

## A Corpus-Based Comparison of Syntactic Complexity in Spoken and Written Learner Language

Shinjae Park  
*Sehan University*

### Abstract

Despite writing and speaking being related activities, their end-products are entirely different. However, previous studies have not shown consistency in terms of grammar use in these two modes. Accordingly, in the present study, I aim to define the syntactic characteristics in these two modes with large-scale data and organized research designs. This study examined 14 indices of syntactic complexity and specific grammar factors in 224 monologues and 139 writings of Korean EFL undergraduates. The results revealed that learners tended to use more finite complement clauses and relative clauses while writing but used because-fragments independently and ‘and’ sentence-initially more frequently while speaking. When compared with previous studies, the characteristics of syntactic complexity of Korean EFL learners, regardless of age, are defined by the use of coordination in speaking and the use of subordination in writing.

### Résumé

L’écrit et l’oral sont des activités clairement liées, mais le résultat final est tout à fait différent. Toutefois, des études antérieures n’ont pas montrées de cohérence dans l’utilisation de la grammaire dans les deux modes. Par conséquent, dans la présente étude, le but est de définir les caractéristiques syntaxiques des deux modes avec des données à grande échelle et des plans de recherche organisés. Cette étude a examiné 14 indices de complexité syntaxique et des facteurs grammaticaux spécifiques dans 224 monologues et 139 écrits d’étudiants coréens de premier cycle EFL. Les résultats ont révélé que les apprenants ont tendance à utiliser des clauses complémentaires limitées et des clauses relatives lorsqu’ils écrivent, mais qu’ils utilisent les fragments ‘parce que’ de manière indépendante et les fragments ‘et’ en début de phrase plus fréquemment à l’oral. En comparaison des études précédentes, les caractéristiques de la complexité syntaxique des apprenants coréens de l’EFL, quel que soit leur âge, sont définies par l’utilisation de la conjonction de coordination dans la parole à l’oral et de la conjonction de subordination par à l’écrit.

## **A Corpus-Based Comparison of Syntactic Complexity in Spoken and Written Learner Language**

Syntactic complexity can be generally construed as the variety and degree of sophistication of the syntactic structures deployed in written production (Bulté & Housen, 2014; Lu, 2011; Ortega, 2003) and has been widely adopted as a reliable measure for second language (L2) writing proficiency. It is often used as an index of language proficiency and development status of L2 learners. Various studies have proposed and investigated measures of syntactic complexity, as well as examined whether it could serve as a reliable predictor of language proficiency. For instance, based on the holistic ratings of the essays from secondary-level writers of varying levels of proficiency, Martínez (2018) demonstrated a significant link between syntactic complexity and writing quality. Specifically, the author reported that the use of longer units on the clausal and sentential level was a strong indicator of high-quality writing; in contrast, the frequent use of simple sentences was found to be associated with lower writing quality. Similarly, previous research using syntactic complexity has focused mostly on written data (Barker et al., 2015; Myles, 2015).

On the other hand, the nature of the relationship between spoken and written language has been an interesting subject to linguists, psychologists, and educators for decades. Despite differences in focus, scholars agree that the end-product is entirely different: while speaking involves producing sounds, writing involves producing marks on a page. However, the same set of grammatical and lexical features seem to be acceptable in written or spoken language. For example, Cleland and Pickering (2006) found that a group of UK undergraduates tended to repeat syntactic form between modalities (from speaking to writing and writing to speaking) to the same extent that they did within either modality. The authors demonstrated that syntactic priming<sup>1</sup> is unaffected by whether prime and target sentences are produced in similar or different modalities and concluded that syntax is accessed in the same way in both spoken and written production (Cleland & Pickering, 2006). This might suggest that the underlying mechanisms are shared, and it is only the output that differs. In this context, it needs to be established which grammatical characteristics are shared or represented differently by learners in the two modes (Hwang et al., 2020; Park & Yoon, 2021).

To date, few studies have compared syntactic complexity of written and spoken L2 productions with a focus on EFL learners (Hwang et al., 2020; Kormos, 2014; Park & Yoon, 2021). Although these studies identified quantitative and qualitative differences between written and spoken data along various syntactic complexity indices reflecting the distinct syntactic features of the two modalities, the results are inconsistent. For example, Hwang et al. showed that learners used longer sentences, more subordination, more verb phrases per T-unit, and less coordination in writing than in speaking. On the other hand, Park and Yoon showed that syntactic complexity did not significantly differ between monologues and writing of 40 Korean EFL learners, except for complex nominals per clause.

In summary, until recently, studies on L2 learners' syntactic characteristics focused more on writing modes, so few studies have compared the syntactic characteristics across written and spoken data supplied by learners of English as a foreign language (EFL). Moreover, the results of the few relevant studies were not consistent. Therefore, in order to fill in these gaps, in the present study, I will analyze in detail the syntactic complexity of

Korean learners' English monologues and writing using a large-scale dataset. My aim is to capture and compare the grammatical characteristics represented differently depending on the production modes by measuring syntactic complexity from various dimensions through L2SCA (L2 Syntactic Complexity Analyzer) presented by Lu (2011) using 244 monologues produced for 2 minutes and 139 essays for 30 minutes by Korean EFL learners. Furthermore, detailed analysis is performed by selecting important grammatical elements within the factors that have been found to have significant differences in global complexity measures in spoken and written data. Furthermore, the results of this study will be compared with those reported in previous studies on syntactic complexity, which will increase the reliability of clarifying the characteristics of syntactic complexity in monologues and writings produced by Korean learners.

## Background

Numerous studies in the past have tended to focus on grammatical structures and their usage of syntactic complexity to see language proficiency and development status of L2 learners. In fact, syntactic complexity has been recognized as an important construct in L2 writing teaching and research, as the growth of syntactic repertoire is an integral part of a learner's development in the target language (Ortega, 2003; Lu, 2011). Most studies have relied on quantifiable complexity indices such as sentence complexity, length of production unit, and frequency of specific sentence structures. Of these, the concept of the T-unit (Hunt, 1965) is defined as the shortest grammatical chunk of a sentence as a unit of analysis. Various studies have proposed and investigated measures of syntactic complexity and examined whether they serve as a predictor of language proficiency (Ai & Lu, 2013; Jiang et al., 2019; Khushik & Huhta, 2020; Lan & Sun, 2019; Lu, 2011; Martínez, 2018). Regarding the syntactic complexity indices, the present study follows Lu (2010, 2011) and others' recommendations (Biber et al., 2016; Hwang et al., 2020; Kyle, 2016; Kyle & Crossley, 2017) to examine syntactic complexity as a global dimension. Table 1 lists the 14 indices of syntactic complexity adopted from Lu (2011). The indices consist of five sets of measures to represent "a different but interrelated aspect of complexity" (Bulté & Housen, 2014, p. 47). They also show the methods of calculation of syntactic factors: length of production, sentence complexity, subordination, coordination, and particular structures.

**Table 1***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

Among recent research using such indices, a recent study on English argumentative essays written by 868 Pakistanis and 287 Finish teenagers, Khushik and Huhta (2020) found that the length of production units, subordination, and phrasal density differed according to proficiency level. Furthermore, in a study that examined the relationship between syntactic complexity and writing quality in research papers produced by 280 ESL undergraduates, Casal and Lee (2019) found that phrasal measures and mean length of T-units differed across levels. Likewise, Lan and Sun (2019) compared the arguments of Chinese English learners to academic journal articles in terms of the use of noun modifiers, as well as examined the correlation between the use of noun modifiers and students' writing proficiency measured by their TOEFL writing scores. The results of this study revealed that the frequency of noun modifiers in students' writings was much less than their use in academic journal articles.

On the other hand, available studies on the subject of syntactic complexity using spoken data are scarce. Commenting on this lack of research on the analysis of spoken data, Chen and Zechner (2011) and Park and Yoon (2021) said that using spoken data is much more difficult because researchers need to complete very complicated pre-work, including transcribing and cleaning disfluencies such as false starts, repetitions, filled pauses, and so forth. However, some studies have attempted to characterize the syntax complexity used in speaking and writing (Biber, et al., 2011; Hwang et al., 2020; Kormos, 2014; Park and Yoon, 2021). Specifically, in an empirical analysis to identify syntactic characteristics of writing by comparing the use of complexity in conversation, Biber et al. found that clausal complexity is a characteristic of speaking, rather than writing and

phrasal complexity is characterized more so in writing. In another study on the effect of product modes on linguistic performance with Hungarian learners of English, Kormos (2014) found that the learners' written productions contained significantly more modifiers per noun phrase than their spoken productions. In addition, the results of Hwang et al.'s (2020) corpus-based analysis of syntactic complexity with written and spoken data provided by 122 beginning-level Korean EFL children revealed that, among the seven syntactic complexity indices, four (MLS, DCT, CPT, VPT) differed significantly between written and spoken production. The written data included longer structures (MLT), more subordination (DCT), and more verb phrases (VPT) than the spoken data, whereas the spoken data involved a greater amount of coordination (CPT) than the written data. Furthermore, in Park and Yoon's (2021) comparison of three production modes i.e., conversation with two or more people, monologue, and essays of 40 Korean learners of English, both monologue and writing modes were found to elicit significantly greater syntactic complexity than conversation in all indices, whereas there was no significant difference in the use of complex structures in monologue and writing, except for CNC. The authors inferred that setting, such as the procedure set during the tasks, may have affected the results, focusing on the execution of the process whereas the two modes were performed. Actually, in the corpus collection, when given several everyday topics, the participants chose the topics of their own interest. Even if they were supposed to respond to the topics as immediately as possible, they still spent from a few seconds to a few minutes planning while choosing the topic before their actual tasks.

In this regard, an interesting possibility is that some components of writing and speaking may be shared, while others may be distinct. In models of spoken production, it is normally assumed that language production involves various stages, with a fundamental division into conceptualization, formulation, and articulation (Levelt, 1989). In written production, some researchers assume that Levelt's account for speaking is applicable to writing, suggesting that both involve stages for planning of contents, linguistic encoding, execution or articulation, and monitoring (Kellogg, 1996; Levelt, 1989). For instance, Bonin et al. (1998) tentatively concluded that some syntactic and semantic information would be shared between modalities in word production. However, despite such correspondences, the cognitive processes of the two modalities guide learners differently (Hwang et al., 2020). Additionally, Ravid and Tolchinsky (2002) pointed out that writing and speaking also differ with respect to context dependency. Written production is less dependent on context, which allows writers a higher degree of control over the product. Furthermore, the fact that the text as a whole is visually accessible throughout the writing process helps writers to closely attend to linguistic forms (Niu, 2009).

The present study aims to capture the characteristics of spoken and written data produced by Korean EFL learners by conducting a carefully-designed methodology and by comparing these results with previous influential studies. To this end, I measure syntactic features using 14 indices across five dimensions presented by Lu (2011) discussed above (see Table 1). The 14 measurements adopted by Lu (2010; 2011) meet the criteria of this study—namely, several measurements should be reviewed to reflect the syntax complexity of various aspects. In addition, a detailed analysis is performed on certain grammatical elements within the indices that have been found to significantly differ in syntactic complexity in spoken and written data. Through the use of not only the holistic measures (i.e., multifaceted indices of syntactic complexity), but also the specific measures (i.e., the detailed grammar factors, and combined with comparing the results with those reported in

influential previous research), this study aims to clarify the syntactic characteristics of L2 speaking and writing. The specific research questions addressed in this study are as follows:

RQ1. Are there differences between monologue and writing of Korean EFL undergraduates in terms of syntactic complexity? If so, in which aspects and to what extent do the two modes differ?

RQ2. If the indices derived from RQ1 are analyzed in more detail, what are the characteristics of spoken and written production?

RQ3. How do the results on the syntactic complexity of Korean learners' spoken and written production compare to findings reported in previous studies?

## Method

### Corpus description

The data analyzed in the present study included 139 writings and 224 monologues in the *Multi-Language Learner Corpus* (hereafter, MULC) of Korean university students (Park and Yoon, 2021). The participants could choose and participate in one or both of the tasks and also choose one of the four daily topics for each task. In fact, the time it took for a learner to select a topic was never in excess of two minutes per task. They weren't given time to plan beforehand, but at the same time, it means they could afford to do so if even for a little while (Park and Yoon, 2021). The writing task was assigned 30 minutes, and the monologue task was assigned 2 minutes. Writing was conducted using a Note application on a desktop computer so that an online dictionary would not be used, whereas monologues were conducted in a soundproof lab, and all data were recorded digitally in real time under the present author's supervision. The topics are shown in Table 2. The collected monologue recordings were manually transcribed by dozens of trained researchers and finally confirmed by English native linguistic experts. Furthermore, prior to the actual evaluation, the linguistic experts went through a pilot test process for 5% of the data, and all discrepancies in data evaluation results were solved through discussion.

To determine the 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 recognized as a standard for L2 language progression throughout Europe since 2001 and has gradually expanded in its use worldwide (Glover 2011; Hulstijn 2007). The CEFR is divided into six levels of proficiency, A1, A2, B1, B2, C1 and C2, which are further subdivided using a traditional classification system that separates proficiency into beginner (A1, A2), intermediate (B1, B2) and advanced (C1, C2) levels (i.e., from A1, the lowest to C2, the highest level). During the evaluation period, the evaluators including the present author, held weekly meetings to apply the above-mentioned evaluation criteria. If an agreement could not be reached by the native experts, a reevaluation was conducted until agreement was reached.

In addition, a one-way ANOVA was conducted to analyze the impacts on proficiency scores (from A1 to C2, i.e., 1 to 6 points) due to differing topics, and as a result, there was no significant difference in the response variable per topic (monologue:  $F(3,240) = .116$ ,  $p = .951$ ; Writing:  $F(3,135) = 2.121$ ,  $p = .102$ ). In other words, the four distinct topics within each task did not lead to a statistically significant difference in L2 proficiency.

**Table 2***Topics Provided in Each Task*

Monologue	Writing
1. What do you usually do in your free time? Hobby, etc.	1. Should everyone get married?
2. What is your favourite genre of movies?	2. Is it essential to wear school uniforms in middle and high schools?
3. Do you think there can be friendship between opposite genders?	3. Should elementary, middle, and high school students be allowed to carry phones in class?
4. Is it better to have a dog than a cat?	4. Should any college student join a club?

Table 3 represents information about the students who participated in the monologues task. Most students majored in English (78, 35%), followed by engineering colleges (42, 19%), natural science colleges (24, 11%), social science colleges (23, 10%), and other foreign language majors (22, 10%). The reason why most students majored in English was that the data were collected by conducting public advertisements, and predominantly those students who were relatively confident in their English production volunteered to participate. The sample had also a balanced gender distribution; the mean age of the participants was 20.9 years old (SD: 1.951).

The English-speaking proficiency of the participants as measured by the Common European Framework of Reference for Languages (hereafter, CEFR) standard was 2.81 (SD=0.870), which is close to B1 (based on a 6-point scale with A1=1 and C2=6; A1, A2, B1, B2, C1, and C2, from the lowest to the highest level). None of the participants had C2 proficiency, which is a native speaker's level. In particular, a large number of students were in the mid- and low-level proficiency groups, i.e., B1 and A2 (B1: 86 (38.6%); A2: 82 (36.8%), followed by B2 (42, 18.8%).

**Table 3***Participant Information*

Majors								
Natural Science	Business Admin.	Engineering	Education	Law	Social Science	English	Other Languages	Arts & Physics
24	15	42	9	8	23	78	22	3
11%	7%	19%	4%	3%	10%	35%	10%	1%
Male: 122, Female: 122								
Age: 20.9								
A1 (7, 3%), A2 (82, 37%), B1 (86, 38%), B2 (42, 19%), C1 (7, 3%), C2 (0, 0%)								
Total: 224 (100 %)								

**Analytic procedures and statistical analysis**

This study analyzed both modes of speaking and writing, so reliable measuring related to syntactic complexity was an important issue. As a basic grammatical unit, a simple clause is a unit with a subject, a finite verb (Lu, 2010), and an optional object or complement; in addition, the T-unit, i.e. a unit that consists of one main clause and (optional) subordinate clauses and non-clausal units or sentence fragments attached to it (Hunt, 1965), was used as an omnibus measure of grammatical complexity of student

writing development (Hunt, 1965). Additions or modifications to these patterns result in complex grammar, with the implicit understanding that more additions result in more complexity (Biber et al., 2011).

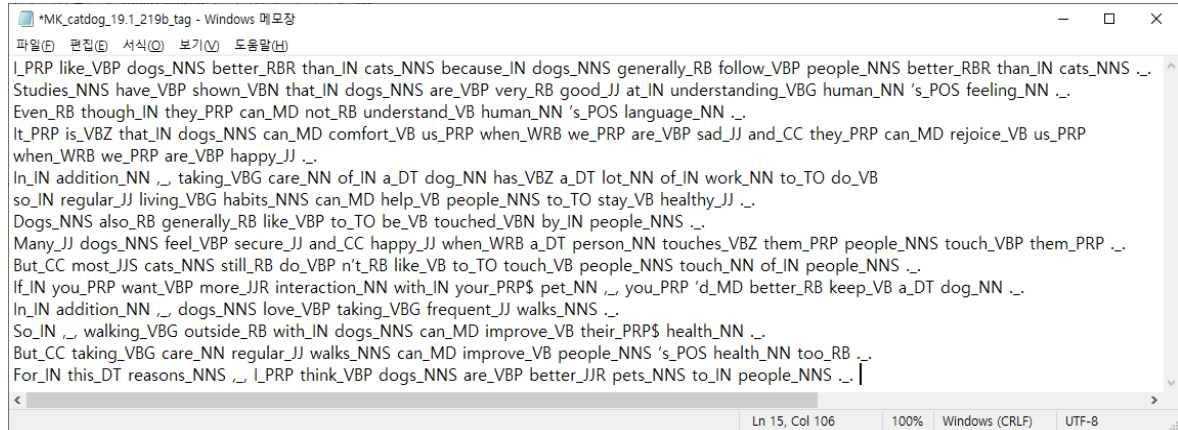
However, since the present study focuses on comparing spoken and written production, the application of the T-unit concept is usually done on written work and would require a lot of time-consuming labour on spoken data. (Litunen & Mäkillä, 2014). Accordingly, the present study relies on sentence segmentation suggested by Litunen and Mäkillä, Foster et al. (2011) and Nippold et al. (2017) —namely that the one sentence contains coordinated clauses. Another unit boundary criterion suggested by Litunen and Mäkillä was the duration of 1.5 seconds. The transcription used in this study marks a pause longer than 1.0 seconds by number, and the sentence ending punctuations were used when there is a clear falling intonation or rising intonation. Therefore, applying these criteria to the current corpus was possible, and this work allowed me to examine the ratio of coordinated structures and the measure of sentence complexity ratio in both modes. In addition, Litunen and Mäkillä asserted that the segmentation unit used in their study might carry spoken language complexity closer to written language complexity, and the unit also may reveal the learner's intended idea in a way that the traditionally used spoken language units may not, as frequent and long pauses in learners' spoken production should not affect the amount of their syntactic complexity use in spoken language. Furthermore, in spoken data, how to deal with false starts, repetitions, and self-corrections is a very important matter (Foster et al., 2000). Therefore, I eliminated the disfluencies (Chen & Zechner, 2011; Lu, 2012); however, following the suggestion of Foster et al. (2000), response tokens such as *oh*, and *hmm* were considered as a word.

The spoken and written data were submitted to the Part of Speech (POS) tagging process using the Stanford NLP tagger (Figure 1). A POS Tagger is a piece of software that reads text and assigns parts of speech to each word (and other tokens), such as noun, verb, adjective, etc. For tagging purposes, five trained researchers including the present author automatically completed the tagging operation using the tagger and manually verified the error of the POS tagged data based on the Vienna Oxford International Corpus of English (VOICE) Tagging Manual (Seidlhofer et al., 2001). In other words, for POS tagging, five researchers conducted a total of two sequential inspection processes. In other words, three junior researchers tagged their assigned share first and based on the results, a meeting was held for a second inspection every week, including two senior researchers one of which is the present author. In addition, the weekly meetings were also intended to minimize discrepancies in the criteria for tagging among researchers, and as the weekly meetings continued, all researchers eventually reached a common set of standards.

Stanford NLP tagger has the advantage of providing most of the common core natural language processing steps, from tokenization through to co-reference resolution (Manning et al., 2014). Furthermore, Tian and Lo's (2015) study discovered that the Stanford POS tagger was one of the best taggers, achieving an accuracy of 83.6%-90.5% on bug reports. Therefore, this process enhanced the reliability and accuracy of the present findings as compared to those reported in previous studies where this process was not used.



**Figure 1**  
*Tagging Instance of MULC (MK\_catdog\_19.1\_tag)*



To address the research questions, firstly, L2SCA (Lu, 2010; 2011) was used to gauge 14 syntax complexity indices presented by Lu (2010, 2011) for global measures in Table 1 using spoken and written data. To find out whether there is a statistically significant difference between the mean scores of the complexity of monologues and writings calculation, the independent-samples t-test was conducted for each index (RQ1). Second, in order to analyze the indicators that show statistically significant differences in spoken and written data in detail, the specific grammar measures found to be different between the two modes in Biber et al.'s (2011) study were analyzed using the Antconc3.5.8. In addition, UCREL's online LL Calculator for computing log-likelihood (LL) values and the Bayes factor was used as the statistical analysis software to compare the results from the two modes. As demonstrated by previous several studies, the log-likelihood test can be used for corpora comparison research and is more reliable than Pearson's chi-squared test (Pojanapunya & Todd, 2018; Rayson & Garside, 2000; Seog, 2018; Seog et al., 2019). Overall, the log-likelihood value is high wherever there is a great variance in frequency. Said differently, high log-likelihood value suggests that a form has a more significant relative frequency difference between the two corpora (Park, 2020; Pojanapunya & Todd, 2018). Finally, the findings were compared to the results reported in previous influential studies (RQ2).

## Results

This study compared the use of syntactic structures between the two modes (writing and speaking) based on 14 syntactic complexity indices and certain grammar factors. The data comprised monologues (n=224) and essays (n=139) produced by Korean college-level students, which included the production data of 40 participants used by Park and Yoon (2021).

### Research Question 1

Table 4 summarizes the mean values of syntactic complexity indices of monologues and essays. The first research question concerned determining whether there is any

significant difference in syntactic complexity between the spoken data and the written data, and if so, in which aspects. An independent-samples t-test was run to determine whether the mean value of complexity for the two modes significantly differed. Since 14 tests (one per index) were simultaneously run on the same dataset, the p-values, and Bonferroni correction was applied to the *p*-values to avoid spurious positives.

**Table 4**

*Mean complexity values and differences for the spoken and written production*

Measure	Mean		SD		<i>t</i>	<i>p</i>	Effect Size
	Monologue (n=224)	Writing (n=139)	Monologue	Writing			
Mean word (Total)	125.67 (28,149)	234.40 (32,581)	52.528	102.552	-13.309	0.000	1.437
<i>Length of production</i>							
MLS	11.715	14.166	4.007	3.969	-5.687	0.329	0.614
MLT	10.536	12.897	3.064	3.162	-7.048	0.231	0.761
MLC	7.297	7.883	1.488	1.251	-3.868	0.198	0.418
<i>Sentence complexity</i>							
CS	1.630	1.809	0.531	0.486	-3.207	0.496	0.346
<i>Subordination</i>							
CT	1.464	1.647	0.398	0.365	-4.412	0.928	0.476
CTT	0.353	0.460	0.208	0.176	-5.052	0.075	0.546
DCC	0.299	0.360	0.129	0.102	-4.797	0.002*	0.518
DCT	0.474	0.623	0.304	0.287	-4.616	0.644	0.498
<i>Coordination</i>							
TS	1.114	1.097	0.170	0.133	0.965	0.003*	0.104
CPT	0.274	0.267	0.226	0.154	0.323	0.004*	0.035
CPC	0.196	0.165	0.173	0.096	1.931	0.000*	0.209
<i>Particular Structures</i>							
CNT	0.898	1.353	0.444	0.469	-9.298	0.434	1.004
CNC	0.612	0.825	0.261	0.238	-7.807	0.185	0.843
VPT	1.811	2.245	0.512	0.538	-7.706	0.193	0.832

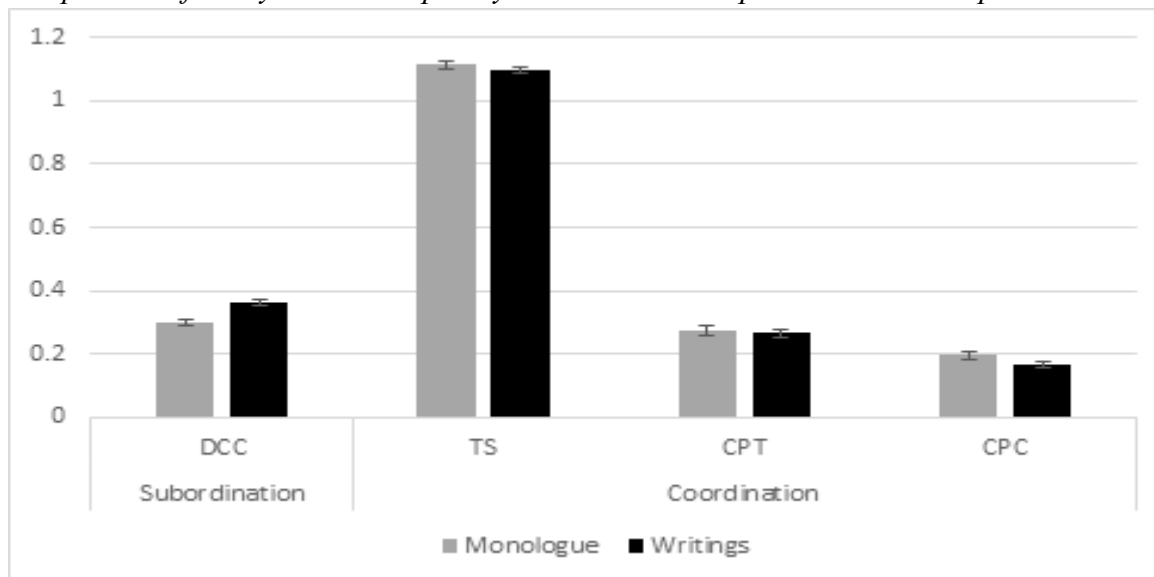
Note: a) P values are after Bonferroni correction, b) The measure of Hedges' *g* is used because sample sizes of each task is different.

As can be seen in Table 4, there were significant differences in DCC in subordination and TS, CPT, and CPC in coordination ( $p < .05$ ). In addition, for 11 of the 14 syntactic complexity indices (i.e., all but TS, CPT, and CPC), the mean value of the spoken data was lower than that of the written data. Of the indices that differed significantly, only DCC was used more complexly in writing, while TS, CPT, and CPC were used more complexly in monologue (see Figure 2). Of note, all three indices that appeared to be more complex in speaking were in the category of the amounts of coordination. Therefore, it is necessary to

pay attention to the coordination that learners use more complexly in speaking than in writing, as well as to the subordination typical of L2 writing. These results are consistent with those reported by Hwang et al. (2020)—namely, that learners used longer sentences, more subordination, and less coordination in writing than in speaking.

**Figure 2**

*Comparison of the syntactic complexity indices between spoken and written production*



On the other hand, to compare the mean differences of the outcome variables across the tasks from each participant, a within-subject design analysis was conducted. The number of learners who performed both tasks was 106, and the results of the within-subject design comparing both modes was slightly different from the results in Table 4. In three of the upper five categories, the results showed more complexity in writing in three categories and, in speaking, more complexity was only present in one category. That is, 9 out of 14 measurements showed significant differences between modes, and 8 of them showed significant complexity in writing. MLS and MLT in Length of Production, DCC, DCT, CTT, and CT in Subordination, VPT and CNT in Particular Structures were found to be more complex in writing. On the other hand, in CPC in Coordination, as with the results of the former analysis, the use of complex structures was more frequent in speaking. In other words, the use of subordination was prominent in writing, while the coordination was more so in speaking, as was the result of former analysis of the entire sample.

Given these results, the reason why the complexity of writing is more prominent is that the difference in the number of words between the tasks among 106 learners is 121.38, (speaking: 109.25 (SD: 47,707); writing: 230.63 (SD: 100.872) and in terms of the entire sample, the difference is 108.73 (125.67 (SD: 23,149), 234.4 (SD: 32,581), respectively). The cause may be found in the number of words per subject when speaking. In other words, it may be inferred that in the within-subject design, more complexity was shown in writing because of the relatively fewer words that were used when speaking.

**Table 5**  
*Within-Subjects Design*

		Paired Differences					t	df	Sig.
		Mean	Std. Deviation	Std. Error mean	95% Confidence Interval of the Difference				
					Lower	Upper			
<b>Length of Production</b>									
MLS	Writing	1.418	4.892	0.475	0.476	2.361	2.985	105	0.004
	- Monologue								
MLT	Writing	1.712	3.944	0.383	0.953	2.472	4.471	105	0.000
	- Monologue								
<b>Subordination</b>									
DCC	Writing	0.064	0.141	0.013	0.037	0.091	4.683	105	0.000
	- Monologue								
DCT	Writing	0.127	0.375	0.036	0.054	0.199	3.481	105	0.001
	- Monologue								
CTT	Writing	0.093	0.249	0.024	0.045	0.141	3.875	105	0.000
	- Monologue								
CT	Writing	0.132	0.549	0.053	0.026	0.238	2.470	105	0.015
	- Monologue								
<b>Coordination</b>									
CPC	Writing	-0.047	0.221	0.021	-0.090	-0.005	-2.194	105	0.030
	- Monologue								
<b>Particular structure</b>									
CNT	Writing	0.329	0.686	0.067	0.197	0.461	4.945	105	0.000
	- Monologue								
VPT	Writing	0.421	0.725	0.070	0.282	0.561	5.981	105	0.000
	- Monologue								
N. of words: Writing (230.63, SD: 109.25), Monologue (100.872, SD: 47.707) t: 11.655 (p=.000)									

## Research Question 2

Having established that there are statistically significant differences in syntactic complexity between the spoken data and written data of Korean undergraduates in coordination and subordination, in this section, I will analyze the results in detail according to each category.

### *Coordination in Spoken Production*

The analysis for RQ1 showed that the use of coordination was statistically more frequent in the spoken production than in the written production. With this in mind, to address RQ2, for statistical significance testing, log-likelihood values were calculated to compare the frequencies of the use of coordinate conjunctions in detail between the two modes (see Table 5). The LL shows a plus or minus symbol before the log-likelihood value to indicate overuse or underuse respectively in Corpus 1 (in the previous column) relative to Corpus 2 (in the next column). The log-likelihood calculations revealed that the Korean EFL learners significantly overused 'and' while beginning utterances as compared to the written task (LL=125.95). The use of coordinating 'and' within a sentence was also more frequently observed in speaking, but there was no significant difference in the use of 'but'.

**Table 6**

Log-likelihood (LL)<sup>2</sup> and *the* Bayes factor results for coordination in monologues and writing

Linguistic feature	Monologue	Writing	LL*	Bayes Factor
And				
Sentence-initial	280	94	+125.95*	114.03*
Sentence-medial	1004	731	+92.32*	81.30*
But				
Sentence-initial	109	93	+4.69	-6.33
Sentence-medial	158	123	+10.99	-0.03

Note: '\*' indicates significant: LL > 15.13 is significant at  $p < 0.0001$  level (also called the 99.99% level) and Bayes Factor > 10: very strong evidence against H0.

The examples of using 'and' when learners begin to construct sentences in monologue are as follows (see Table 6). The first example of 'and' was followed by 'yeah', and the second example of 'and' was followed by 'um', indicating that the learners tend to use 'and' combining other meaningless fillers to have time for constructing each sentence at the beginning of the sentence, which supports the results of Hwang et al. (2020) in that learners use the coordinate syntactic unit using 'and' because they are under more cognitive pressure during the spoken task than the written task.

**Table 7**

*Language Instances of Coordination in a spoken production*<sup>3</sup>

File name	Examples
MK_moviegenre_19.1_227b	<i>And</i> yeah, let's more talk about horror.
MK_catdog_19.1_116b	<i>And</i> um this topic is so difficult but, um I love a cat and I will have a cat someday after I have a job.

### *Subordination in written production*

In this study, the use of dependent clauses in the written production was found to be

statistically significantly higher than in the spoken production. For a detailed analysis of the dependent clause, I analyzed specific grammar factors, i.e., finite dependent clauses that showed statistically differences in the spoken and written production in Biber et al.'s (2011) research. In the aforementioned study, the finite dependent clause was found to serve three major syntactic functions: adverbial, complement, and noun modifier (Biber et al., 2011). Table 7 classifies them by function and presents some examples in the corpus used in the present study.

**Table 8***Functions and Examples of Finite Dependent Clause*

Syntactic function	Examples
Adverbial	At first, I was scared to read in English, but surprisingly, it was not that kind of hard work, <i>because</i> I could imagine scenes while reading. (W_literature_18.2_32) Also, I realized that we must be genuine to others <i>if</i> I want to communicate truly. (W_literature_18.2_169)
Complement	I surely can argue <i>that</i> college student should join at least one club in their campus. (W_club_19.1_118)
Noun modifier	Let's go back to opinion <i>that</i> the most important duty of students is studying. (W_mobile phone_19.1_88)

The log-likelihood calculations in Table 8 revealed that the Korean EFL learners significantly overused *finite complement clauses* and *relative clauses* in writing as compared to the spoken task (LL=21.31 and 85.70, respectively). In contrast, the Korean EFL writers significantly underused *because-clauses* (LL=-37.16) as compared to the spoken production.

**Table 9***Comparison of Dependent Clause by Modes in MULC*

Linguistic feature	Writing	Monologue	LL	Bayes Factor
Finite adverbial clause				
Because-clause	176	272	-37.16*	26.15*
If-clause	188	110	+10.84	-0.17
Although-clause	11	5	+1.51	-9.50
Finite complement clauses				
verb+that clause	171	80	+21.31*	10.30*
Finite noun modifier clauses				
relative clauses	312	100	+85.70*	74.69*

### Research Question 3

I further compared these results with previous studies (i.e., Biber et al., 2011; Hwang et al., 2020; Kormos, 2014; Park and Yoon, 2021) to characterize the syntactic complexity of Korean EFL learners in speaking and writing. To this end, the main features of the previous studies are presented in Table 9.

**Table 9**

*Previous research comparing written and spoken production*

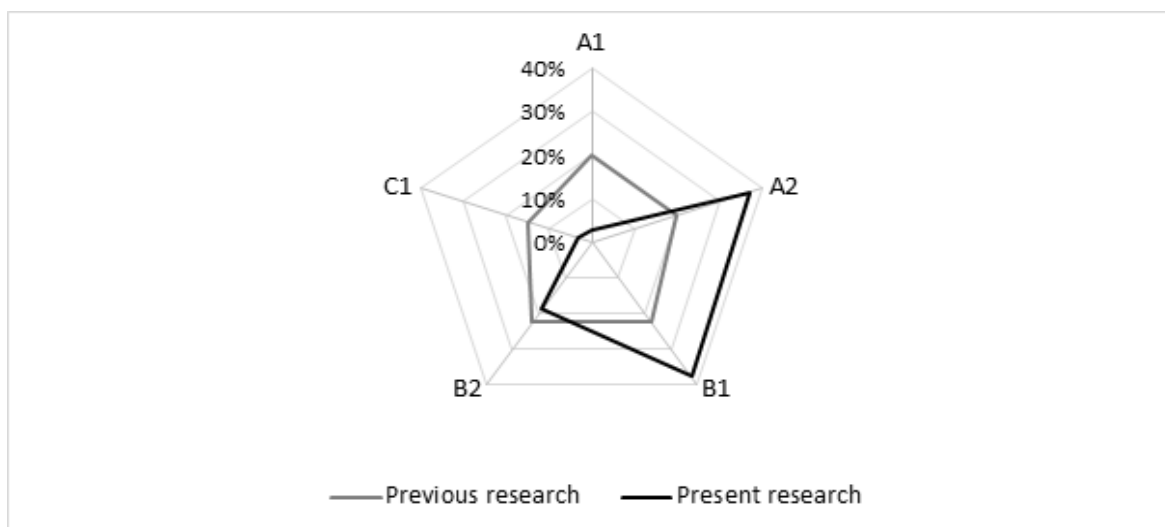
Data	Tasks	Results
<i>Park and Yoon (2021)</i>		
40 Korean undergraduates	<ul style="list-style-type: none"> <li>- Equal number of participants were chosen by English proficiency</li> <li>- Everyday topics were given in both tasks</li> <li>- Writing: 30 minutes</li> <li>- Monologues: 2 minutes</li> </ul>	Syntactic complexity was not different between the two modes, except for complex nominals per clause
<i>Hwang et al. (2020)</i>		
122 beginner-level Korean children	<ul style="list-style-type: none"> <li>- Everyday topics were given in both tasks</li> <li>- Writing: 30 minutes</li> <li>- Conversation: 25 minutes (e.g., describing a person that you want to introduce to your partner)</li> <li>- Preplanning time was given</li> </ul>	Learners used longer sentences, more subordination, more verb phrases per T-unit, and less coordination in writing than in speaking.
<i>Biber et al. (2011)</i>		
Texts from academic journals & conversation from Americans	429 research articles from 11 academic journals (participants' L1 were not provided) 723 American English conversations	Clausal subordination indices were characterized in speaking, while more noun phrases were observed in writing
<i>Kormos (2014)</i>		
44 students (15-18-year-old) in a Hungarian-English bilingual secondary school	Two narrative tasks (i.e., cartoon description, picture narration) in speech and writing	The two modes did not show any significant differences in subordination, while writing contained more modifiers per noun phrase.

The present study is most similar to that of Hwang et al. (2020) in that the Korean children produced more subordination and less coordination in writing than in speaking; however, it is difficult to compare the results of the other studies. For example, Biber et al. (2011) gathered academic articles and samples of face-to-face conversation between

American English speakers to construct their written and spoken corpora and found that the study participants used several different kinds of subordination, such as because-clauses and if-clauses, more often in speaking than in writing, while more noun phrases were observed in writing. Of course, the fact that English native speakers in their study used more because-clauses in speaking than in writing is congruent with the results of the present study; however, no common points in other grammatical factors were found. The use of the because-clauses by Korean learners will be dealt with in detail in the following section. Furthermore, the bilingual learners in Kormos' (2014) study were Hungarian teenagers who performed a cartoon description task and a picture narration task. The results showed that there was no significant difference in the ratio of subordinate clauses across the two modalities, while writing contained more modifiers per noun phrase. Furthermore, compared to Park and Yoon's (2021) study, the results of the two studies were completely different, although the same corpus was used in each study (but, there was a difference in the number of analyzed files). A major reason is the use of different methodologies: specifically, while the sample of Park and Yoon's (2021) study was intentionally selected to collect an equal amount of data by proficiency level, all collected data was used in the present study. Figure 3 shows that the rate of middle and low levels is notably higher (A2: 37%; B1: 38%) in this study than Park and Yoon's (2021) study; however, in their study, the rate of the lowest and highest level learners (A1: 20%, C1: 15%) was relatively high (Figure 3). Therefore, the results of this study were obtained from a larger and included all, which provides more reliability in characterizing speaking and writing of Korean EFL undergraduates in terms of syntactic complexity.

### Figure 3

*Data ratio of each proficiency in Park and Yoon (2021) and the present study*



In summary, the samples differed in terms of the participants' age, L1, learning environment, and proficiency. Therefore, it can be inferred that methodological differences across studies may have affected the definition of the characterization of syntactic complexity in the spoken and written production modes.

### Discussion



In the present study, I conducted a corpus-based analysis of syntactic complexity in written and spoken data provided by college-level Korean EFL learners. The following global parameters were considered: length of production units, overall sentence complexity, amounts of subordination, amounts of coordination, and phrasal sophistication. This analysis allowed me to answer important questions concerning whether and, if so, to what extent writing and speaking differ in each of these five areas of syntactic complexity. In addition, specific grammar factors in the corresponding categories through the analysis of the global measures were analyzed, and the comparison of these results with other influential previous studies was used to reliably define the characteristics of syntactic complexity in Korean learners' spoken and written languages. The important findings related to the three research questions addressed in this study can be summarized as follows.

### **Research Question 1**

The results showed that, among the 14 analyzed syntactic complexity indices, four (DCC, TS, CPT, CPC) differed significantly between written and spoken production. Specifically, subordination scores were higher for the written data, while coordination scores were higher for the spoken data. These results suggest that Korean EFL undergraduates use significantly more coordination and less subordination in speaking than in writing. These results are mostly consistent with Hwang et al.'s (2020) findings using Korean children that written data show longer sentences, more subordination, and less coordination than spoken data. However, the present findings differ from other previous studies (i.e., Biber et al., 2011; Kormos, 2014; Park & Yoon, 2021).

On the other hand, among 14 complexity indices, my participants tended to use more subordinating clauses (i.e., DCC) in writing and more coordinating clauses (i.e., TS, CPT, CPC) in speaking. Furthermore, in most syntactic indices except for the amount of coordination, L2 learners used more complicated syntax in writing than in speaking, because 11 among 14 indices were higher in writing than in speaking.

### **Research Question 2**

For a close examination of the participants' use of subordination and coordination, I analyzed the use of grammatical factors in the two categories. To this end, the log-likelihood value was calculated using AntConc3.5.8 and UCREL calculator after tagging using the Stanford NLP tagger, which is recognized to be highly accurate. In order to further improve the accuracy of tagging, five researchers including the present author performed manual final inspections based on the VOICE guidebook after the automatic tagging process. In speaking, the use of coordination was found to be more frequent than writing, so this category was closely examined for coordinating conjunctions including 'and' in the beginning of the sentence (see also Hwang et al., 2020). In writing, the use of subordination was more frequent than in speaking, so the use of its grammatical functions was carefully examined for the following functions: adverbial (i.e., if, because, although), complement (i.e., that-subordinating conjunction), and noun modifier (i.e., relative pronoun). The results showed that the participants tended to use more finite complement clauses and relative clauses in writing and more because-clause and coordinating clauses

and phrases beginning with ‘and’ in speaking.

Hwang et al. (2020) speculated that the production of a subordinate structure requires learners to constantly monitor the semantic relationship between it and the main clause; in the present study, this monitoring was easier for my participants during the written task, because writing allows more time for planning and control over linguistic forms. In contrast, since the learners were under more cognitive pressure during the spoken task, they appeared to have adopted processing strategies that allowed them to produce longer syntactic units with less cognitive effort. One such strategy was to coordinate syntactic units using ‘and’ while beginning main clauses (Hwang et al., 2020). As in a sample of Korean children analyzed by Hwang et al. (2020), Korean EFL undergraduates’ use of subordinating and coordinating clauses in the present study clearly reflects the distinct cognitive process involved in writing and speaking.

On the other hand, because-clause is likely to be used independently more in the spoken production than in the written production among Korean learners (Table 10: LL > 15.13 and Bayes Factor > 10, as noticed below Table 6). In other words, they used ‘because’ as an adverb, rather than as subordinating conjunction more in speaking than in writing (Table 11). This can be so because a Korean causal connective word, ‘waynyahamen,’ is more like a connective adverbial. In other words, it is a kind of interlanguage effect, i.e., since Korean has ‘waynyahamen’ S+V (because S+V) fragment and ‘nazenara’ S+V exists in the Japanese language, both Korean and Japanese learners tend to adopt the pattern which because-clause is used as an independent clause (Hong, 2018). Therefore, the L1 transfer seems to be most immediately salient in the scenario (Hong, 2018), especially in monologues where subordinating clauses are rare and occurrences of unfinished utterances and hesitations are frequent (Litunen & Mäkillä, 2015).

**Table 10**

*Log-likelihood (LL) and Bayes Factor results for because-clause in Monologues and Writing*

Linguistic feature	Monologue	Writing	LL*	Bayes Factor
Because-independent clause	116	42	+47.72*	36.70*
Because-dependent clause	156	134	+6.44	-4.58

**Table 11**

*Language instances of coordination in a spoken production*

File name	Examples
MK_leisure_19.1_221b	And I want to visit the England to watch the soccer game. <i>Because</i> Spain is famous too. But I’m more prefer England so I want to go there and watch some soccer game.
MK_catdog_19.1_166b	I think it’s better to raise a dog than cat. <i>Because</i> most of all there are many noise when I sleep. I can’t sleep well because of many cats.
MK_club_19.1_168b	I think all students join club is not necessary. <i>Because</i> students have their opinion. But I want to recommend club activity for every

---

students. *Because* club activity has so many advantage for students.

---

### Research Question 3

The results of the present study on spoken-written data differences in syntactic complexity with the global measures are mostly consistent with previous findings reported by Hwang et al. (2020) examining Korean EFL children. Hwang et al. (2020) showed that learners use longer sentences, more subordination (i.e., CDT), and particular structures (i.e., VPT), but less coordination (i.e., TS) in the writing task than in the speaking task. Hwang's (2020) results on the use of subordination and coordination were similar to those found in the present study. Furthermore, the results of the present study and Hwang et al. (2020) were significantly different from those reported by Biber et al. (2011), Kormos (2014), and Park and Yoon (2021). The reason behind these inconsistencies is related to the use of different methodologies, including the use of participants with different ages, L1, learning environment, proficiency, etc., as well as the use of different data selection methods. For example, the present study and Park and Yoon's (2021) study each used the same corpus, but the results were completely different. The major difference can be attributed to the different data collection processes: while Park and Yoon (2021) intentionally selected an equal amount of data by proficiency when collecting the sample, all collected data was used in the present study.

In summary, the significance of this study is that the characteristics of syntactic complexity of L2 Korean learners were defined regardless of age in global measures since it is mostly in line with Hwang et al.'s (2020) study targeting Korean EFL children. The present study revealed that the characteristics of Korean learners in speaking and writing can be characterized by the use of coordination and subordination, respectively. In addition, this study further analyzed the grammatical elements in detail and found that the frequency of learners using 'and'-utterance and 'because'-fragment at the beginning of an utterance was significantly higher than that in writing. It was also established that the use of that-clause as complement function and related pronouns was significantly more frequent in writing than in speaking. Taken together, it is generally agreed upon among scholars, that the use of more complex structures in writing than speaking is commonplace (Lintunen & Mäkilä, 2014). The reason for the difference between these two modes can be found in the developmental progression described in the Biber et al.'s (2011) study, although the type of participants and modes are not exactly comparable to the present study:

Conversation is acquired first; the grammar of writing is acquired later, and not always successfully. Grammatical structures that are readily acquired (at relatively early stages) and frequently produced in conversation by all native speakers of a language are obviously not difficult; therefore, these structures do not represent a high degree of production complexity. In contrast, many types of complex phrasal embedding are produced in only the more specialized circumstances of formal writing. These styles of discourse are not acquired naturally, and many native speakers of English rarely (or never) produce language of this type. Further, when these stages of acquisition do occur, they are late, typically in adulthood. Considering all these factors, it is reasonable to hypothesize that these grammatical structures represent a considerably higher degree of production complexity than

the conversational complexity features.

In addition, the authors of the study mentioned above (Biber et al., 2011) assumed similar developmental processes for L2 learners of English, reflecting natural progression from conversational capability to ability in academic writing. It is not always the case though as some L2 learners never acquire conversational skills, being taught written skills rather than spoken English in the first place. However, even for certain groups of learners, aptitude in English academic writing comes later in life, and thus complex features usually found within academic writing will be established in later developmental stages.

## Conclusions and Implications

An important research-oriented implication of the present study is that there are benefits to considering syntactic complexity as a multidimensional construct and carefully assembling a set of grammar features when addressing complexity-related research questions. The present study has advantages in the methodological issue: specifically, I assessed both global and specific measures of syntactic complexity. It is also significant that the characteristics of syntactic complexity of Korean EFL learners were identified through comparison with previous studies. With regard to global measures of syntactic complexity, I used 14 syntactic indices and measured with L2SCA suggested by Lu (2010, 2011). Importantly, DCC (in Subordination) was used more in L2 writing than in speaking, while more TS, CPT, and CPC (in Coordination) were used in L2 speaking. These results are meaningful and capture the characteristics of Korean learners' syntactic complexity represented differently in English monologues and writing. This study also suggests that it is very important for researchers to properly design research methodology to meet their research purposes, such as the type of participants and modes because it has a profound and direct impact on the results.

On the other hand, in terms of specific measures, I analyzed finite dependent clauses between the two modes. The results revealed that Korean L2 learners significantly more frequently used 'and' sentence-initially and 'because'-clause independently and less frequently used finite complement clauses and relative clauses in speaking than writing. The inappropriate use of 'and' and 'because' is a phenomenon that is widely seen in spoken production by Korean L2 learners, especially among learners with lower proficiency who need to constantly monitor their production (Kormos, 2014). In speaking, 'and' is one of the strategies to take time when planning sentence composition due to cognitive pressures (Hwang et al., 2020). On the other hand, because-clause is likely to be used independently rather than dependently among Korean L2 learners, which is acknowledged as the L1 transfer and seems most salient in monologues where subordinating clauses are rare, and where occurrences of unfinished utterances are frequent. As such, the results of the present study with both global and specific measures provide insights into the unique characteristics of Korean L2 learners in terms of grammar complexity.

Findings from this study point to the importance for second language teachers to be aware of the significant gap in two global (i.e., Subordination and Coordination) and four detailed aspects in finite clauses (i.e., 'and'-utterance, 'because'-fragment, complement clause, and relative clause) of syntactic complexity between L2 learners' speaking and writing. This gap calls for the design of relevant pedagogical interventions by teachers to enhance L2 university students' syntactic development.

Given the scope and the design of this research, several issues were not dealt with in the present study. First, this study adopted the segmentation method used in Litunen and Mäkillä's study (2014) in spoken language, which might have had a big effect on the metrics I examined. However, no matter what research was taken into account, one segmentation method should be selected and the present method is considered appropriate for spoken language. As for the effectiveness of this method which requires researchers to do additional manual processing when being applied to spoken data, Litunen and Mäkillä's study (2014) argued that the use of pauses in learners' spoken production should not affect the amount of the use of complex structures in their spoken language. Second, it would also be intriguing to focus exclusively on the effects of the mode on syntax complexity in speaking and writing. For instance, the use of complexity can be compared among learners by using the same prompt in each mode. Finally, it would be very beneficial to systematically investigate the effects of making use of different learners (i.e., L1 and L2 proficiency), style of tasks (i.e., conversation and interview), or task settings (i.e., timing condition, whether or not a topic is provided) on the outcome of complexity usage.

Correspondence should be addressed to Shinjae Park.

Email: muhando@sehan.ac.kr

### Notes

<sup>1</sup> Syntactic priming is the tendency for speakers to reuse previously processed syntactic structure (Cleland & Pickering, 2006: 187).

<sup>2</sup> a) The log-likelihood value is always a positive number. b) The UCREL log-likelihood wizard by Rayson inserts '+' for overuse and '-' for underuse of corpus 1 (Monologue) relative to corpus 2 (Writing).

<sup>3</sup> The texts extracted from the L2 learner corpus have errors but remain uncorrected.

### Acknowledgements

This paper was supported by the Sehan University Research Fund in 2022. Sincere thanks to Dr. Soyeon Yoon for devoting herself to compiling INU-MULC.

### References

- 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.
- Barker, F., Salamoura, A., & Saville, N. (2015). *Learner corpora and language testing*. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 511–533). Cambridge: Cambridge University Press.
- 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.
- Bonin, P., Fayol, M., & Gombert, J. E. (1998). An experimental study of lexical access in the writing and naming of isolated words. *International Journal of Psychology*, 33, 269-286
- 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.
- Chen, M., & Zechner, K. (2011, June). *Computing and evaluating syntactic complexity features for automated scoring of spontaneous non-native speech*. In Proceedings of the 49th annual meeting of the Association for Computational Linguistics: Human Language Technologies (pp. 722-731).
- Chun, J. H., & Hong, S. 2018. Pedagogical reflections on teaching because clauses in Korean EFL writing (to appear in). *Modern Studies in English Language and Literature*, 62(4), 235-257.
- 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.
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied linguistics*, 21(3), 354-375.
- Glover, P. (2011). Using CEFR level descriptors to raise university students' awareness of their speaking skills. *Language Awareness*. Vol. 20(2), 121-33.  
<https://doi.org/10.1080/09658416.2011.555556>
- Hong, S. (2018). What Japanese and Korean EFL Learners Have in Common: the Acquisition of BECAUSE-Clauses. *English Language and Literature*, 23(4), 37-56.
- 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.
- Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*. Vol. 91(4), 663-7. <https://www.jstor.org/stable/4626094>
- Hunt, K. W. (1965). *Grammatical Structures Written at Three Grade Levels*. Champaign, Ill: National Council of Teachers of English.
- Ito, F., & Misumi, H. (2016). The distinctive use of coordinating and subordinating conjunctions in Japanese EFL academic compositions. *The Gunma-Kohsen Review*, 35, 71-81.
- Jiang, J., Bi, P., & Liu, H. (2019). Syntactic complexity development in the writings of EFL learners: Insights from a dependency syntactically-annotated corpus. *Journal of Second Language Writing*, 46, 100666.
- Kellogg, R. T. (1996). *A model of working memory in writing*. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 57-71). Mahwah, NJ: Lawrence Erlbaum.
- 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.

- Kormos, J., Byrnes, H., & Manchón, R. M. (2014). Differences across modalities of performance. *Task-based language learning: Insights from and for L2 writing*, 193-216.
- 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, Atlanta, GA.
- Kyle, K., & Crossley, S. (2017). Assessing syntactic sophistication in L2 writing: A usage-based approach. *Language Testing*, 34(4), 513-535.
- 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.
- Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.
- Lim, S. B. (2018). A study of grammatical errors in college students' English writings. *The Journal of Modern British & American Language & Literature*, 36(1), 301-328.
- Lintunen, P. & Mäkilä, M. (2014). Measuring syntactic complexity in spoken and written learner language: comparing the incomparable?. *Research in Language*, 12(4), 377-399.
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15, 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. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190-208.
- Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., & McClosky, D. (2014, June). The Stanford CoreNLP natural language processing toolkit. In Proceedings of 52nd annual meeting of *The association for computational linguistics: system demonstrations* (pp. 55-60).
- Martínez, A. C. L. (2018). Analysis of syntactic complexity in secondary education EFL writers at different proficiency levels. *Assessing Writing*, 35, 1-11.
- Myles, F. (2015). *Second language acquisition theory and learner corpus research*. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 309-331). Cambridge: Cambridge University Press.
- Nippold, M. A., Frantz-Kaspar, M. W., & Vigeland, L. M. (2017). Spoken language production in young adults: Examining syntactic complexity. *Journal of Speech, Language, and Hearing Research*, 60(5), 1339-1347.
- Niu, R. (2009). Effect of task-inherent production modes on EFL learners' focus on form. *Language Awareness*, 18, 384-402.
- 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.
- Park, S. (2020). A corpus study of modal verbs in Korean learners' speech. *Journal of Linguistic Studies*, 25(2), 121-137.
- Park, S. & Yoon, S. (2021). Syntactic complexity of EFL learners' casual conversation, monologue, and writing. *The Journal of Studies in Language*, 37(1), 75-89.
- Pojanapunya, P. & Todd, R. (2018). Log-likelihood and odds ratio: Keyness statistics for different purposes of keyword analysis. *Corpus Linguistics and Linguistic theory*, 14(1), 133-167.

- Ravid, D., & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of Child Language, 29*, 419–448.
- Rayson, P. & Garside, R. (2000). Comparing corpora using frequency profiling. In Proceedings of *The Workshop on Comparing Corpora (ACL 2000)*. Hong Kong: Association for Computational Linguistics.
- Seidlhofer, B., Breiteneder, A., Klimpfinger, T., Majewski, S., Osimk-Teasdale, R., Pitzl, M. L., & Radeka, M. (2001). *Vienna Oxford International Corpus of English*.
- Tian, Y., & Lo, D. (2015, March). A comparative study on the effectiveness of part-of-speech tagging techniques on bug reports. In 2015 IEEE 22nd *International Conference on Software Analysis, Evolution, and Reengineering (SANER)* (pp. 570-574). IEEE.
- Seog, D. S. Y. (2018). A Corpus-based Study of the Use of English Articles by Korean EFL Learners. *The Linguistic Association of Korean Journal, 26*(1), 57-73.
- Seog, D. S. Y., Choi, I. C., & Lee, Y. S. (2019). Modal Usage of Advanced Korean EFL Learners, Young Korean EFL Learners, and Native English Speakers: A Corpora Comparison Study, *Language and Linguistics, 85*, 23-54.