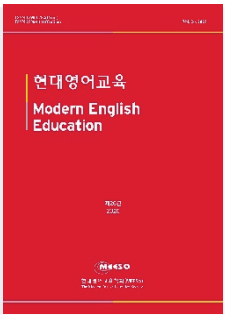
Modern English Education, Vol. 26, (2025) pp. 290-305
doi: https://doi.org/10.18095/meeso.2025.26.1.290

Modern English Education

ISSN: 1598-0782 / e-ISSN: 2586-6141
http://journal.meeso.or.kr



Examining Syntactic Complexity Across Proficiency Levels in Argumentative Essays by Korean EFL Learners

Jung-eun Year (University of Seoul)

Received: 2 June 2025
Revised: 17 June 2025
Accepted: 19 June 2025

Year, Jung-eun. (2025). Examining syntactic complexity across proficiency levels in argumentative essays by Korean EFL learners. *Modern English Education*, *26*, 290-305.

Keywords
syntactic complexity, second language writing development, Korean EFL writing development, argumentative essays
통사적 복잡성, 제2언어 쓰기 발달, 한국인 영어 학습자 쓰기 발달, 논증적 글쓰기

Jung-eun Year
Visiting Professor
Department of General English
University of Seoul
jeyear1118@uos.ac.kr

Abstract

This study investigated syntactic complexity of argumentative essays written by Korean university students with three English proficiency levels-beginning, intermediate, and upper-intermediate-in an English as a foreign language (EFL) context. Drawing on a corpus of 128 timed essays and analyzing them using the L2 Syntactic Complexity Analyzer (Lu, 2010), this study examined 12 indices representing four dimensions of syntactic complexity: length-based measures, subordination, coordination, and phrasal elaboration. Findings revealed a general upward trend in syntactic complexity with increasing proficiency, particularly in mean length of T-unit (MLT), dependent clauses per T-unit (DC/T), and complex nominals per T-unit (CN/T). Subordination showed a modest increase at the intermediate level and continued to increase through the upper-intermediate level, where it significantly distinguished learners from beginning-level peers. Phrasal elaboration became more prominent at the upper-intermediate level. However, coordination measures remained stable across groups, possibly reflecting L1 influence. While results suggested a developmental shift from clause-based to phrase-based structures, the lack of change in coordination suggested that this shift might not apply uniformly across syntactic dimensions. This study concludes with pedagogical implications for EFL writing instruction, emphasizing the need for targeted support in developing phrasal complexity and understanding learners' syntactic growth.

INTRODUCTION

Syntactic complexity in second language (L2) writing generally refers to the range and degree of sophistication of syntactic structures produced by L2 users in their texts (Ortega, 2003; Wolfe-Quintero et al., 1998). Over the past few decades, this construct has attracted considerable attention among researchers, as L2 writers' ability to produce more complex, though not necessarily more complicated (Larsen-Freeman, 2013), structures, characterized by structural variation and elaboration, is assumed to reflect reliable indices of their writing performance and development. By extension, it is also considered to reflect their overall proficiency in the target language (Lu, 2010).

L2 research to date has indeed revealed that, as proficiency increases, L2 writers use a greater variety of word combinations, phrases, and sentences, progressively employing more elaborate forms in ways that are both clear and informative in communicating complex thoughts (Foster & Skehan, 1996). More specifically, empirical studies have shown that syntactic development in academic writing tends to follow a continuum, shifting from clausal-level constructions, particularly coordinate and subordinate clauses commonly found in conversational discourse, toward increased use of phrasal-level constructions, such as dense noun and prepositional phrases, that typify written discourse (Biber et al., 2016; Norris & Ortega, 2009).

Research further highlights that this progression, as with other aspects of L2 development, is fluid and often non-linear, exhibiting substantial individual variation and interacting in complex ways with several factors, such as the writer's first language (L1) (Lu & Ai, 2015), discourse genre (e.g., argumentative vs. narrative) (Zheng & Barrot, 2024), and task conditions, including planning time (timed vs. untimed) (Lu, 2011), task complexity (Biber et al., 2020), and instructional context (e.g., ESL vs. EFL) (Ortega, 2003), among others. This complexity calls for a more nuanced understanding of syntactic development and thoughtful application of research findings in instructional settings.

To contribute to this line of research, and as a step toward designing and improving writing programs that effectively support L2 writers' syntactic development, the present study explores the syntactic features of texts produced by Korean adult learners of English in foreign language contexts-an area that remains relatively underexplored in the literature (Ortega, 2003). More specifically, this study, both exploratory and descriptive in nature, seeks to identify the syntactic features of argumentative essays written by learners at three different proficiency levels: beginning, intermediate, and upper-intermediate. These proficiency levels are determined by a standardized test, rather than being based on writing performance, as few studies have explicitly examined how syntactic features vary according to learners' general proficiency (see Bae & Min, 2020). The study further examines whether, and to what extent, syntactic features across proficiency levels reflect developmental patterns observed in previous research. This line of inquiry is expected to provide valuable pedagogical implications, helping L2 writing instructors and course designers better understand students' specific needs and challenges related to syntactic development and fine-tune their interventions to more effectively support learners' academic writing.

LITERATURE REVIEW

Syntactic complexity, often used interchangeably with grammatical complexity in the L2 research literature (Wolfe-Quintero et al., 1998), constitutes a multifaceted construct encompassing discrete components and distinct dimension levels that emerge at different stages of language development (Bulté & Housen, 2014). This means it cannot be captured by a single measure alone, but requires a range of indices that adequately represent the various structural aspects and developmental patterns of learner language. Reflecting this multicomponential and multidimensional nature, and facilitated by automated computational tools, recent studies investigating syntactic complexity in L2 writing have employed a diverse set of analytical measures, which are broadly categorized into length-based, subordination, coordination, and phrasal complexity measures.

Importantly, these measures not only reflect formal structural variation, but also align with broader functional trajectories of syntactic development in L2 writing, as proposed within Systemic Functional Linguistics (Halliday & Matthiessen, 1999, as cited in Norris & Ortega, 2009). According to this framework, syntactic choices and their subsequent development are motivated by the communicative purposes and situational contexts in which language is used. At the early stages of writing, L2 writers create a more explicit and sequential flow of information by linking ideas through coordinate conjunctions and embedding dependent clauses. This style of expression is termed "dynamic," reflecting clause-based, flexible, and explicitly connected structures that unfold in real time. It is most prevalent in everyday oral discourse and is characteristic of the intermediate stage in L2 writing. With increased proficiency, writers tend to rely on more informationally condensed features, such as noun and prepositional phrases, which support the compactness expected in formal academic genres. Referred to as

the "synoptic" style of expression, this style is closely associated with formal written discourse and is characteristic of advanced stages in L2 writing.

Together, the shift from dynamic to synoptic expression outlines a developmental trajectory in syntactic complexity-progressing from simple and sometimes fragmented clauses to coordinated and embedded structures, and ultimately culminating in dense phrasal constructions (Norris & Ortega, 2009). Situated within this conceptual framework and informed by prior studies, this review examines key measures of syntactic complexity, their theoretical underpinnings, and major findings derived from empirical research.

Length-Based Measures

Length measures are among the most commonly employed metrics for quantifying syntactic complexity in L2 research (Norris & Ortega, 2009; Ortega, 2003). Often referred to collectively as "mean length of unit" (MLU), these measures typically involve calculating the mean length of a production unit such as a clause, a T-unit, or a sentence (e.g., the average number of words per sentence). While operational definitions of these production units vary across studies, a clause is generally defined as a structure consisting of a subject and a finite verb; a T-unit as "one main clause plus any subordinate clauses attached to or embedded in it" (Hunt, 1970, p. 4); and a sentence as a group of words that conveys a complete thought and ends with terminal punctuation (Wolfe-Quintero et al., 1998). Commonly used measures in writing research include mean length of sentence (MLS), mean length of clause (MLC), and mean length of T-unit (MLT). Among these, MLT is often favored in L2 studies because it focuses on a core syntactic unit (i.e., the T-unit) while avoiding inflation from run-on sentences (Larsen-Freeman, 1978). It also integrates multiple aspects of complexity that contribute to unit length, including clause length, subordination, and phrasal elaboration, thereby providing a parsimonious yet omnibus summary measure (Biber et al., 2020; Lu, 2011; Ortega, 2003).

Originating in L1 writing research (Brown, 1973; Hunt, 1970), length-based measures have been widely used as generic or global indices of overall syntactic complexity and, in some cases, as proxies for general language proficiency (Bulté & Housen, 2014). Theoretically, these measures rest on the premise that as writers advance in language proficiency, they acquire new syntactic constructions and express increasingly complex ideas with more words per production unit and more elaborate syntactic features. In this view, longer syntactic units are construed as more complex, and syntactic development in writing is reflected in gradual increases in these length-based metrics (Norris & Ortega, 2009; Ortega, 2003). Historically, these measures have remained central in both L1 and L2 research, largely due to their practical utility; before the advent of automated tools such as the L2 Syntactic Complexity Analyzer (Lu, 2010), earlier studies relied extensively on length-based measures, as the number of words per production unit was relatively easy to compute manually. Recently, the development of automated computational tools has enabled researchers to calculate a broader array of syntactic complexity measures with greater ease and efficiency. This notwithstanding, L2 research continues to rely on traditional length-based indices, recognizing their usefulness for comparison with previous studies.

Numerous empirical studies have confirmed the robustness of length-based indices, demonstrating a strong correlation between increased length of production units and both L2 proficiency and writing quality. In a synthesis of L2 writing research, Ortega (2003) concluded that length-based measures such as mean length of sentence (MLS), T-unit (MLT), and clause (MLC) are valid indicators of syntactic development, revealing that high-proficient writers, in both second and foreign language contexts, tend to produce longer syntactic units than low-proficient writers. Ortega further noted that ESL learners, who have more extensive writing experience and, consequently, higher proficiency, produced longer syntactic structures than their EFL counterparts. Findings from studies with EFL writers have reinforced the predictive value of length-based measures for assessing syntactic complexity. In a study analyzing Korean university learners' timed argumentative essays across beginning, intermediate, and advanced levels, J. Kim (2014) found that length-based measures were reliable indicators that differentiated all three proficiency groups; learners produced longer MLS, MLT, and MLC as proficiency increased. Similarly, Khushik and Huhta (2020), analyzing timed essays by adolescent EFL learners from Sindhi and Finnish L1 backgrounds, showed that sentence and T-unit length effectively differentiated learners at lower proficiency levels.

On the other hand, some studies have provided evidence that their predictive validity is limited at higher proficiency levels. Instead, these studies report that advanced writers tend to adopt alternative syntactic strategies that do not necessarily involve lengthening structures. For example, Casal and Lee (2019), in their examination of first-year L2 writing across low-, mid-, and high-rated papers, found significant differences in MLT between low- and mid-rated papers. By contrast, no significant differences were observed between mid- and high-rated papers in terms of MLT. Rather, they identified significant differences in the use of specific nominal structures containing attributive adjectives and prepositional, postnominal, or participial modifiers-features that were characteristic of high-rated academic writing in their corpus. This

finding suggests that L2 writers at advanced stages tend to rely more on structurally complex, fine-grained syntactic features rather than on length alone.

Moreover, research indicates that the relationship between length-based measures and proficiency or writing qualities may be confounded by genre, given that different genres elicit distinct types of syntactic structures to fulfill their specific communicative purposes. A case in point is Bae and Min (2020), who investigated the impact of genre on the syntactic complexity used by Korean L2 college students in four different genres: narrative, comparison, cause-effect, and argumentative essays. While MLS significantly distinguished beginning from intermediate learners in comparison essays, it did not reliably differentiate intermediate from advanced learners. In other genres, such as narrative, cause-effect, and argumentative writing, length-based measures (e.g., MLS and MLT) generally failed to yield statistically significant distinctions across proficiency levels. Instead, the authors reported that proficiency differences across different genres were more effectively captured by fine-grained indices of syntactic complexity, such as clausal or phrasal complexity.

Overall, research findings support the view that length-based measures are robust indicators of syntactic complexity across broad linguistic contexts, as they tend to correlate with general L2 proficiency and writing quality. However, their validity may be limited, particularly at higher proficiency levels and across varied genres, where their sensitivity appears to diminish. As large-grained indices, therefore, length-based measures are most effective only when used in conjunction with more fine-grained measures that target specific structural features, such as subordination and phrasal elaboration.

Subordination

Subordination measures focus on the clausal density of dependent or subordinate clauses embedded within larger production units. These subordinate clauses include adverbial clauses (e.g., *because students arrived late*), relative clauses (e.g., *students who major in engineering*), and nominal clauses (e.g., *I believe that students will complete the assignment*). Such measures are typically assessed using ratios that capture the frequency of subordination, including clauses per T-unit (C/T), complex T-units per T-unit (CT/T), dependent clauses per clause (DC/C), and dependent clauses per T-unit (DC/T).

Subordination is considered a key indicator of syntactic complexity, because subordinate clauses reflect the ability to attach or embed clauses within clauses and create hierarchical relationships. It is instrumental in transforming a sequence of independent clauses (e.g., *I went to California. I saw a movie. It was a new movie. Jeremy had recommended it*) into a more cohesive and informative structure (e.g., *When I went to California, I saw the new movie that Jeremy had recommended*) (examples from Nippold et al., 2005, p. 1049). Due to this increased clausal density, such subordinate structures are often regarded as cognitively more demanding to process than other types of syntactic structures (though see Bulté & Housen, 2012, for a critical discussion of this assumption). Accordingly, the traditional view holds that a higher density of subordination reflects greater syntactic complexity, which in turn signals higher writing proficiency during development. This perspective has led earlier studies to accord greater importance to the analysis of subordinate structures.

Research findings, however, suggest that subordination may not serve as a consistent indicator of syntactic growth across all proficiency levels, as its density does not increase linearly beyond the intermediate stage. Instead, the use of subordinate structures is often found to level off or even decline at more advanced stages. For instance, L1 writing research has shown that the frequency of subordination tends to plateau relatively early, around age 11, after which it no longer reliably reflects syntactic development (Nippold et al., 2005). Similar patterns have been observed in L2 writing. Lu (2011), analyzing timed argumentative essays by Chinese EFL university students, found that subordination measures such as dependent clauses per clause (DC/C) and dependent clauses per T-unit (DC/T) were more prominent at lower to intermediate levels, while advanced learners relied less on clausal subordination. A study examining college-level German L2 writing by Byrnes et al. (2010) reported comparable results: subordination, as measured by clauses per T-unit, increased significantly from lower to intermediate levels but tapered off at very advanced stages.

Yet not all studies support this non-linear trend. Some have documented a relatively steady increase in subordination across proficiency levels, potentially influenced by learners' L1 backgrounds. For example, studies focusing on Korean EFL writers (J. Kim, 2014; Park, 2012) found that complex T-units per T-unit (CT/T), a proxy for subordination, clearly distinguished advanced-level writers from those at lower proficiency levels, suggesting that subordination remained a valid indicator of advanced writing. The findings of Bae and Min (2020) also mirrored these results, reporting a linear relationship between subordination and proficiency among college-level Korean L2 writers. These patterns may be attributed to L1 influence, as Korean, which features a rich inventory of verbal suffixes to link clauses, favors subordination through clause chaining in academic writing (Jendraschek & Shin, 2011). Similarly, Neff et al. (2004, as cited in Ortega, 2012) found that Spanish L1 EFL learners produced unexpectedly high levels of subordination compared to L1 English writers, possibly due to cross-rhetorical transfer from Spanish, which favors heavy subordination in formal writing. These findings highlight the

complexity of using subordination as a universal marker of syntactic development and suggest that its diagnostic value may vary depending on the writer's L1 background.

Furthermore, research has shown that the use of subordination differs significantly across discourse modes and genres. As aforementioned, subordination features, such as finite dependent clauses, are more frequent in conversation and other spoken registers than in formal academic writing, which increasingly favors dense nominal phrases and other phrasal grammatical features (Biber & Gray, 2016). Extending this observation to academic genres, levels of syntactic complexity, and specifically the extent to which subordinate clauses are used, vary substantially across genres of academic writing. Narrative writing, which typically involves recounting events or telling stories in a manner similar to spoken discourse, has been shown to contain higher ratios of subordinate clauses. In contrast, argumentative essays, which aim to express abstract ideas and logical relationships, tend to rely less on clausal subordination and more on phrasal elaboration (Beers & Nagy, 2009, for L1 studies; see also Mazgutova & Kormos, 2015; Yoon & Polio, 2017, for L2 studies).

To recapitulate, it may be fair to say that subordination, at least in English writing, marks a dominant syntactic structure at intermediate stages. It stands as a useful indicator for distinguishing beginning and intermediate learners but becomes less reliable for fully capturing progress at more advanced levels. As L2 writers progress, they often reduce their reliance on subordinate clauses and shift gradually toward more compact and information-dense phrasal constructions. However, it is also worth emphasizing that the deployment of subordinate clauses in academic writing can be mediated not only by proficiency level but also by the writer's L1 background and genre-specific rhetorical demands.

Coordination

As a distinct subconstruct of syntactic complexity, coordination is seen as a primary indicator of complexification in the early stages of language development (Bardovi-Harlig, 1992). It reflects how frequently independent clauses or phrases are linked using coordinating conjunctions such as *and*, *but*, or *so*. Coordination is measured through indices such as T-units per sentence (T/S), which captures clausal coordination at the sentence level, that is, the number of clauses coordinated within sentences, and coordinate phrases per clause (CP/C) or per T-unit (CP/T), both of which assess phrasal coordination at the clause or T-unit level.

Coordination remains an important, though sometimes overlooked, dimension of syntactic complexity, partly because earlier studies often centered on intermediate or advanced learners in ESL contexts and therefore prioritized subordination and more sophisticated forms. However, Norris and Ortega (2009) point out that L2 writers rely heavily on sentential coordination to combine ideas at incipient levels of L2 writing proficiency. As learners' syntactic repertoires expand at intermediate stages, they tend to shift toward subordination, while their use of coordination declines. It is therefore essential to include coordination-based measures if the study aims to capture developmental patterns at beginning proficiency levels and to better understand the broader trajectory of syntactic development across proficiency levels.

In support of this view, Bardovi-Harlig (1992), one of the first researchers to draw attention to the significance of coordination measures at early stages of development, found that coordination index scores in ESL learners' compositions decreased steadily from beginners (39.9%) to advanced learners (10.3%). In a cross-sectional and longitudinal study of German L2 learners, Vyatkina (2012) likewise observed that the frequency of coordinating conjunctions (CC) per 100 words declined over time, suggesting that reliance on coordination tends to diminish with advancing writing proficiency. The study further reported a strong negative correlation between the use of coordinating conjunctions and subordination, indicating that as German L2 writers progress, they used fewer coordinated structures while increasing their use of subordinate construction, in line with established developmental patterns in L2 writing.

However, empirical research has presented a more complex picture, challenging the assumption that increased coordination is solely indicative of lower proficiency. Ai and Lu (2013), for instance, found no statistically significant differences in the mean values of the CP/C and T/S measures among their three groups: lower-proficiency NNS, higher-proficiency NNS, and NS university students. They also noted that the two NNS proficiency levels did not differ significantly from each other in the amount of coordination used. Similarly, Khushik and Huhta (2020, 2022) reported that coordination indices generally failed to distinguish between the Common European Framework of Reference for Languages (CEFR) levels A1, A2, and B1, although a slight increase was observed between A1 and A2.

Several researchers have also pointed to potential L1 effects. Lu and Ai (2015), for example, reported that advanced learners from certain L1 backgrounds (e.g., French, German, Russian) demonstrated greater use of sentential coordination (T/S) than their native English-speaking peers, possibly as a result of transfer from their first languages. In stark contrast, they found that L1 Chinese college students used significantly fewer coordinate sentences than their L1 English counterparts. This pattern may be ascribed to the structural feature of Chinese, where two independent clauses can be joined by

punctuation rather than coordinating conjunctions. A similar finding was reported by Yin et al. (2021), who also documented reduced use of coordinate structures among Chinese L1 learners.

This body of research collectively suggests that the use of coordination is a complex feature influenced not only by proficiency but also significantly by the learner's L1 background. Coordination measures, particularly sentential coordination (T/S), are most indicative of syntactic complexity at beginning proficiency levels, and a decline in clausal coordination over time is often interpreted as a sign of linguistic development. Yet, this developmental trend may not hold universally. For some L1 groups, coordination may persist or even increase at higher proficiency levels, reflecting cross-linguistic influence rather than a lack of syntactic advancement.

Phrasal Elaboration

Phrasal elaboration or phrasal complexity, recognized as a distinct, subclausal dimension alongside clausal and global complexity, pertains to the internal structure and modification, most notably within noun phrases. It reflects a learner's ability to compress information into more compact and syntactically sophisticated forms, particularly through nominalization, dense noun phrase modification, and the use of prepositional phrases. As illustrated in the following examples, these features allow writers to condense three clauses (e.g., *When he announced that he was leaving, it surprised them*) into a succinct yet information-dense structure (e.g., *The announcement of his departure surprised them*) (examples from Beers & Nagy, 2009, p.189). This dimension of complexity involves features such as attributive adjectives (e.g. the *challenging* assignment), pre-modifying nouns (e.g., *course* materials), post-modifying prepositional phrases (e.g., the result *of the group project*), and nonfinite clauses (e.g., the article *assigned for next week*). Common indices used in previous studies include complex nominals per clause (CN/C), and complex nominals per T-unit (CN/T), among others.

Recently, a growing body of empirical evidence has demonstrated that phrasal elaboration is a defining feature of academic writing at advanced stages. A series of studies conducted by Biber and his colleagues (Biber & Gray, 2013; Biber et al., 2011, 2016, 2020; Staples et al., 2016) has rigorously investigated the linguistic features of conversation and various forms of academic writing, challenging the traditional view that grammatical complexity in L2 writing is primarily marked by increased use of subordinate clauses. Their findings showed that formal academic writing relies heavily on nouns and other phrasal modifications such as premodifying nouns, *of*-constructions, attributive adjectives, and prepositional phrases. They further noted that L2 writers gradually increase their use of phrasal features while decreasing their reliance on clausal features, such as finite subordinate clause, over time and with increasing proficiency.

Further supporting this view, Casal and Lee (2019), in their assessment of first-year L2 writing, found that higher-rated writing is closely associated with phrasal complexity. Specifically, they observed that complex nominal densities and the use of particular complex nominal types (e.g., adjective pre-modification, preposition post-modification, and participle modification) were significantly associated with high quality academic writing. In contrast, low-rated papers exhibited significantly lower densities of complex nominals. Similarly, in a study analyzing independent essays from the Test of English as a Foreign Language (TOEFL), Kyle and Crossley (2018) found that phrasal complexity accounted for approximately 20.3% of the variance in holistic writing scores.

Taken together, research in L2 writing has shown that phrasal complexity is a hallmark of academic writing and an important index of language development. While clausal complexity tends to plateau or even decrease at intermediate to advanced proficiency levels, phrasal complexity continues to grow and become increasingly prominent with proficiency. This trend suggests that phrasal elaboration represents a key area of syntactic growth at more advanced stages of writing development (Bulté & Housen, 2014; Lu, 2011). In light of this, understanding phrasal complexity is crucial for accurately describing the linguistic features of academic writing and for tracking the development of L2 writers towards academic proficiency (Biber et al., 2011).

The Present Study

Extensive research to date has established a wide range of measures for examining the multidimensional nature of syntactic complexity (Bulté & Housen, 2012). With the development of automated computational tools such as the L2 Syntactic Complexity Analyzer (Lu, 2011), recent L2 studies have been able to efficiently examine a range of syntactic complexity indices, each reflecting distinct dimensions of syntactic development. Building on this body of research, and supported by automated analysis, the present study was conducted to investigate the syntactic features manifested in argumentative essays written by Korean university students in an EFL context and to examine developmental differences in syntactic complexity across three proficiency levels: beginning, intermediate, and upper-intermediate. Two research questions guided the present study:

Research Question 1: Which syntactic features characterize the argumentative essays produced by Korean EFL learners at the beginning, intermediate, and upper-intermediate proficiency levels?

Research Question 2: To what extent do the developmental patterns observed in this cross-sectional study indicate a shift from dynamic, clause-level syntactic structures to synoptic, phrasal-level constructions in the writing of Korean EFL learners?

METHOD

Participants and Proficiency Groups

The Korean EFL participants in this study consisted of 152 students enrolled in six intact English writing classes at a university in South Korea. The course was a mandatory, first-year requirement designed to provide students with output-oriented writing practice in an EFL context. Each class met once a week for a 100-minute session over a 15-week semester. The participants, ranging in age from 18 to 24, represented a diverse range of majors, including humanities, business, sociology, art, and music. Most students had received little to no formal instruction in English writing prior to the course. The data thus reflect authentic learner output produced under typical instructional conditions, offering a pedagogically grounded perspective on syntactic development at the tertiary level.

To explore the syntactic profiles of argumentative essays across different proficiency groups, students were divided into three language proficiency levels according to their scores on the General Tests of English Language Proficiency (G-TELP), an international English proficiency assessment developed by the International Testing Services Center (ITSC). The test, which evaluates general English ability, was administered during the first week of the semester and included three sections: grammar, listening, and reading. Of the 152 students, participants were excluded from the study if they did not complete the standardized English proficiency test (n = 18) or if their written text contained fewer than 100 words (n = 6).

The remaining 128 students were classified into three proficiency groups based on their performance on the proficiency test, which was aligned with the Common European Framework of Reference for Languages (CEFR). This CEFR-based grouping provides a widely recognized framework for interpreting syntactic complexity across developmental stages and facilitates meaningful comparisons with existing L2 writing research (Hulstijn et al., 2010; Khushik & Huhta, 2022). Cut-off points were determined using established CEFR descriptors (M. Kim, 2023), allowing for the classification of learners as beginning-level (A2: below 59), intermediate-level (B1: 60–74), and upper-intermediate-level (B2: above 75). Two students who scored above 95 (C1) were included in the B2 group, as their essays did not exhibit features warranting separate treatment as C1-level writing. Specifically, a comparative analysis of their essays and those of the upper-intermediate group across four dimensions of syntactic complexity revealed no statistically significant differences. The resulting group sizes were 28 students in A2, 44 in B1, and 56 in B2, reflecting distribution patterns in urban universities in EFL contexts where more learners cluster around intermediate and upper-intermediate levels (cf. S.Kim, 2021; Yoon, 2017). While this grouping scheme offers strong external validity, its uneven group sizes were taken into account when selecting statistical procedures.

To assess the validity of the grouping method, a preliminary correlation analysis was conducted. The correlation between participants' general proficiency levels and their holistic writing scores, as measured by IntelliMetric, an automated essay scoring tool (Rudner et al., 2006), and instructor-assigned ratings, was statistically significant and indicated a moderate-to-large effect size (r = .49, p < .001). This result supports the convergent validity of the proficiency grouping, suggesting that general English proficiency, as measured by the G-TELP, corresponds reasonably well with holistic writing quality as assessed by both automated scoring and a human rater.

Corpus Data and Participants

The corpus for the present study comprises 128 argumentative essays written by Korean university students. These essays were distributed across three proficiency levels: 28 from beginning-level learners (A2), 44 from intermediate-level learners (B1), and 56 from upper-intermediate learners (B2). The essays were collected during the fifteenth week of the semester as part of the course requirements, serving as the final in-class writing assessment. The writing task was completed within a 60-minute time limit, during which students were not permitted to consult external resources such as dictionaries. A total of three writing prompts were used. These prompts addressed general knowledge topics related to students' college experiences and did not require specialized content knowledge. Example prompts included: (1) Do you agree or disagree with the following statement? Students should pursue internship opportunities before graduation; (2) Should university students work

part-time during the academic semester?; and (3) Do you prefer to work in groups or individually for class projects? All essays were handwritten in class and subsequently transcribed electronically for analysis, with spelling errors corrected, run-on sentences segmented, and missing sentence-final punctuation added, following preprocessing practices recommended in corpus-based Natural Language Processing (NLP) research (Lu, 2014).

Computational Tool

The students' essays were analyzed using the L2 Syntactic Complexity Analyzer (L2SCA), developed by Lu (2010), an automated tool that measures syntactic complexity in college-level L2 writing. This tool has been widely used in recent L2 writing studies and has gained popularity among researchers in the field (e.g., Bae & Min, 2020; Bi & Jiang, 2020; Casal & Lee, 2019; Yoon & Polio, 2017; Zheng & Barrot, 2024) for three main reasons. First, it generates 14 indices of syntactic complexity, which are grouped into five broad categories: length-based measures, subordination, coordination, phrasal features, and overall sentence complexity (see Table 1). These indices tap into different aspects of syntactic development and reflect the types of quantitative measures commonly used in current corpus-based research on L2 writing. As such, they allow for meaningful comparisons across different proficiency levels (see Lu, 2010, 2011, for detailed descriptions). Second, the tool has demonstrated strong reliability in validation studies. Lu (2010), in a study comparing manually annotated essays with L2SCA outputs, found strong agreement, and this finding was later corroborated by Jiang et al. (2019) and Yoon and Polio (2016), suggesting that the tool yields reliable syntactic complexity scores when applied to L2 learners' texts. Third, L2SCA is freely available both as a web-based interface and a downloadable package, and it supports batch processing, which makes it especially useful for analyzing large sets of learner texts.

Given these strengths, the analyzer serves as a useful starting point for describing and comparing syntactic features in the present corpus of EFL essays written at three different levels of English proficiency. Of the 14 indices available in the system, this study focused on 12, selected for their relevance to the research questions and their frequent use in prior studies. C/S (clauses per sentence) and VP/T (verb phrases per T-unit) were excluded due to their limited added value. C/S was omitted due to its high correlation with core subordination measures (e.g., C/T, $r = .94$; DC/T, $r = .91$), which rendered its interpretive value largely redundant. VP/T (verb phrases per T-unit) was not included because this study concentrated on noun phrase complexity in examining phrasal complexity in academic writing, and because it was also strongly correlated with another length-based metric, MLT ($r = .93$). Moreover, these two indices have not been widely employed in previous research as reliable indicators of syntactic development (Lu, 2011; Yoon, 2017).

TABLE 1
Syntactic Complexity Measures Evaluated (Lu, 2010, 2017)

Measure	Code	Definition
Length of production unit		
Mean length of clause	MLC	# of words / # of clauses
Mean length of sentence	MLS	# of words / # of sentences
Mean length of T-unit	MLT	# of words / # of T-units
Sentence complexity		
Clauses per sentence	C/S	# of clauses / # of sentences
Amount of subordination		
T-unit complexity ratio	C/T	# of clauses / # of T-units
Complex T-unit ratio	CT/T	# of complex T-units / # of T-units
Dependent clause ratio	DC/C	# of dependent clauses / # of clauses
Dependent clauses per T-unit	DC/T	# of dependent clauses / # of T-units
Amount of coordination		
Coordinate phrases per clause	CP/C	# of coordinate phrases / # of clauses
Coordinate phrases per T-unit	CP/T	# of coordinate phrases / # of T-units
Sentence coordination ratio	T/S	# of T-units / # of sentences
Degree of phrasal sophistication		
Complex nominals per clause	CN/C	# of complex nominals / # of clauses
Complex nominals per T-unit	CN/T	# of complex nominals / # of T-units
Verb phrases per T-unit	VP/T	# of verb phrases / # of T-units

Data Analyses

To identify the syntactic complexity features associated with the three proficiency groups and to determine which measures most clearly differentiate these groups, a series of statistical analyses was conducted. First, a one-way ANOVA was performed to examine group differences across 12 syntactic indices. When the ANOVA revealed statistically significant effects, Bonferroni-adjusted post hoc tests (p = .0167) were conducted to determine whether differences existed between specific pairs of groups.

When the assumption of normality did not hold, as indicated by Shapiro–Wilk tests (p < .001), non-parametric procedures were performed alongside the parametric analyses to support a more robust interpretation. While the one-way ANOVA served as the main procedure for detecting overall group differences, Kruskal–Wallis tests were applied to the non-normally distributed variables, followed by Mann–Whitney U tests with Bonferroni correction for post hoc comparisons, offering a complementary rank-based view.

RESULTS

Descriptive Statistics

To present an overview of the dataset, descriptive statistics were compiled for the three proficiency groups, including their proficiency test scores, and the average number of words (W) and sentences (S) produced per argumentative essay (see Table 2). These variables provide useful background information about learners' general writing behavior across proficiency levels (J. Kim, 2014; Lu, 2010).

An ANOVA revealed a significant difference in proficiency test scores among the three groups, $F(2, 125) = 378.71$, $p < .001$, with a very large effect size ($\eta^2 = .86$). Post hoc analyses confirmed that all three groups significantly differed from one another (all $p < .001$), indicating a clear linear progression in general English proficiency: upper-intermediate students demonstrated higher proficiency than intermediate students, who, in turn, had higher proficiency than those at the beginning level.

Similarly, the results for total word count (W) showed a significant group difference, $F(2, 125) = 17.29$, $p < .001$, with a large effect size ($\eta^2 = .22$). Post hoc comparisons revealed that the upper-intermediate group differed significantly from both the beginning and intermediate groups ($ps < .01$), while the difference between the intermediate and beginning groups did not reach significance after Bonferroni correction (p = .030). These findings suggest a general trend: upper-intermediate students wrote longer texts than intermediate students, who, in turn, wrote more than beginning-level students, although the latter difference was not statistically significant.

The results for total sentence count (S) showed a significant group difference, $F(2, 125) = 4.74$, p = .01, with a small to moderate effect size (η^2 = .07). Post hoc comparisons revealed a marginally significant difference between the beginning and upper-intermediate groups (p = .016), relative to the Bonferroni-adjusted threshold (p = .0167), while no significant differences were found between the other adjacent groups (A2 vs. B1, p = .02, and B1 & B2, p = 1.0). These results suggest a trend in which both intermediate and upper-intermediate groups tended to produce more sentences than the beginning group, although the differences were not statistically robust.

TABLE 2

Means and Standard Deviations of Proficiency Test Scores, Word Counts, and Sentence Counts by Proficiency Group

Variable	Beginning (A2) (n = 28)		Intermediate (B1) (n = 44)		Upper-Intermediate (B2) (n = 56)		Total (n = 128)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Average scores of proficiency test	53.29	5.16	68.00	4.05	81.34	4.42	70.62	11.79
Average number of words (W)	256.96	64.50	306.39	49.88	337.82	63.53	309.27	66.63
Average number of sentences (S)	18.50	3.53	20.93	3.62	20.93	3.82	20.40	3.80

Note. The values for total word count (W) and total sentence count (S) were automatically calculated using the L2 Syntactic Complexity Analyzer (L2SCA; Lu, 2010).

Syntactic Complexity Measures

Table 3 summarizes the means (*M*) and standard deviations (*SD*) of 12 syntactic complexity values in the argument essays at each proficiency level and the results of one-way ANOVAs conducted on the group means. Results are presented across four categories: length-based measures, subordination, coordination, and phrasal elaboration.

Length of Production Units

All three measures related to the length of production units, mean length of clause (MLC), mean length of sentence (MLS), and mean length of T-unit (MLT), showed a gradual upward progression across the three proficiency levels, indicating that learners produced longer syntactic units in their writing as a function of their general proficiency.

Among these, MLS and MLT, both of which demonstrated statistically significant differences, were found to be highly correlated (r = .88, p < .001). This strong collinearity suggests that the two indices capture similar aspects of syntactic complexity in learners' writing, likely because most sentences in the dataset contained only a single T-unit. Given this redundancy, MLT was selected as the primary length-based complexity measure for subsequent analyses. This decision reflects both theoretical and practical considerations. Theoretically, MLT offers a more structurally meaningful indicator of syntactic complexity, as it captures the length of an independent syntactic unit (i.e., a T-unit) without being influenced by sentence-level variation; practically, it is more widely recognized in L2 writing research as a robust and reliable index of syntactic development.

For MLT, a one-way ANOVA revealed a significant group difference, $F(2, 125) = 11.15$, $p < .001$, with a large effect size ($\eta^2 = .15$). Post hoc analyses with Bonferroni correction revealed significant differences between the beginning and upper-intermediate groups, $p < .001$, and a marginally significant difference between intermediate and upper-intermediate groups ($p = .016$). No statistically significant difference was found between the beginning and intermediate groups. The mean MLT values of the beginning group (M = 13.30), the intermediate group (M = 14.41), and the upper-intermediate group (M = 15.71) suggest that a difference of approximately two or more words was sufficient to yield statistical significance, whereas a mean difference of one word or less was not.

Amount of Subordination

All four subordination-related measures, T-unit complexity ratio (C/T), complex T-unit ratio (CT/T), dependent clause ratio (DC/C), and dependent clauses per T-unit (DC/T), were found to be strongly correlated (r >.80). A preliminary manual inspection suggests that this high correlation may stem from a structural tendency among the Korean EFL learners in the present study to produce T-units containing a single dependent clause. Given this collinearity, DC/T was retained as the primary subordination index, as it most directly reflects the frequency with which learners embed subordinate clauses within larger syntactic units and has been shown to be developmentally sensitive in previous research. The other indices were excluded from further analysis to reduce redundancy and improve model interpretability.

For DC/T, the mean values increased steadily from the beginning group (M = 0.47) to the intermediate (M = 0.54) to the upper-intermediate groups (M = 0.64), indicating a developmental trend in the use of subordinate structures. A one-way ANOVA revealed a significant group difference, $F(2, 125) = 7.43$, $p < .001$, with a moderate effect size ($\eta^2 = .11$). Post hoc analyses with Bonferroni correction revealed significant differences between the beginning and upper-intermediate groups ($p = .001$), while no significant differences were observed between the other adjacent groups (As vs. B1 and B1 vs. B2). Overall, these findings suggest that as learners progress in proficiency, they increasingly incorporate subordinate clauses into their writing, reflecting greater syntactic complexity.

TABLE 3

Means and Standard Deviations of Syntactic Complexity Measures by Proficiency Group, with One-Way ANOVA Results

		Beginning (A2) (n = 28)		Intermediate (B1) (n = 44)		Upper-Intermediate (B2) (n = 56)		ANOVA	
Measure	Code	*M*	*SD*	*M*	*SD*	*M*	*SD*	*F*	*Sig*
Length of production unit									
Mean length of clause	MLC	9.29	1.38	9.35	1.36	9.47	1.29	0.20	0.815
Mean length of sentence	MLS	13.87	2.01	14.88	2.66	16.30	2.56	9.75	**<0.001***
Mean length of T-unit	MLT	13.30	1.79	14.41	2.48	15.71	2.32	11.15	**<0.001***
Amount of subordination									
T-unit complexity ratio	C/T	1.44	0.18	1.55	0.21	1.67	0.24	10.88	**<0.001***
Complex T-unit ratio	CT/T	0.38	0.12	0.43	0.12	0.48	0.15	5.24	**0.007***
Dependent clause ratio	DC/C	0.31	0.09	0.34	0.08	0.37	0.08	5.20	**0.007***
Dependent clauses per T-unit	DC/T	0.47	0.17	0.54	0.20	0.64	0.22	7.43	**0.001***
Amount of coordination									
Coordinate phrases per clause	CP/C	0.20	0.14	0.21	0.10	0.18	0.10	0.89	0.413
Coordinate phrases per T-unit	CP/T	0.28	0.19	0.32	0.14	0.30	0.19	0.31	0.734
Sentence coordination ratio	T/S	1.04	0.08	1.04	0.09	1.04	0.09	0.09	0.911
Degree of phrasal elaboration									
Complex nominals per clause	CN/C	0.97	0.22	1.00	0.22	1.05	0.29	1.04	0.356
Complex nominals per T-unit	CN/T	1.40	0.30	1.55	0.41	1.70	0.46	5.26	**0.006***

Note. *p*-values marked with an asterisk (*) indicate statistical significance at the Bonferroni-adjusted level of $p < .0167$.

Amount of Coordination

When examining the extent of coordination, the three coordination-related measures, CP/C, CP/T, and T/S, did not exhibit statistically significant differences across proficiency levels. The mean values of T/S, a sentential-level coordination index, remained virtually unchanged across the three groups (1.04 for all three groups). A mean value close to 1 suggests that most sentences contained only a single T-unit, indicating limited use of coordination through conjunctions such as *and*, *but*, or *so*.

Similarly, the mean values for CP/C (0.20 for beginning, 0.21 for intermediate, and 0.18 for upper-intermediate) and CP/T (0.28, 0.32, and 0.30, respectively) remained largely stable across groups. These consistently low values across all coordination measures imply that both sentential and clausal coordination are not particularly sensitive indicators of syntactic development among the Korean EFL university students examined in the present study. On balance, the results indicate that coordination, unlike subordination or phrasal elaboration, may offer limited insight into syntactic development in this learner population.

Degree of Phrasal Elaboration

With regard to phrasal elaboration, complex nominals per T-unit (CN/T) differed significantly across proficiency groups, $F(2, 125) = 5.26$, $p = .006$, with a small to moderate effect size ($\eta^2 = .07$). As the assumption of normality was violated for the upper-intermediate group ($p = .014$), nonparametric analyses were conducted. A Kruskal–Wallis H test revealed a statistically significant group difference, $H(2) = 8.29$, $p = .016$. Post hoc Mann–Whitney U tests with Bonferroni correction ($\alpha = .0167$) identified a significant difference between the beginning and upper-intermediate groups ($U = 490.5$, $p = .005$), while no significant differences were observed between the beginning and intermediate groups ($p = .164$) or between the intermediate and upper-intermediate groups ($p = .104$). These findings suggest that developmental differences in complex nominal usage were primarily evident between the beginning and upper-intermediate proficiency levels, whereas adjacent-level differences were not statistically reliable.

In contrast, complex nominals per clause (CN/C) did not differ significantly across groups, indicating that clause-level

nominal elaboration may be less sensitive to proficiency level than T-unit-based measures in this learning context.

DISCUSSION

The first research question sought to identify which specific syntactic features or measures differed across proficiency groups in Korean EFL learners' argumentative writing. Analyses of 12 syntactic complexity indices revealed distinct patterns of variation across proficiency levels. Beginning-level learners produced fewer words, used fewer subordinate clauses, and demonstrated limited use of nominal phrases. As proficiency increased, intermediate and upper-intermediate learners wrote longer texts and T-units and increasingly incorporated subordinate clause structures into their essays. Additionally, upper-intermediate writers began to employ more compact nominal phrases, such as attributive adjectives (e.g., *individual academic* success) and postnominal phrases (e.g., the amount *of work to finish*).

Building on these group-based differences, the second research question explored how these patterns reflect broader developmental trends in syntactic organization. The results suggest a gradual progression from dynamic, clause-driven to synoptic, phrase-based constructions, consistent with developmental patterns propounded in Systemic Functional Linguistics (Norris & Ortega, 2009). While subordination increased steadily across proficiency levels, phrasal elaboration emerged more prominently only at the upper-intermediate level, indicating that the transition toward synoptic structures is underway but not yet fully established. Interestingly, the early reliance on coordination often reported in developmental models was not evident among these adult EFL writers. To further elucidate how this developmental shift unfolds across proficiency levels, each of the four principal dimensions of syntactic complexity, namely, length, subordination, coordination, and phrasal elaboration, is examined in turn.

For the global index of syntactic complexity, the results of the length-based measures provide support for their robustness as reliable indicators, corroborating previous findings (Lu, 2011; Ortega, 2003; Wolfe-Quintero et al., 1998); the writers at higher proficiency levels produced more words per T-unit. This pattern is reflected in the mean values of mean length of T-unit (MLT): the Korean EFL writers at the upper-intermediate level produced, on average, 15.71 words per T-unit, compared to 14.41 at the intermediate level and 13.30 at the beginning level. The mean difference of approximately 2.41 words between the beginning and upper-intermediate groups was statistically significant, while the 1.31-word difference between the intermediate and upper-intermediate groups was only marginally significant. In contrast, the smaller 1.11-word difference between the beginning and intermediate groups did not reach significance. This result echoes findings from Ortega's (2003) synthesis of 19 studies, which reported the mean differences of 2 or more words per T-unit are typically associated with significant developmental gains in L2 writing. It is also comparable with the results of J. Kim (2014), who examined argumentative essays by Korean university EFL writers and found statistically significant MLT differences of approximately 2.5 words between adjacent proficiency groups (basic–intermediate and intermediate–advanced), with figures that closely resemble those observed in the present study.

The study also reaffirmed that subordination serves as an informative indicator of syntactic development, particularly in characterizing writing at the upper-intermediate proficiency level (Wolfe-Quintero et al., 1998). The mean values of dependent clauses per T-unit (DC/T) increased steadily across the three proficiency levels: on average, beginning-level writers produced 4.7 dependent clauses per 10 T-units, intermediate learners 5.4, and upper-intermediate learners 6.4. The finding that DC/T significantly distinguished the upper-intermediate group from the beginning group is consistent with empirical findings that subordination increases across beginning to intermediate levels and continues to rise through upper-intermediate levels, where it plays a major role in the development of sentence complexity (Byrnes et al., 2010).

While the upper-intermediate group produced more dependent clauses per T-Unit (M = 0.64) compared to those at the intermediate level (M = 0.54), the difference was not statistically significant. This pattern may reflect a trend noted in some studies suggesting that the increase in subordination becomes less pronounced at higher levels (Vyatkina, 2013). It may indicate that upper-intermediate writers are potentially nearing a plateau for this feature or beginning to rely more on other forms of syntactic complexity, such as phrasal elaboration. However, this interpretation remains tentative for two key reasons. First, the present study did not include advanced-level writers. Since a decline or plateau in subordination is most commonly observed at advanced stages, more definitive conclusions would require data from learners at higher proficiency levels. Second, as noted in the literature review, studies analyzing argumentative essays by Korean EFL university students (e.g., Bae & Min, 2020; J. Kim, 2014) have reported a consistent and significant increase in the use of subordinate clauses from beginning to intermediate and even into higher proficiency levels-despite contrasting findings in other L2 contexts where subordination tends to decline at more advanced stages (e.g., Byrnes et al., 2010). One interpretation of this pattern is that Korean EFL learners' reliance on subordinate clauses, even at higher proficiency levels, may reflect typological

transfer from Korean, a language in which clausal subordination is both frequent and structurally flexible (Choi, 2020). To capture a fuller picture of how subordination functions in Korean EFL learners' argumentative writing, further research involving a broader range of proficiency levels and cross-linguistic comparisons is needed.

In addition, while the L2 Syntactic Complexity Analyzer (L2SCA) is a valuable tool for examining broad patterns of syntactic development, it is limited in its ability to provide fine-grained analyses of subordinate structures. Subordination in English comprises three primary structures: nominal clauses, adverbial clauses, and relative clauses. Empirical evidence suggests that these structures emerge at different stages of development. Nominal clauses tend to appear earliest, owing to their syntactic simplicity and high frequency in learner input, followed by adverbial clauses, which require more complex, discourse-level reasoning. Relative clauses, by contrast, are structurally more demanding and typically emerge later in development (Biber & Gray, 2016; Nippold et al., 2005; Ortega, 2003). On this basis, Lambert and Kormos (2014) caution researchers against treating subordination as a unified category, as doing so may obscure important developmental distinctions among different types of subordinate clauses. In line with this view, a preliminary manual inspection of subordination types revealed that the Korean EFL writers in this study predominantly produced adverbial and nominal clauses, while relative clauses were notably infrequent across all proficiency levels. Therefore, a future study drawing on a dataset similar to the current one, with a substantial number of intermediate and upper-intermediate L2 writers, would be well-served by employing a more fine-grained analysis of subordinate structures to further illuminate developmental patterns of subordination (e.g., Staples et al., 2016).

Turning to coordination measures, coordination has often been regarded as a salient indicator of early syntactic development, as L2 writers with limited syntactic resources tend to link ideas using relatively simple coordinating conjunctions such as *and*, *but*, or *so*, after which their use is expected to decline with increasing proficiency (Bardovi-Harlig, 1992; Norris & Ortega, 2009). Contrary to this expectation, the present study did not find empirical evidence supporting this pattern. Its findings revealed neither a clear downward trend nor statistically significant differences across proficiency groups, suggesting that coordination may not serve as a reliable developmental marker in the context of argumentative writing. In particular, the mean values of the sentential coordination measure (T/S), which hovered close to 1 across all groups, indicate that Korean university writers produced few coordinated compound sentences in their essays, regardless of proficiency level.

This outcome may be shaped, at least in part, by L1 influence. In Korean writing, coordinating conjunctions such as *and*, *but*, or *so* typically appear at the beginning of a sentence to connect two independent sentences, rather than linking two clauses within a single compound sentence. This structural preference may shape how Korean learners approach coordination in English, leading to less frequent and limited use of coordinated structures, even at early stages of English writing. This structural preference, well-documented in Kang (2008), may lead Korean EFL learners to underuse coordinated structures in English or favor sentence-initial conjunctions that do not form compound sentences. Such L1 influence may help explain the absence of statistically significant differences in sentential coordination reported in Bae and Min (2020) and J. Kim (2014), both of which analyzed argumentative essays written by Korean EFL university students. Clearly, further empirical research involving L2 writers from diverse L1 backgrounds, particularly those at the very early stages of general language proficiency, is needed to better understand the developmental trajectories of coordination in L2 syntactic complexity.

The results concerning phrasal complexity, which reflects the ability to employ structurally compressed phrases, showed that complex nominals began to emerge gradually in the writing of Korean EFL learners at the upper-intermediate level (e.g., Biber & Gray, 2013; Kyle & Crossley, 2018). This study found that complex nominals per T-unit (CN/T) increased progressively across proficiency levels and significantly distinguished beginning-level writers from upper-intermediate writers. As an illustration, upper-intermediate students produced more complex nominals per T-unit (e.g., *Working together on assignments* can improve *the students' cooperating skills* and reduce *the amount of work per student*) than beginning-level learners (e.g., *Collaborating in group work* can upgrade *communication capacity*). However, the small to moderate effect size for CN/T ($\eta^2 = 0.07$), along with the absence of statistically significant differences between the adjacent proficiency groups (A2 vs. B1 and B1 vs. B2), suggests that phrasal elaboration may only be beginning to emerge, rather than reflecting a strong command of these structures.

On a final note, the study found that CN/T effectively distinguished the beginning and upper-intermediate groups, whereas complex nominals per clause (CN/C) did not (see Ali & Lu, 2013, for a similar result). This implies that increases in the density of complex nominals are more reliably detected when measured against the T-unit (i.e., a main clause plus any associated dependent clauses) than against individual clauses (i.e., both main and subordinate). These findings point to the possibility that CN/T may serve as a more sensitive index of early-stage phrasal complexity, particularly for EFL learners who are still in the process of developing clause-internal elaboration (Kyle & Crossley, 2016; Lu, 2011).

CONCLUSION

The present study sought to inform improvements in a university-level writing curriculum for foreign language learners, most of whom aim to communicate effectively in written discourse that meets the expectations of academic and professional communities (Byrnes et al., 2010). A central concern, therefore, is how to support both learners and instructors in developing a foundational understanding of how syntactic elements are structured in English, and in building the capacity to make informed choices about sentence construction-an essential component of the course instruction. Equally important is the ability to predict and understand learners' syntactic patterns in their written texts, so as to provide feedback that is tailored to the specific strengths and weaknesses of their writing.

Viewed in this pedagogical context, the classroom-based design of the present study offers practical implications regarding the syntactic complexity of argumentative essays produced by Korean university EFL learners. As learners' proficiency increases, they tend to produce longer texts and longer production units (i.e., sentences and T-units) while demonstrating greater control over subordinate clause structures. Beyond developmental considerations, the prominence of subordinate clauses at the intermediate levels may, in part, reflect the effects of focused instruction, as the writing course is designed for intermediate-level learners and explicitly emphasizes subordinate clauses, including nominal, adverbial, and relative clauses. At this stage, the upper-intermediate learners gradually enhance their phrasal complexity, particularly through the increased use of complex nominal phrases, although their skills in this area are not yet fully developed. This finding is noteworthy, as classroom instruction rarely addresses features related to phrasal complexity. To support these learners more effectively, pedagogical interventions that incorporate explicit instruction on phrasal elaboration, such as complex noun phrases, remodifying structures, and postnominal modifiers, may help promote the development of more advanced syntactic constructions.

To conclude, this exploratory study comes with several limitations that should be acknowledged and addressed in future research. The first limitation concerns the temporal gap between the administration of the proficiency test and the collection of writing samples. While proficiency scores were obtained during the first week of the semester, the argumentative writing tasks were completed in the fifteenth week. Most students had hardly any exposure to the target language outside the classroom and were not concurrently enrolled in other English courses. Still, it is possible that their writing instruction during the semester may have contributed to improvements in their overall English proficiency. Consequently, the initial test scores may not have accurately reflected their proficiency at the time of writing. This particular writing task was selected because, among the three writing tasks completed during the semester, it was the only one that required students to write under time constraints and without access to external resources such as dictionaries**.** To mitigate this limitation, a follow-up study is planned to analyze essays collected during the early weeks of the semester.

A second shortcoming relates to the sample size. As an initial attempt to examine the syntactic characteristics of first-year college students' writing, the present study included a relatively small number of essays, which may limit the generalizability of the findings. In particular, the small number of beginning-level learners may constrain a more comprehensive understanding of the syntactic complexity features that are characteristic of those at the early stages of L2 development. Nevertheless, the observed distribution of students across proficiency levels reflects what is typically seen in competitive urban universities, where intermediate- or upper-intermediate-level learners tend to be more heavily represented than beginning-level learners. In this regard, the sample supports the ecological validity of the study.

Finally, the study grouped learners into three proficiency levels based on a standardized English proficiency test rather than holistic evaluations of their writing. As there have been relatively few studies investigating the relationship between syntactic complexity and L2 writers' general language proficiency (but see Bae & Min, 2020), the present findings may offer insights into how grammatical development unfolds across broader measures of proficiency, not just writing-specific ability. Even so, the study could have been strengthened by grouping students according to their actual writing performance, which could have yielded a more direct comparison of syntactic complexity in relation to writing ability. As noted earlier, participants' general proficiency test scores were significantly correlated with their holistic writing scores, as measured by IntelliMetric and by instructor ratings. That said, future research would benefit from further examining how general language proficiency and writing-specific performance independently and jointly contribute to the development of syntactic complexity in L2 writing.

Acknowledgements

I sincerely thank the two anonymous reviewers for their insightful feedback and constructive suggestions, which greatly

improved the clarity and overall quality of this manuscript. Any remaining errors are my own.

I'm also grateful to IntelliMetric® for providing automated essay scoring support, which offered useful perspectives on the holistic assessment of student writing.

References

Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. In A. Díaz-Negrillo, N. Ballier, & P. Thompson (Eds.), *Automatic treatment and analysis of learner corpus data* (pp. 249-264). John Benjamins.

Bae, Jiyoung, & Min, Sujung. (2020). Genre-based analysis of syntactic complexity in L2 college students' writing: Pedagogic scope and directions. *Journal of Asia TEFL, 17*(3), 937-953.

Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly, 26*(2), 390-395.

Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing, 22*(2), 185-200.

Bi, P., & Jiang, J. (2020). Syntactic complexity in assessing young adolescent EFL learners' writings: Syntactic elaboration and diversity. *System, 91*, Article 102248. https://doi.org/10.1016/j.system.2020.102248

Biber, D., & Gray, B. (2013). *Discourse characteristics of writing and speaking task types on the TOEFL iBT® test: A lexico-grammatical analysis* (ETS Research Report No. RR-13-04). Educational Testing Service

Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. 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.

Biber, D., Gray, B., Staples, S., & Egbert, J. (2020). Investigating grammatical complexity in L2 English writing research: Linguistic description versus predictive measurement. *Journal of English for Academic Purposes, 46*, Article 100869. https://doi.org/10.1016/j.jeap.2020.100869

Brown, R. (1973). *A first language: The early stages*. Harvard University Press.

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (Vol. 32, pp. 21-46). John Benjamins.

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

Byrnes, H., Maxim, H. H., & Norris, J. M. (2010). Realizing advanced foreign language writing development in collegiate education: Curricular design, pedagogy, assessment. *Modern Language Journal, 94*(2), 227-235.

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.

Choi, Susie. (2020). Development of clause chaining in Korean. *Frontiers in Psychology, 11*, Article 256. https://doi.org/10.3389/fpsyg.2020.00256

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299-323.

Halliday, M. A. K., & Matthiessen, C. (1999). *Construing experience through meaning: A language-based approach to cognition*. Cassell.

Hulstijn, J. H., Alderson, J. C., & Schoonen, R. (2010). Developmental stages in second-language acquisition and levels of second-language proficiency: Are there links between them? In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (Eurosla Monographs Series, No. 1, pp. 11-20). European Second Language Association.

Hunt, K. W. (1970). Syntactic maturity in schoolchildren and adults. *Monographs of the Society for Research in Child Development, 35* (1, Serial No. 134).

Jendraschek, G., & Shin, Y. M. (2011, September). *Suffixal clause-linking in Korean on the subordination-coordination continuum*. Paper presented at the 44th Annual Meeting of the Societas Linguistica Europaea, Logroño, Spain.

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,* Article 100666. https://doi.org/10.1016/j.jslw.2019.100666

Kang, Yusun. (2008). The use of cohesive conjunctions in Korean EFL learners' written narrative discourse. *Studies in English Education, 13*(2), 97-122.

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.

Khushik, G. A., & Huhta, A. (2022). Syntactic complexity in Finnish-background EFL learners' writing at CEFR levels A1-B2. *European Journal of Applied Linguistics, 10*(1), 142-184.

Kim, Ji-young. (2014). Predicting L2 writing proficiency using linguistic complexity measures: A corpus-based study. *English Teaching, 69*(4), 3-30.
Kim, Minjung. (2023, July). *Linking an English language test (G-TELP) to the CEFR: A comparison of Modified Angoff and Bookmark methods*. Paper presented at the 21st Asia TEFL Conference, Seoul, South Korea.
Kim, Susie. (2021). Generalizability of CEFR criterial grammatical features in a Korean EFL corpus across A1, A2, B1, and B2 levels. *Language Assessment Quarterly, 18*(3), 273-295.
Kyle, K., & Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *Modern Language Journal, 102*(2), 333-349.
Lambert, C., & Kormos, J. (2014). Complexity, accuracy, and fluency in task-based L2 research: Toward more developmentally based measures of second language acquisition. *Applied Linguistics, 35*(5), 607-614.
Larsen-Freeman, D. (1978). An ESL index of development. *TESOL Quarterly, 12*(4), iii-67.
Larsen-Freeman, D. (2013). Complexity theory: A new way to think. *Revista Brasileira de Lnguística Aplicada*, *13*, 369-373.
Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15*(4), 474-496
Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45*(1), 36-62.
Lu, X. (2014). *Computational methods for corpus annotation and analysis*. Springer.
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.
Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for academic purposes programme. *Journal of Second Language Writing, 29*(1), 3-15.
Neff, J., Dafouz, E., Díez, M., Prieto, R., & Chaudron, C. (2004). Contrastive discourse analysis: Argumentative text in English and Spanish. In C. L. Moder & A. Martinovic-Zic (Eds.), *Discourse across languages and cultures* (pp. 267-283). John Benjamins.
Nippold, M. A., Hesketh, L. J., Duthie, J. K., & Mansfield, T. C. (2005). Conversational versus expository discourse: A study of syntactic development in children, adolescents, and adults. *Journal of Speech, Language, and Hearing Research, 48*(5), 1048-1064.
Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*(4), 555-578.
Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*(4), 492-518.
Ortega, L. (2012). Interlanguage complexity: A construct in search of theoretical renewal. In B. Kortmann & B. Szmrecsanyi (Eds.), *Linguistic complexity: Second language acquisition, indigenization, contact* (pp. 127-155). De Gruyter.
Park, Seon-yeong. (2012). A corpus-based study of syntactic complexity measures as developmental indices of college-level L2 learners' proficiency in writing. *Korean Journal of Applied Linguistics, 28*(3), 139-160.
Rudner, L. M., Garcia, V., & Welch, C. (2006). An evaluation of IntelliMetric™ essay scoring system. *The Journal of Technology, Learning, and Assessment, 4*(4), 1-30.
Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication, 33*(2), 149-183.
Vyatkina, N. (2012). The development of second language writing complexity in groups and individuals: A longitudinal learner corpus study. *Modern Language Journal, 96*(4), 576-598.
Vyatkina, N. (2013). Specific syntactic complexity: Developmental profiling of individuals based on an annotated learner corpus. *Modern Language Journal, 97*(Supplement 1), 11-30.
Wolfe-Quintero, K., Inagaki, S., & Kim, Hae-Young. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity* (Technical Report No. 17). Second Language Teaching & Curriculum Center, University of Hawai'i at Mānoa.
Yin, S., Gao, Y., & Lu, X. (2021). Syntactic complexity of research article part-genres: Differences between emerging and expert international publication writers. *System, 97*, Article 102427. https://doi.org/10.1016/j.system.2020.102427
Yoon, Hyung-Jo. (2017). Linguistic complexity in L2 writing revisited: Issues of topic, proficiency, and construct multidimensionality. *System*, *66*, 130-141.
Yoon, Hyung-Jo, & Polio, C. (2016). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly, 50*(2), 275-301.
Yoon, Hyung-Jo, & Polio, C. (2017). The linguistic development in the study abroad context: A synthesis of recent research. *Language Teaching, 50*(3), 335-349.
Zheng, Y., & Barrot, J. S. (2024). Syntactic complexity in second language (L2) writing: Comparing students' narrative and argumentative essays. *System, 123*, Article 103342. https://doi.org/10.1016/j.system.2024.103342