

## Predicting L2 Speaking Proficiency Using Syntactic Complexity Measures: A Corpus-Based Study

PARK, SHINJAE  
Sehan University, South Korea  
tlswo@naver.com

### ABSTRACT

*This paper discusses the syntactic complexity factors contributing to the achievement of a higher proficiency level in English speech. Here I have examined complexification at the sentential, clausal, phrasal and nominal level of syntactic organisation in a Korean learner spoken corpus using quantitative measures and compared the scores with holistic ratings of learners' overall speaking quality. After the normality assumption analysis confirmed that the logistic regression was appropriate, an analysis was performed to ascertain the effects of complexity measures on participants' L2 proficiency. First, length-based complexity features, namely, MLT and coordinated phrases, namely, CPT and CPC were found to be predictors for English speaking proficiency. Next, the logistic regression model was statistically significant and explained 36.3% of the variance in classification according to L2 proficiency and correctly classified 75.4% of cases. Results also showed that when learners come to use the coordinated phrases per clause proficiently, they were over 24 times more likely to achieve higher proficiency in spoken English. Finally, an effective equation was proposed to help educators classify EFL learners according to proficiency in L2 speech after gauging the selected complexity dimensions. However, more comprehensive studies which consider other methods of unit segmentation for spoken data or include more measures to predict L2 speech proficiency, are necessary to verify the results of this study.*

*Keywords: Learner spoken corpus; Syntactic complexity; Monologue; EFL; Logistic regression*

### INTRODUCTION

Syntactic complexity is defined as the extent to which language produced while performing a task is elaborate and diverse (Ellis, 2003, p. 340); furthermore, it has long been recognised as a measure of proficiency in second language acquisition (SLA; Ai & Lu, 2013; Taguchi et al., 2013; Yang et al., 2015; Kyle & Crossley, 2018). The present study adopts the definition of Bulté and Housen (2014, p. 46) that the more components a feature or system comprises—in addition to denser relationships between its components—the more complex the feature or system is. This study also measures syntactic complexity indices because they represent the expansion of the capacity for using additional language in mature and skillful ways, thus making use of the complete range of given linguistic resources by the given grammar such that various communicative goals are realised successfully (Ortega, 2015, p. 82).

While the analysis of the syntactic complexity of language has mostly been used in writing until recently, the same complexity analysis can be adopted in the same way by examining the transcription materials of spoken data. Studies examining the relationship between a speaker's speaking ability and syntactic complexity scores have emerged, but the number of studies is fewer than those of writing (Iwashita, 2006, pp. 151–169; Chen & Zechner, 2011, pp. 722–731; Lintunen & Makila, 2014, pp. 377–399; Neary-Sundquist, 2017, pp. 242–262). Analysing syntactic complexity indices in a spoken corpus is generally no simpler than using a written corpus. First, physical reasons should be taken into consideration why data collection and transcription are difficult, which often consume a lot of time, human resources and effort (Yoon & Park, 2021, pp. 603–604). In addition, spoken output is generally less clean than written data even after the transcription is systematically completed by linguistic experts. That is, when learners are not proficient in L2 speaking, it can lead to repetition, pauses, stuttering, false starts and self-correction, which may then negatively affect the

accurate analysis of syntactic complexity, as it can result in significant meaningless utterances by learners with less proficiency in English (Chen & Zechner, 2011, pp. 722–731; Foster et al., 2000, pp. 354–375). This is because the length of utterance is a factor of syntactic complexity measurement. Therefore, researchers need to eliminate these disfluencies first when gauging syntax complexity indices if adopting a spoken corpus as the data set in their research (Chen & Zechner, 2011, pp. 722–731; Foster et al., 2000, pp. 354–375). Furthermore, few studies have suggested practical ways to predict L2 proficiency using spoken data to gauge syntactic complexity measures (Iwashita, 2006; Iwashita et al., 2008; Chen & Zechner, 2011; Park & Yoon, 2021; Neary-Sundquist, 2017).

In addition, several L2 studies have used an insufficient number of complexity measures although there are some complexity measures available in the L2 acquisition literature (Bulté & Housen, 2014, pp. 42–65). As a result, L2 complexity research studies often suffer from low content validity (Bulté & Housen, 2014). Hence, an approach is needed that regards complexity as an increasingly complex construct comprising of several components. Further, more measures capturing various facets of syntactic complexity at the sentential, clausal and phrasal level are also required in studies on L2 data. (Norris & Ortega, 2009, pp. 555–578; Ai & Lu, 2013, pp. 249–264; Bulté & Housen, 2014, pp. 42–65; Yang et al., 2015, pp. 53–67; Kyle & Crossley, 2018, pp. 333–349).

Accordingly, this study attempts to fill in some gaps in the current research on complexity by examining the relationship between 14 syntactic complexity measures of L2 speaking and proficiency and providing mathematical equations to help effectively determine L2 proficiency. Unlike previous studies that adopted insufficient complexity measures to investigate written data, this study not only analyses multidimensional complexity but also focuses on spoken data, which has not been studied sufficiently to date. Furthermore, this study suggests a novel and alternative equation that could help educators determine learners' proficiency in L2 speaking, which was produced after comparing the holistic ratings of speaking quality and quantitative measures gauging complexification at the sentential, the clausal, the phrasal and the nominal level of syntactic organisation.

## LITERATURE REVIEW

Language complexity often includes concepts of syntactic and lexical complexity and is used as a measure of linguistic performance and proficiency in native or SLA studies (Wang & Slater, 2016, pp. 81–86). This present study focuses on syntactic complexity, which is understood broadly as 'the range and the sophistication of grammatical resources exhibited in language production' (Ortega, 2015, p. 82). Syntactic complexity has received the most attention in L2 writing studies (Lu, 2010, pp. 474–496; Lu & Ai, 2015, pp. 16–27).

Norris and Ortega (2009, pp. 555–578) argue that all three levels of syntactic organisation, namely, sentential, syntactic and phrasal must be measured in order to examine L2 development since L2 learners integrate these levels of syntactic complexity at different stages of development. In line with this argument, a series of recent research studies selected various measuring sets to gauge syntactic complexity. According to Yang et al. (2015, p. 55), syntactic complexity should be approached as a multidimensional construct; therefore, they selected eight measures, which represented several interconnected sub-constructs including clausal coordination, clausal subordination, phrasal coordination and noun phrase complexity. In addition, Bulté and Housen (2014, p. 47) have selected 10 measures of sentence complexity including the mean length of sentential units, clauses and propositions combining clause integration strategies.

On the other hand, Lu (2010; 2011) has selected 14 measures, which were then further divided into five categories of syntax complexity, each representing different but interrelated aspects of syntax complexity. The five categories included length of production, sentence complexity, amount of subordination, amount of coordination and the degree of syntactic sophistication. All of these measures meet the criteria I proposed, that is, that multiple measures should be examined to measure different levels of syntactic complexity. Therefore, this present study adopts the set of measures chosen in Lu's study (2010; 2011) together with descriptions that have been used in SLA literature when necessary. In addition, this current study uses the T-unit level as well as the sentence level to assess syntactic complexity because the SLA literature suggests that the analysis at sentence level is adequate for the description of the syntactic complexity of adults and that T-unit level is particularly useful for describing the syntactic complexity of oral production (Bardovi-Harlig, 1992).

#### COMPLEXITY ACROSS PROFICIENCY LEVELS: WRITING

A series of studies have investigated the relationship between syntactic complexity and writing quality by examining the development of syntactic complexity. They tend to show that this construct increases over time and L2 writing ability develops with more instruction. The extensiveness of various syntactic complexity measures indicating L2 writing quality has been investigated by certain cross-sectional studies (Taguchi et al., 2013; Yang et al., 2015). In a study designed to select measures of syntactic complexity that contribute to the quality of L2 writing, Taguchi et al. (2013, pp. 420–430) have analysed argumentative essays written by English learners and found that noun phrase modification contributed to writing quality. Meanwhile, Yang et al. (2015) investigated the relationship between syntactic complexity and the writing quality of ESL graduate students' essays using the TOEFL iBT Writing Scoring Guide (2015, p. 58). They found that two measures of the mean length of sentence and the mean length of T-unit were significant predictors of the quality of writing.

For decades, syntactic complexity has been quantified with regard to length of a unit (e.g., sentence, clause, or T-unit) and clausal subordination, and recently, phrasal complexity (Biber et al., 2011; Kyle 2016) has also been added (Yoon & Park, 2021, pp. 600-603). The T-unit (e.g., minimal terminable unit) has been widely considered as a standard segmentation unit, proposed by Hunt (1965, p. 21) and he defined the T-unit as a unit that consists of one main clause and (optional) subordinate clauses and non-clausal units or sentence fragments attached to it. If I adopt the definition, the following sentence has two T-units (i.e., *a* and *b*).

Example: "Since he got so upset, I didn't think we would want to wait for Tina to come back, and I wanted to leave as soon as possible." (adopted from Yoon & Park, 2021, p. 601)

This sentence is segmented into two T-units as follows.

- a. Since he got so upset, I didn't think we would want to wait for Tina to come back
- b. I wanted to leave as soon as possible.

The relationship between syntactic complexity and proficiency in L2 writing has been explored by some recent studies (Ai & Lu, 2013; Kim, 2014; Lorenzo & Rodríguez, 2014; Lu & Ai, 2015). Their common findings tended to indicate that L2 proficiency can impact syntactic complexity. Lu (2011) analysed several English essays written by Chinese university learners using 14 syntactic complexity measures and found that some of these measures (i.e. complex nominals per clause; mean length of a clause; complex nominals per T-unit; mean length of sentence; mean length of T-unit) were related to proficiency. Kim (2014) demonstrated that more proficient L2 writers were capable of creating longer texts using more diverse vocabulary and were able to write more complex nominalisations and more words per

sentence than less proficient learners. A corpus of narratives of subjects ranging from the third year of secondary education to the second year of post-compulsory education was analysed by Lorenzo and Rodríguez (2014, pp. 64–72). As a result, learners in the lowest grade produced texts that lacked coordinate phrases and dependent clauses. However, over time, their essays improved in terms of quality, and their writing became more syntactically complex. The data showed significant progress in the mean length of clauses, sentence subordination, as well as complex nominals in each clause and in verb diversity and verb tenses (Lorenzo & Rodríguez, 2014, p. 70).

Wang and Slater's (2016, pp. 81–86) study on the written data of English learners in China has revealed that complex nominals, the mean length of sentences and the mean length of clauses were correlated to English proficiency. Ai and Lu (2013) compared 600 English writing samples by Chinese English learners with the Louvain Corpus of Native English Essays and discovered that the two groups had differences in the mean length of clauses, mean length of sentences, mean T-unit length and complex nominals. However, they also determined no difference in the use of sentential coordination. Martínez (2018) examined complexification at three levels of syntactic organisation, that is, sentential, clausal and the phrasal level of syntactic organisation of English writing in lower intermediate and intermediate writers and then compared the scores of the selected measures with holistic ratings of their overall writing quality. Data revealed that intermediate students outperformed lower intermediate students as far as the general quality of the compositions were concerned, and in all syntactic complexity measures for all but one measure (i.e. the compound complex sentence ratio), and the increase was significant.

#### COMPLEXITY IN LEARNER SPEECH

This section reviews not only those studies where the target is English learners (Iwashita, 2006; Iwashita et al., 2008; Chen & Zechner, 2011; Park & Yoon, 2021), but also those that are learners of other languages (Neary-Sundquist, 2017). In addition, studies comparing L2 learner performance based on the complexity of written and oral tasks also indicate a relatively small subset (Lintunen & Makila, 2014), and some studies have argued that the complexity of the two modes is typically different (Kyle & Crossley, 2018; Biber et al., 2011; Halliday, 2006), while some studies suggest the opposite conclusion (Park & Yoon, 2021, pp. 82-87).

For instance, Iwashita (2006) found that the number of T-units and the number of clauses per T-unit (i.e. length-based complexity measures) were good predictors of Japanese L2 learners' speaking ability. Iwashita et al. (2008) examined four measures of complexity (i.e. clauses per T-unit, dependent clause ratio, mean length of utterance and verb phrase ratio) in oral L2 English data. They found that only one measure, that is, the mean length of utterance, was positively correlated with proficiency. In this study, it was noteworthy that the dependent clause ratio remained fairly flat regardless of proficiency levels. Neary-Sundquist (2017) examined the amount of subordination and coordination and phrasal complexity in the oral data of German learners at the intermediate, advanced and superior proficiency levels. As per her findings, different patterns of use emerged in all three complexity measures as proficiency levels increased. The results also revealed that the mean length of clauses was the most effective measure for distinguishing between proficiency levels. Halliday (2006) asserted that the complexities of speech are dramatically different from those of academic writing and, specifically, that clausal subordination features are the major grammatical complexities of speech, whereas complexities of academic writing are phrasal (Kyle & Crossley, 2018). Moreover, Biber et al. (2011) asserted that clausal subordination features do not illustrate an increased degree of production complexity because conversation is acquired first. Conversely, they hypothesised that many types of complex phrasal embedding represent a considerably

higher degree of production complexity as these grammatical structures are produced only in more specialised circumstances of formal writing (Biber et al., 2011).

Park and Yoon (2021), on the other hand, examined a learner corpus obtained from 40 undergraduate learners of English, which comprises casual conversations, monologues, and writings. They found that both monologue and writing modes elicited greater syntactic complexity than conversation. Moreover, syntactic complexity was not significantly different between monologues and writings only, except for Complex Nominals per Clause. They assumed that similar execution environments in the two modes are the reason why there is little difference in the degree of the use of complex structures. In the data collection for monologue and writing, the participants were given various everyday topics different from the spontaneous conversation task, and they chose their own topics of interest. Even if they were supposed to respond to the topics as soon as possible, they could spend seconds to minutes planning their production while choosing the topic. They concluded that a monologue-like manner such as presenting motivating topics may enhance the use of complex structures in L2 conversations.

#### GAPS FILLED BY THIS STUDY

Some of the previous studies on the complexity of EFL learner data at different proficiency levels discussed above are relevant to this study, revealing several gaps that this present study should fill. This analysis contributes new findings to the study of L2 complexity by examining oral data from Korean EFL learners while using as many indices of complexity as possible. Furthermore, this study presents an effective and alternative method for determining learners' proficiency by deriving a discriminant equation with selected complexity variables.

Therefore, this current study attempts to fill several gaps in the research on the complexity of L2 speech. First, most previous studies focusing on complexity have examined written data. Second, many recent studies have used only some of the matrices of syntactic complexity. Third, most studies have only tried to find predictors of L2 proficiency while comparing complexity scores within different proficiency groups. The author is unaware of any extant study that has produced discriminant equations for categorising EFL learners according to L2 proficiency while gauging complexity measures. Considering the existing research on the complexity of the learner spoken corpus, the current study aims to investigate the following research questions:

Research Question 1: Does the dataset use basic assumptions sufficient to perform a discriminant analysis?

Research Question 2: If not, how can an equation derived through logistic regression analysis be used to categorise L2 learners with low proficiency from a high proficiency group?

Research Question 3: To what extent can the derived equation predict the categorisation of the data observed?

#### METHODOLOGY

The main focus of this study was to identify the valid factors that indicate categories of learners according to their speaking proficiency by calculating syntactic complexity measure scores and then producing practical and mathematical equations to help determine learners' L2 speaking proficiency.



## SAMPLES

The data were obtained from the INU Multi-Language Learner Corpus that had compiled spoken and written data from more than 300 EFL undergraduates in South Korea from 2018 to 2020. When compiling the spoken corpus, learners were recruited regardless of their major, gender and age, so spoken data from students with various metadata were compiled. The collected data were then systematically transcribed by linguistics experts in three stages and were strictly managed each year by an IRB review and approval from the institution concerned. An exhaustive illustration of the sampling design is available in other studies of the present author (Park & Yoon, 2021, pp. 80-81; Yoon & Park, 2021, pp. 607-608). This study was restricted to 138 subsamples of spoken data for detailed analysis.

The sample consisted of 69 males and 69 females (see Table 1). Of the total, language majors accounted for the largest number (70), followed by engineering (26) and social science (25). The reason there were many language majors was that the task advertised on campus for the compilation of the corpus was recording English speaking. Many language majors were more confident about speaking English than other majors who applied. Students' proficiency in English speaking was assessed by three native English linguistics experts, with a low proficiency group (hereafter, LP) of 89 and high proficiency group (hereafter, HP) of 49. In addition, if the proficiency assessments did not match, they were asked to re-evaluate to converge their opinions.

TABLE 1. Students' information in the sample

Gender		Major					Proficiency	
Male	Female	Science	Engineering	Sociology	English	Other languages	Low	High
69	69	17	26	25	56	14	89	49
50 %	50 %	12.3 %	18.8 %	18.0 %	40.6 %	10.1 %	64.5 %	35.5 %
Total: 138 (100 %)								

All the 138 college students who participated in the data collection performed a 2-minute English monologue task. They chose one of the four topics offered, which were topics on everyday issues that anyone could answer in English without difficulty. The task was performed in a soundproof lab and stored in real time. Hence, to ensure that the results were comparable, both constraints of time as well as topic were controlled (Wolfe-Quintero et al., 1998). The following were the subjects of the monologues:

- What do you usually do in your free time? Hobbies, etc.
- What is your favorite movie genre?
- Do you think there can be true friendship between opposite genders?
- Is it better to have a dog than a cat?

## MEASURES

### SYNTACTIC COMPLEXITY INDICES

Recognising the importance of measuring syntactic complexity as a multidimensional construct to analyse the complexity of the transcribed spoken corpus, this current study used the computational tool, L2 Syntactic Complexity Analyser (L2SCA; Lu, 2010, pp. 474-96) to gauge the full set of 14 measures presented in L2SCA. The five dimensions of syntactic complexity, that is, the length of the production unit, the amount of subordination, the amount of coordination, degree of phrasal sophistication and overall sentence complexity are tested by the measures, with each measure testing one of them (Ai & Lu, 2013; Lu, 2010; 2011; Norris

& Ortega, 2009; Wolfe-Quintero et al., 1998). Table 2 presents the definitions of the different production units and syntactic structures that are involved in computing the measures adapted from Lu (2010; 2011) and Lu and Ai (2015, pp. 16–27).

The first set of measures gauged complexity in terms of the mean length of production. This is an overall production length measure defined as the length of production at the clausal, sentential or T-unit level, namely, the mean length of clause (MLC), the mean length of sentence (MLS) and the mean length of T-unit (MLT). The second type measured sentence complexity or clauses per sentence (CS). The third type reflected the amount of subordination, including clauses per T-unit (CT), complex T-units per T-unit (CTT), dependent clauses per clause (DCC) and dependent clauses per T-unit (DCT). The fourth set of measures gauged the amount of coordination, namely, coordinate phrases per clause (CPC), coordinate phrases per T-unit (CPT) and T-units per sentence (TS). The last type considered a particular structure in relation to a larger production unit and consisted of three ratios, including complex nominals per clause (CVC), complex nominals per T-unit (CNT) and verb phrases per T-unit (VPT). Therefore, the syntactic complexity measures in this study were targeted not only at sentence complexity, but also phrasal and clausal levels and length measures (i.e. the mean length of a sentence), as well as sophistication measures (i.e. coordinate clause ratios, dependent clause ratios and complex nominals).

TABLE 2. The 14 syntactic automated complexity measures (Lu 2010)

<i>Length of production unit</i>	
	Mean length of clause (MLC: words/clause)
	Mean length of sentence (MLS: words/sentence)
	Mean length of T-unit (MLT: words/T-unit)
<i>Sentence complexity</i>	
	Clauses per sentence (CS: clauses/sentence)
<i>Subordination</i>	
	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)
<i>Coordination</i>	
	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)
<i>Particular structures</i>	
	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)

#### ASSESSMENT OF L2 LEARNERS' SPEAKING QUALITY

To determine learners' L2 speaking levels, three native English linguistics experts were recruited and asked to evaluate learners' L2 speaking levels clearly and objectively based on the Common European Framework of Reference for Languages (CEFR), which has been recognised as a standard for L2 assessment throughout Europe since 2001 and has gradually expanded in use worldwide (Glover, 2011, pp. 121–133). The CEFR divides the level of proficiency into six levels, A1, A2, B1, B2, C1 and C2, which are further subdivided using a traditional classification system that divides proficiency into beginner, intermediate and advanced levels (Figure 1). In addition, if the proficiency assessments by the three native experts did not match, they were asked to re-evaluate to converge their opinions. The six proficiency levels were eventually divided into two groups for binary logistic regression (i.e. LP: A1, A2, B1; HP: B2, C1, C2) as presented in the next section.

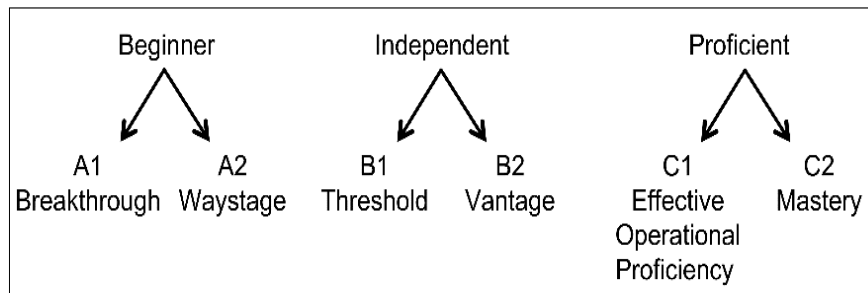


FIGURE 1. The vertical dimension (Council of Europe 2001: 217–25)

The most widely used statistical methods for analysing categorical variables are linear discriminant analysis and logistic regression. Both are appropriate in terms of building linear classification models, but linear discriminant analysis makes more assumptions about the data. In other words, if the populations are normal with identical covariance matrices, researchers prefer discriminant analysis to logistic regression in solving discriminant analysis problems. However, logistic regression is appropriate if these assumptions are violated (DeCoster & Claypool, 2004). Therefore, it was essential to examine the data first to determine if the assumptions of normality were satisfied.

## RESULTS

### DISCRIMINANT ANALYSIS

To determine the values of the univariate ANOVA's Box's M in order to investigate whether the essential assumptions, that is, normality assumptions, were satisfied, a discriminant analysis was performed. Table 3 presents the results of Box's test of equality of covariance matrices. A significance value of .000 would indicate that the data differed significantly from multivariate normal. This would mean that the study could not proceed with the analysis because normality assumptions were violated. With non-normality as in this case, the logistic regression method is used instead to analyse categorical outcome variables (Park, 2020; Sio & Ismail, 2019, p. 29).

TABLE 3. Box's test of equality of covariance matrices

Box's M		549.221
F	Approx.	5.347
	df1	91
	df2	32003.643
	Sig.	.000

### LOGISTIC REGRESSION

In this current study, the goal of logistic regression was to search for the best fitting and most parsimonious model that describes the relationship between the outcome (i.e. proficiency in English speech) and a set of independent (i.e. complexity measures) variables (Pohar et al., 2004). The dependent variable for English proficiency was categorised into two groups: LP and HP. This investigates the predictive validity of independent variables to confirm the predictive quality of complexity indices. The regression analysis was conducted to determine whether the selected indices among 14 complexity variables could predict L2 proficiency in



the spoken data. In this study, regression analysis indicated that the significant independent variables in the model were as follows: MLS, MLT, CPT and CPC ( $p < .05$ ).

Table 4 presents the overall test of the model using the likelihood ratio. The model was statistically significant,  $\chi^2(4) = 42.359$  ( $p < .05$ ), indicating that the model not only identified the sub-scales that influenced L2 proficiency but also demonstrated an adequate fit of data to the model.

TABLE 4. Omnibus test of model coefficients

		Chi-square	df	Sig.
Step 1	Step	42.359	4	.000
	Block	42.359	4	.000
	Model	42.359	4	.000

Table 5 presents the two methods (i.e. Cox & Snell  $R^2$  and Nagelkerke  $R^2$ ) to calculate the explained variation. The explained variation in the dependent variables in this model was 36.3 % according to Nagelkerke  $R^2$  values (Nagelkerke  $R^2 = .363$ ).

TABLE 5. Model summary

Step	-2 Log likelihood	Cox & Snell R-square	Nagelkerke R-square
1	137.187	.264	.363

Logistic regression estimates the probability of the event occurring, that is, the observed number of students in each proficiency group and the predicted number according to logistic regression ( $p < .05$ ). Table 6 presents the assessment of the effectiveness of the predicted classification as opposed to the actual classification. It indicates that 87.6 % of the total 89 cases in the lower group, or 78 cases, were correctly classified into the lower group, and 53.1 % of the total 49 cases in the upper group, or 26 cases, were correctly classified into the upper group. The overall classification accuracy was 75.4 %.

TABLE 6. Category prediction

	Observed	Predicted		
		L2 Proficiency		Correct %
		Low	High	
Step 1	Low	78	11	87.6
	High	23	26	53.1
Overall Percentage				75.4

Table 7 presents a summary of the logistic regression results used to determine the influence of predictors on L2 proficiency in a spoken corpus. First, when considering the odds ratio (i.e.  $\text{Exp}(B)$ ), it shows that the strongest predictor of L2 proficiency is CPC with an odds ratio of 24. This indicated that L2 learners were over 24 times more likely to achieve high proficiency when they acquired CPC. Second, considering the level of significance, learners' English speaking proficiency showed significant differences in MLT, CPT and CPC, and no significant differences were found in MLS. Third, it showed the coefficients and the Wald values, which are used to determine the significance of each of the independent variables. MLT (Wald = 8.146), CPT (Wald = 8.698) and CPC (Wald = 4.207), excluding MLS (Wald = 0.268,  $p > 0.05$ ), were statistically significant ( $p < .05$ ). Finally, the estimated logistic regression equation for English proficiency levels was derived as follows:

Log (proficiency level):  $-4.857 + 0.045(\text{MLS}) + 0.433(\text{MLT}) - 10.135(\text{CPT}) + 7.795(\text{CPC})$

Using this equation, two groups of high and low proficiency levels were divided by the boundary of 0.5. In other words, when the values of learners' MLS, MLT, CPT and CPC were entered into the derived logistic regression equation, values of 0.5 or less were allocated to the low proficiency group, and those scoring higher entered the high proficiency group.

TABLE 7. Variables in the equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 1	MLS	.045	.268	1	.604	1.046
	MLT	.433	8.146	1	.004	1.541
	CPT	-10.135	8.698	1	.003	.000
	CPC	7.795	4.207	1	.040	24.293
	Constant	-4.857	20.426	1	.000	.008

## DISCUSSION AND CONCLUSION

This study investigated the syntactic complexity of EFL undergraduates' spoken corpus and examined the predictors of complexity that lead to high L2 proficiency. In total, 14 measures were used to gauge complexity at the sentential, the clausal, the phrasal and the nominal levels of syntactic organisation. After the normality assumption analysis confirmed that logistic regression was more appropriate than linear discriminant analysis, the analysis was performed to ascertain the effects of syntactic complexity measures on the likelihood of participants' classification as high proficiency. The logistic regression model was statistically significant,  $\chi^2(4) = 42.359$ ,  $p < .05$ , and explained 36.3 % of the variance in classification according to L2 proficiency and correctly classified 75.4 % of cases. Based on the dataset analysed here, the complexity measures that were chosen, that is, MLS, MLT, CPT and CPC, predicted L2 speech proficiency. The results also indicated that learners were over 24 times more likely to achieve high proficiency in English speaking when they focus on improving their CPC grammar ability. That is, CPC means coordinate phrases per clause, and coordinate phrase is counted as the sum of the number of coordinate adjective, adverb, noun, and verb phrases (Ai & Lu, 2013). To use coordinate conjunctions to group two or more words into a single unit, learners must know that the target words are grammatically equivalent, meaning that they have achieved a level of understanding the words' parts of speech. Finally, an effective equation was proposed to help educators classify EFL learners according to their proficiency in L2 speech after measuring the selected dimensions of complexity.

Some recent research into syntactic complexity using spoken data are partially supported with this finding (Iwashita, 2006; Iwashita et al., 2008; Neary-Sundquist, 2017). In terms of mean length of production (i.e. MLS and MLT), Iwashita (2006) and Iwashita et al. (2008) have showed that length-based complexity features are good predictors of speaking ability. Iwashita et al. (2008) have examined language testing data collected from ESL learners performing tasks where they had to express opinions or recount information and discovered that only mean length of utterance among CT, dependent clause ratio, verb phrase ratio, and mean length of utterance had the expected positive correlation with proficiency level. Furthermore, the results of this study, in which the amount of coordination (i.e. CPC and CPT) is a strong predictor of learner speaking proficiency, are partially consistent with that of Neary-Sundquist (2017), who argues that all indices of the amount of subordination and coordination and the phrasal sophistication show different patterns of use according to proficiency level.

This study's findings have some implications for L2 speaking pedagogy. First, the importance of the selected measures of syntactic complexity should be noteworthy to teachers who are engaged in developing an effective grammar curriculum. Specifically, the present results point to length-based complexity features such as MLT and coordinated phrases such as CPT and CPC as predictors of English speaking proficiency. Awareness of and attention to findings like these could help L2 teachers understand that these dimensions of syntactic complexity should be taught thoroughly and may need more emphasis during appropriate activities and speaking tasks. Second, an effective equation has been proposed to help educators classify EFL learners according to their proficiency in L2 speech by measuring four complexity dimensions only. This equation will help teachers determine learners' L2 proficiency, especially when there is no official English test score or when conducting speaking tasks is not possible due to limited time, place or insufficient budget (Park, 2020).

Before summing up, it is necessary to recognise the limitations of the current research that may be addressed in future studies. First, this study used samples of monologues for analysis. Future studies will benefit from diversifying genres in examining spontaneous conversations and essay-writing with the same analytic process. Second, while this study gauged a set of 14 complexity measures, though it may be sufficient on a complexity scale, future research will be more beneficial when it extends to the analysis of accuracy and fluency areas. Finally, to minimise the problem with counting units in speech data, this work deleted repetitions, false starts and fillers from transcriptions; however, it is worth considering other segmentation units presented by some researchers for speech data, such as Analysis of Speech Unit (AS-unit), Conversation Unit (C-unit) and Utterance Unit (U-unit) (Crookes, 1990, pp. 184–190; Foster et al., 2000, pp. 354–375; Lintunen & Mäkilä, 2014, pp. 377–399). Although there are limitations in the scope and segmentation methods, this study offered valuable insights into a new research methodology that sought to extend the body of research on L2 assessment by providing a mathematical way to distinguish L2 proficiency levels by comparing the degree of complexity in a spoken corpus and the holistic English speaking quality. In addition, the study suggests that the amount of coordination per a clause is one of the strongest indicators of L2 fluency, which can be used as a measure for assessing oral fluency and proficiency in L2 speech in EFL classrooms.

#### ACKNOWLEDGEMENT

*The author would like to thank the anonymous reviewers. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (No. 2020S1A5B5A17090102).*

#### REFERENCES

- Ai, H. & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In A. Diaz-Negrillo, N. Ballier & P. Thompson (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 249–64). John Benjamins Publishing Company.
- Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly*, 26(2), 390–5. <https://www.jstor.org/stable/3587016>
- 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. <https://doi.org/10.5054/tq.2011.244483>
- Bulté, B. & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65. <https://doi.org/10.1016/j.jslw.2014.09.005>
- Chen, M. & Zechner, K. (2011). *Computing and evaluating syntactic complexity features for automated scoring of spontaneous non-native speech* [Paper presentation]. 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (pp. 722–31).

- Council of Europe. (2001). *The common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.
- Crookes, G. (1990). The utterance, and other basic units for second language discourse analysis. *Applied Linguistics*, 11(2), 183–99. <https://doi.org/10.1093/applin/11.2.183>
- DeCoster, J. & Claypool, H. (2004). *Data analysis in SPSS*. Retrieved October 2, 2015, from [https://www.academia.edu/15281435/Data\\_analysis\\_in\\_SPSS](https://www.academia.edu/15281435/Data_analysis_in_SPSS)
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21(3), 354–75. <https://doi.org/10.1093/applin/21.3.354>
- Glover, P. (2011). Using CEFR level descriptors to raise university students' awareness of their speaking skills. *Language Awareness*, 20(2), 121–33. <https://doi.org/10.1080/09658416.2011.555556>
- Halliday, M. A. K. (2006). *The language of science*. Bloomsbury Publishing.
- Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. National Council of Teachers of English.
- Iwashita, N., Brown, A., McNamara, T. & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics*, 29(1), 24–49. <https://doi.org/10.1093/applin/amm017>
- Iwashita, N. (2006). Syntactic complexity measures and their relation to oral proficiency in Japanese as a foreign language. *Language Assessment Quarterly*, 3(2), 151–69. [https://doi.org/10.1207/s15434311laq0302\\_4](https://doi.org/10.1207/s15434311laq0302_4)
- Kim, J-Y. (2014). Predicting L2 writing proficiency using linguistic complexity measures: A corpus-based study. *English Teaching*, 69(4), 27–51.
- Kyle, K. (2016). *Measuring Syntactic Development in L2 Writing: Fine Grained Indices of Syntactic Complexity and Usage-based Indices of Syntactic Sophistication* [Unpublished doctoral dissertation]. Georgia State University.
- 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), 333–49. <https://doi.org/10.1111/modl.12468>
- Lintunen, P. & Mäkilä, M. (2014). Measuring syntactic complexity in spoken and written learner language: comparing the incomparable? *Research in Language*, 12(4), 377–99.
- Lorenzo, F. & Rodríguez, L. (2014). Onset and expansion of L2 cognitive academic language proficiency in bilingual settings: CALP in CLIL. *System*, 47, 64–72. <https://doi.org/10.1016/j.system.2014.09.016>
- 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. <https://doi.org/10.1016/j.jslw.2015.06.003>
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International journal of corpus linguistics*, 15(4), 474–96. <https://doi.org/10.1075/ijcl.15.4.02lu>
- 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>
- Neary-Sundquist, C. A. (2017). Syntactic complexity at multiple proficiency levels of L2 German speech. *International Journal of Applied Linguistics*, 27(1), 242–62. <https://doi.org/10.1111/ijal.12128>
- Norris, J. & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–78. <https://doi.org/10.1093/applin/amp044>
- Ortega, L. (2015). Syntactic complexity in L2 writing: Progress and expansion. *Journal of Second Language Writing*, 29, 82–94. <https://doi.org/10.1016/j.jslw.2015.06.008>
- Park, S. (2020). Determining L2 Proficiency in the production and perception of consonant clusters. *3L: Language, Linguistics, Literature*, 26(3). <http://doi.org/10.17576/3L-2020-2603-04>
- Park, S. & Yoon, S. (2021). Syntactic Complexity of EFL Learners' Casual Conversation, Monologue, and Writing. *The Journal of Studies of Linguistics*, 37(1), 75-89. <https://doi.org/10.18627/jslg.37.1.202105.075>
- Pohar, M., Blas, M., & Turk, S. (2004). Comparison of logistic regression and linear discriminant analysis: a simulation study. *Metodoloski zvezki*, 1(1), 143–61.
- 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), 22-37. <http://doi.org/10.17576/3L-2019-2502-02>
- 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(2), 420–30. <https://www.jstor.org/stable/43267799>
- Wang, S. & Slater, T. (2016). Syntactic complexity of EFL Chinese students' writing. *English Language and Literature Studies*, 6(1), 81–6.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity* (No. 17). University of Hawaii Press.
- 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. <https://doi.org/10.1016/j.jslw.2015.02.002>
- Yoon, S. & Park, S. (2021). Grammatical complexity of EFL learners' casual conversation at different proficiency levels. *Korean Journal of English Language and Linguistics*, 21, 599-616. <http://doi.org/10.15738/kjell.21..202107.599>