in Korea and is currently being compiled. Thus far, the MULC contains 109 conversations, 224 monologues, and 139 essays. To write the essays, the participants were asked to select one among five topics, and the essays were later electronically stored in the MULC database. In the corpus, each participant's speaking proficiency was registered based on the rubrics of Common European Framework of Reference for Languages (CEFR). The ratings were given by three native English speakers who were specifically trained for that purpose: They are currently lecturers at the university that has been compiling this corpus, and continuously have helped for this project. They rated 5% of the speaking corpus in the pilot test, and after high consistency between their assessments was achieved, they actually began to conduct the assessment. In this study, following the criteria articulated by Crossley and McNamara (2013), only the texts that exceeded the 100-word cut-off in length were selected for further analysis. Essays shorter than 100 words were deemed unreliable as they do not contain a sufficient number of features owing to their short length.

Although the selection of essays from the corpus was random, I tried to ensure that the number of corresponding items remained as consistent as possible. The writers of the essays included in the sample were 54 females and 35 males (see Table 1). Based on proficiency levels, the essays were

grouped as follows: (1) low proficiency group (LP; N = 47; 52.8%) and (2) high proficiency group (HP; N = 42; 47.2%). In the sample of selected essays, 25 participants (28.1%) wrote essays on marriage, 22 (24.7%) on review of a book read recently, 17 (19.1%) on mobile phones, 16 (18%) on uniforms, and 9 (10.1%) on clubs. The topics of the essays were as follows:

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?

5) Write a brief review of the book you have read recently

Table 1. Participants' characteristics

Gender		Topic					Proficiency	
Male	Female	Marriage	Literature	Mobile Phone	Uniform	Club	Low	High
35	54	25	22	17	16	9	47	42
39.3%	60.7%	28.1%	24.7%	19.1%	18.0%	10.1%	52.8%	47.2%
		Total: 89 (100%)						

3.2. Computational tools and variables

To measure linguistic indices used for in this study, L2SCA was used. Further detail on each index can be found in Lu (2010; 2011). The five major categories were as follows: (1) length of production units, (2) overall sentence complexity, (3) number of subordination clauses, (4) number of coordination clauses, and (5) phrasal sophistication. The analysis focused on 14 indices in the following five categories: (1) length of production (N = 3), (2) sentence complexity (N = 1), (3) number of subordination clauses (N = 4), (4) number of coordination clauses (N = 3), and (5) phrasal sophistication (N = 3). The specific linguistic features and indices gauged by the computational tool (L2SCA) are summarized in Table 2.

Table 2. The 14 syntactic automated complexity measures (Lu, 2010, 2011)

Length of production unit	Mean length of clause (MLC: words/clause)
	Mean length of sentence (MLS: words/sentence)
	Mean length of T-unit (MLT: words/T-unit)
Sentence complexity	Clauses per sentence (CS: clauses/sentence)
Subordination	Clauses per T-unit (CT: clauses/T-unit)
	Complex T-unit per T-unit (CTT: complex T-units/T-unit)
	Dependent clauses per Clause (DCC: dependent clauses/clause)
	Dependent clauses per T-unit (DCT: dependent clauses/T-unit)
Coordination	Coordinate phrases per clause (CPC: coordinate phrases/clause)
	Coordinate phrases per T-unit (CPT: coordinate phrases/T-unit)
	T-units per sentence (TS: T-units/sentence)
Particular structures	Complex nominals per clause (CNC: complex nominals/clause)

Complex nominals per T-unit (CNT: complex nominals/T-unit)
Verb phrases per T-unit (VPT: verb phrases/T-unit)

4. Results

The statistical methods to determine best fitting model to describe the relationship between the outcome (i.e., proficiency in English speech) and a set of independent variables (i.e., complexity measures) include discriminant analysis and logistic regression analysis. To determine whether the normality assumption is satisfied, the Box's M test was conducted (Table 3). The significance value (p=.000) indicates that the data used in this study differ across factors in terms of the multivariate normal distribution (Table 3).

Table 3. Box's M test of equality of covariance matrices

Box's M		181,949
	Approx.	4.557
	df1	36
F	df2	24,795.25
	Sig.	.000

With non-normality as in this case, the logistic regression method is applied to analyse categorical outcome variables (Gong & Sun, 2009: 1366-71; Park, 2020b; Sio & Ismail, 2019: 29). Dependent variables of L2 speaking proficiency were divided into LP and HP for the binary logistic regression. Next, stepwise regression analysis was conducted to find indices that predict L2 speaking quality, and as a result, the significant independent variable of the model was revealed to be complex nominals per T-unit (CNT, $p < .05$) (Table 4).

Table 4. Summary of Stepwise Multiple Regression Analysis

Model	Unstandardized Coefficients		Standardized Coefficients	*t*	Sig.
	B	S.E	Beta		
(Constant)	.947	.157		6.036	.000
CNT	.393	.111	.354	3.528	.001

The results of the overall test of the model using the likelihood ratio are presented in Table 5. The model was statistically significant, $\chi^2(1) = 12.033$ ($p < .05$); therefore, the model not only identified the sub-scales that influenced L2 proficiency but also adequately fit the data.

Table 5. Omnibus test of model coefficients

		Chi-square	df	Sig.
	Step	12.033	1	.001
Step 1	Block	12.033	1	.001
	Model	12.033	1	.001

The two methods used in this study to calculate the explained variation are presented in Table 6. According to Nagelkerke R^2 values (Nagelkerke R^2 = .169), the explained variation in the dependent variables in this model was 16.9%.

Table 6. Summary of Stepwise Multiple Regression Analysis

Step	-2 Log likelihood	Cox and Snell R-square	Nagelkerke R-square
1	111.066	.126	.169

The assessment results of the effectiveness of the predicted classification compared with the actual classification, i.e., the observed number of students in each of the two proficiency groups and the predicted number according to logistic regression (p < .05) are reported in Table 7. As is seen in Table 7, about two-thirds (72.3%) of all cases in the LP group and more than half (52.4%) of all cases in the HP group were correctly classified. The overall classification accuracy amounted to 62.9%.

Table 7. Category prediction

Observed		Predicted		
		L2 Proficiency		Correct
		Low	High	%
Step 1	Low	34	13	72.3
	High	20	22	52.4
	Overall Percentage			62.9

Table 8 summarizes the results of the logistic regression conducted to assess the influence of predictors on L2 proficiency in the data. First, with regard to the odds ratio (i.e., Exp(B)), CNT, with an odds ratio of 6.386, is the strongest predictor of L2 proficiency. This finding proposes that, upon acquisition of CNT, EFL learners were more than six times more likely to achieve HP. Second, I computed the coefficients and the Wals values of CNT, which were used to evaluate the statistical significance of each independent variable (Wals: 9.598, p < .05). Finally, based on the results, I derived the following estimated logistic regression equation (See Eq.1) for English proficiency levels.

Eq.1: **log (speaking proficiency level) = -2.578 + 1.854(CNT)**

Employing this equation, the boundary of 0.5 was used to split the participants into two groups (LP and HP). Specifically, when the values of learners' CNT were entered into the derived logistic regression equation, the participants whose values were equal to or below 0.5 or less were assigned to the LP group, while those scoring higher than 0.5 were added to the HP group.

Table 8. Variables in the equation

		B	S.E,	Wals	df	Sig.	Exp(B)	Low	High
Step 1	CNT	.1.854	.598	9.598	1	.002	6.386	1.976	20.638
	Constant	–2.578	.821	9.860	1	.002	.076		

5. Discussion and Conclusions

In this study, I used a corpus-based approach and adjusted design to investigate Korean L2 learners' writing characteristics. To this end, I systematically examined syntactic complexity features

of EFL learners' essays in the following major areas: (1) length of production units, (2) overall sentence complexity, (3) number of subordination clauses, (4) number of coordination clauses, and (5) phrasal sophistication. Based on the results, I identified the linguistic features in EFL writing that best predict L2 learners' speaking proficiency levels.

In previous research, syntactic complexity features in L2 writing have been widely used to improve writing quality and evaluate writers' proficiency levels. Unlike the aforementioned research, this study contributes to expanding the research scope, proposing a complementary method, i. e., a crossover research method between modes, to distinguish L2 speaking levels by measuring syntactic complexity in writing. The results of the stepwise multiple regression analysis performed in this study suggest that the best determinant is complex nominals per T-unit (CNT) (i. e., category: phrasal sophistication). This study is meaningful in that it revealed that this phrasal complexity, which is closely related to the writing, is also important for predicting speaking proficiency. These results are in line with several previous studies that reported a positive relationship between the use of complex nominals and L2 proficiency (Biber et al., 2011; Biber et al., 2016, Taguchi et al., 2013; Kyle & Crossley, 2018). Furthermore, this study is important in that it demonstrates that the phrasal complexity, such as complex nominal, is an important measure in speaking as well as writing through the results that the complex nominal per T-unit in writing is the best predictor of L2 speaking. It is also suggested that these issues need to be addressed more in future studies. In addition, the logistic regression model explained 16.9% of the total variance in classification and was statistically significant, $\chi2(1) = 12.033$, $p < .05$. The accuracy rate of the classification amounted to 62.9%. It also suggests that, upon acquisition of CNT, learners are 6.4 times more likely to achieve higher proficiency in speaking English. In other words, the model empirically demonstrates that educators can effectively classify L2 learners according to the proficiency of L2 speaking when only measuring CNT among multiple indices in the writings of the learners in the EFL context.

6. Implications

The results of this study provide interesting implications for L2 researchers and educators. First, the present results suggest that measuring complexity in written data can reliably predict students' proficiency in English speaking quality. This finding is important because, unlike speech data that requires a complex process of collection and transcription, written data are relatively easy to obtain and analyze (Park, 2020a; Park & Yoon, 2021). In other words, it is valuable in contexts when no official English test scores are available, or when, owing to limited time, place, or insufficient budget, it is impossible to conduct speaking tasks. Second, EFL instructors ensure that, through explicit vocabulary instruction, Korean L2 learners acquire a rich and diverse vocabulary. According to the present results, the learners who improved their CNT abilities from phrasal sophistication were over 6.4 times more likely to achieve a higher proficiency in speaking English. This finding suggests that, to enhance L2 learner's nominal and phrasal sophistication knowledge, relevant pedagogical interventions would be needed. Additionally, EFL instructors should prioritize the development of phrasal sophistication in their teaching. Finally, making use of the equation proposed in this study, educators can measure only one dimension to effectively classify EFL learners according to levels of proficiency in L2 speaking. The equation is valid since I attained more wide-ranging results than could have been afforded by analyzing a fewer number of syntactic features.

Given the scope of the present study and the accessibility of the corpus used, there are several limitations. First, in terms of the scope of the study, further research is needed to explore whether the results reported herein show the same pattern as those of non-Korean EFL learners or learners with different academic backgrounds. Likewise, it is worth investigating whether the proposed approach would work with other topics, writing types, context-related factors, or test formats, i.e., speaking

tests. Second, the MULC was designed to become publicly accessible inside and outside school environments; however, it is currently in the stage of POS tagging and continues to expand, thereby delaying its release. In addition, a more user-friendly data arrangement is still required for the corpus' widespread use. Nevertheless, despite the limitations outlined above, this study provides valuable insights into the effects of syntactic complexity measures on L2 speaking proficiency. By conducting a thorough analysis of syntactic features in EFL essays and proposing a complementary method to evaluate L2 speaking proficiency, this study makes a meaningful contribution to the field of L2 speaking and writing research.

Funding

References

Ahmadian, M. J. (2012). The effects of careful guided online planning on complexity, accuracy and fluency in intermediate EFL learners' oral production: The case of English articles. *Language Teaching Research, 16*(1), 129-149.

Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. *Automatic treatment and analysis of learner corpus data*, 249-264.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development?. *TESOL Quarterly, 45*(1), 5-35.

Biber, D., Gray, B., & Staples, S. (2016). Predicting patterns of grammatical complexity across language exam task types and proficiency levels. *Applied Linguistics, 37*(5), 639-668.

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of second language writing, 26*, 42-65.

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.

Cleland, A. A., & Pickering, M. J. (2006). Do writing and speaking employ the same syntactic representations?. *Journal of Memory and Language, 54*(2), 185-198.

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.

Crossley, S., & McNamara, D. (2013). Applications of text analysis tools for spoken response grading. *Language Learning & Technology, 17*(2), 171-192.

Espada-Gustilo, L. (2011). Linguistic features that impact essay scores: A corpus linguistic analysis of ESL writing in three proficiency levels. *3L: Language, Linguistics, Literature, 17*(1), 55-64.

Ferris, D. R. (1994). Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly, 28*(2), 414-420.

Flahive, D. E., & Snow, B. G. (1980). Measures of syntactic complexity in evaluating ESL compositions. *Research in language testing*, 171-176.

Gong, J., & Sun, S. (2009, June). A new approach of stock price prediction based on logistic regression model. In *2009 International Conference on New Trends in Information and Service Science* (pp. 1366-1371). IEEE.

Halliday, A. K., & Webster, J. J. (2006). Language of Science (Vol. Volume 5). *London: Bloomsbury Academic*.

Homburg, T. J. (1984). Holistic evaluation of ESL compositions: Can it be validated objectively?. *TESOL quarterly, 18*(1), 87-107.

Hunt, K. W. (1966). Recent measures in syntactic development. *Elementary English, 43*(7), 732-739.

Hwang, H., Jung, H., & Kim, H. (2020). Effects of Written Versus Spoken Production Modalities on Syntactic Complexity Measures in Beginning-Level Child EFL Learners. *The Modern Language Journal, 104*(1), 267-283.

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.

Kim, J. Y. (2014). Predicting L2 Writing Proficiency Using Linguistic Complexity Measures: A Corpus-Based Study. *English Teaching, 69*(4), 27-51.

Kormos, J. (2014). Differences across modalities of performance. In H. Byrnes and R. M. Manchón (Eds), Task-based language learning: Insights from and for L2 writing (pp. 193-216). *Amsterdam: John Benjamins*.

Kuiken, F., Vedder, I., & Gilabert, R. (2010). Communicative adequacy and linguistic complexity in L2 writing. *Communicative proficiency and linguistic development: Intersections between SLA and language testing research, 1*, 81-100.

Kyle, K., & Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *The Modern Language Journal, 102*(2), 333-349.

Lan, G., & Sun, Y. (2019). A corpus-based investigation of noun phrase complexity in the L2 writings of a first-year composition course. *Journal of English for Academic Purposes, 38*, 14-24.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International journal of corpus linguistics, 15*(4), 474-496.

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45*, 36-62.

Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing, 29*, 16-27.

Martínez, A. C. L. (2018). Analysis of syntactic complexity in secondary education EFL writers at different proficiency levels. *Assessing Writing, 35*, 1-11.

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication, 27*, 57-86.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*, 492-518.

Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing, 29*, 82-94.

Park, S. (2020a). A Corpus Study of Modal Verbs in Korean Learners' Speech. *Journal of Linguistic Studies, 25*(2), 121-137.

Park, S. (2020b). Determining L2 Proficiency in the Production and Perception of Consonant Clusters. *3L: Language, Linguistics, Literature®, 26*(3), 41-52.

Park, S. & Yoon, S. (2021). Syntactic Complexity of EFL Learners' Casual Conversation, Monologue, and Writing. *The Journal of Studies in Language, 37*, 75-89

Rajadurai, J. (2007). Capturing L2 Spoken English: Methodological Challenges. *Linguistics Journal, 2*(3), 41-58.

Rimmer, W. (2006). Measuring grammatical complexity: The Gordian knot. *Language Testing, 23*, 497-519.

Sio, J. & Ismail, R. (2019). Binary logistic regression analysis of instructional leadership factors affecting English language literacy in primary schools. *3L: Language, Linguistics, Literature®, 25*(2), 25-37.

Szmrecsanyi, B., & Kortmann, B. (2012). Introduction: Linguistic complexity. In *Linguistic Complexity* (pp. 6-34). de Gruyter.

Taguchi, N., Crawford, W., & Wetzel, D. Z. (2013). What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program. *TESOL Quarterly, 47*, 420-430.

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity* (No. 17). University of Hawaii Press, Hawaii.

Wu, S. & Ortega, L. (2013). Measuring global oral proficiency in SLA research: A new elicited imitation test of L2 Chinese. *Foreign Language Annals, 46*, 680-704.

Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing, 28*, 53-67.

AUTHOR BIODATA

Dr. Shinjae Park is a research professor at Incheon National University, South Korea. Her research interests are corpus studies, phonetics, phonology, and language teaching and assessment.