

# Exploring syntactic complexity and its relationship with writing quality in EFL argumentative essays

Elias Wondemagegne<sup>1,\*</sup>, Birhanu Simegn<sup>2</sup>, Ebabu Tefera Adugna<sup>3</sup>

<sup>1</sup>Doctoral candidate at Bahir Dar University

<sup>2</sup>Associate professor in the Department of English Language and Literature, at Bahir Dar University

<sup>3</sup>Associate professor in the Department of English Language and Literature, at Bahir Dar University where he teaches and advises postgraduate students. His research interests include teaching writing and related communicative skills.

## Abstract

Despite a large number of studies examining syntactic features that are predictive of second language (L2) writing quality, assessed by human raters at the university level, few have systematically investigated this link using a large set of indices in the foreign language learning (EFL) classroom context. The current study sought to determine the extent to which a variety of syntactic complexity and sophistication indices are associated with and may predict writing quality by analyzing 30 argumentative essays written by undergraduate EFL students in an Ethiopian university classroom setting. To represent syntactic complexity as a multidimensional construct, we used conventional absolute measures, fine-grained clausal and phrasal indices, and newly proposed sophistication indices related to the use of verb argument constructions (VACs) indexed by TAASSC (Tool for the Automatic Analysis of Syntactic Sophistication and Complexity; [17]). Essays were graded, and five separate predicted models of writing quality were created utilizing each complex feature index and all of the measures. Robust predictors of writing quality were identified at both syntactic complexity and sophistication dimensions. Regression analyses showed that the combined model including both fine-grained clausal complexity and VAC-based indices could account for 53.6% of the variance (the largest amount of variance in the study) in writing scores. The finding indicates that the inclusion of diversified adverbial modifiers and nonfinite clauses such as modal auxiliaries controlled by less frequent verbs were predictive of higher-quality writing. These findings shed light on some characteristics of L2 learners' writing growth and enable us to draw pedagogical implications for teaching and assessing writing in the Ethiopian EFL context.

## Research Article

## Open Access & Peer-Reviewed Article

## Corresponding author:

Elias Wondemagegne, Doctoral candidate at Bahir Dar University

## Keywords:

Syntactic complexity, Syntactic sophistication, Verb argument constructions, EFL writing quality

**Received:** February 18, 2024

**Accepted:** May 10, 2024

**Published:** May 16, 2024

## Academic Editor:

Hossein Salarian, University of Tehran, Iran

## Citation:

Elias Wondemagegne, Birhanu Simegn, Birhanu Simegn (2024) Exploring syntactic complexity and its relationship with writing quality in EFL argumentative essays. Journal of Language Research - 1(2):13-35.

## Introduction

### *Background information and statement of the problem*

Over the last two decades, L2 writing scholars in the field of instructed second-language acquisition (ISLA) have taken up investigating the elaborated

degree of linguistic features in L2 learners' written output to better understand learners writing competence and L2 growth from a linguistic and psycholinguist perspective [9; 29]. Syntactic complexity refers to the variety and sophistication of syntactic features exhibited in L2 outputs, including the length of production units i.e., the number of words per clause, the amount of subordination, or the frequency of distinct clauses, or phrase structures used [35]. Per se, analyzing L2 output in terms of indices measuring these syntactic features is a common means to gauge L2 growth or writing quality. The default hypothesis has been that as L2 learners' interlanguage develops, they begin to produce less frequent and diverse syntactic features [5; 27]. That said, L2 learners' language-using ability or writing quality is demonstrated through their control over complex and sophisticated levels of syntactic knowledge (Li & Yang, 2023) [34].

Presuming this, investigating the relationship between L2 writing quality human ratings writing scores and syntactic complexity indices has been the focus of much L2 writing research [2; 23]. This focus has been motivated by at least three reasons conceptually, researchers are interested in identifying features that both adequately reflect the multidimensional nature of syntactic complexity and strongly correlate with quality ratings of L2 writing [5]. From the assessment perspective, this line of research has informed the empirical development of writing proficiency scales and facilitated automated scoring [15]. Pedagogically, a detailed understanding of this relationship can offer useful insights into the aspects of syntactic complexity to focus on for different learner groups, genres, and writing tasks (Bulté & Housen, 2019). Additionally, this focus has been motivated by the advent of computational tools such as L2 Syntactic Complexity Analyzer (L2-SCA) [21] has made it possible to examine this relationship on a larger scale quickly and accurately [4; 24; 30].

Despite recent advances, existing L2 research on syntactic complexity has several theoretical, conceptual, and methodological flaws [3] (Housen, De Clercq, Kuiken, & Vedder, 2019). As an illustration, syntactic complexity has traditionally been conceived as the basic internal formal structure of the text [6] and measured using only large-grained complexity indices such as clausal subordination and T-unit (consisting of one main clause plus extra, embedded clauses, Hunt, 1965) which have high predictive power [35]. However, in recent years, the use of these absolute complexity, large-grained, indices has come under criticism for failing to account for syntactic variation, the specific subtypes of clausal and phrase level elaboration that emerge as L2 learners advance (Housen et al., 2019) and for the difficulty in interpretation [5]. Recent studies [30] have indicated that fine-grained indices that capture specific subtypes of phrasal or clausal structures (e.g., clause complements per clause) are better indicators of holistic scores of TOFEL independent essay quality [19].

Another criticism is that the traditional absolute complexities do not account for the relative frequency of syntactic forms and the entangled link between syntactic forms and lexical items [17]. That said, the absolute complexity indices do not strongly coincide with some aspects of L2 acquisition theories, such as usage-based approaches, which posit that linguistic constructions emerge from language use, such that, features related to the frequency, saliency, and contingency of syntactic constructions, not absolute complexity, is the main indicator of L2 acquisition [11; 13]. Aligned with the theory of usage-based approach, some scholars most notably Kyle (2016) proposed features related to the use of Verb Argument Constructions (VACs, which consist of a verb slot and its related arguments) as a reliable method of measuring relative complexity or syntactic sophistication [24] (Li & Yang, 2023) [30]. These studies have confirmed that the newly proposed VACs sophistication indices are better indicators of L2 writing quality. However, the majority of L2 writing studies in this line of inquiry have primarily

focused on argumentative, and narrative essays written by advanced or college-level L2 learners in large-scale standardized tests (e.g., TOEFL) in the English-dominant contexts [19; 36]. Thus, it is unclear as to what absolute and fine-grained along with VACs sophistication features capture the writing quality and syntactic developmental trajectories of Ethiopian EFL. Moreover, little is known about the link between syntactic features and EFL course instructors' holistic ratings of genres commonly assessed in EAP classrooms context. Therefore, the present study then tapped into what specific syntactic features predict EFL teachers' holistic rating score of argumentative essays produced by undergraduate EFL learners in their classroom context using computational tools namely TAASSC [17]. As such, exploring the relationship between L2 writing quality and syntactic features of argumentative essays produced by this group of learners enables us to obtain useful information regarding their linguistic competence and repertoires, helping to create a more comprehensive picture of the writing performance of EFL learners at different stages of L2 learning.

The overall structure of this study consists of five sections, including the introductory part. In the following, section two presents a brief overview of syntactic complexity as a multi-dimensional construct by laying out the theoretical dimensions and looking at how syntactic complexity is evaluated in L2 acquisition. The third section is concerned with the methodology used for this study. The fourth section presents the analysis and results of the study. Finally, the conclusion gives a summary as well as the implication of the findings to pertinent future studies.

#### *Syntactic complexity as a multi-dimension construct*

Recognizing the multidimensional nature of syntactic complexity, Ortega (2015) defined it as the "expansion of the capacity to use the additional language in ever more mature and skillful ways, tapping the full range of linguistic resources offered by the given grammar to successfully fulfill various communicative goals" [29]. Her definition highlights that the term "syntactic complexity" refers to the ability to generate longer and more complex structures (i.e., absolute complexity) as well as knowledge of how to skillfully leverage the complexity and subtleties of language [6; 17]. This implies that these two dimensions are critical in comprehending and analyzing syntactic complexity as a multidimensional construct [9], and this viewpoint is taken into account for the current study.

#### *Syntactic Complexity*

The first construct, "syntactic complexity," as defined in this study, pertains to the hierarchical organization and nesting of the basic formal structures of language units (e.g., phrases and clauses) within a sentence (Housen et al., 2019). Measuring this dimension helps to demonstrate the extent to which EFL/ESL learners can manipulate and combine diverse grammatical patterns to build complex linguistic systems as their interlanguages develop. Given this, in ISLA literature, various indices have been proposed to quantify syntactic complexity as a multidimensional entity, [22], which can be summarized into two approaches: absolute complexity indices [21] and fine-grained indices [17]. The former assesses the internal complexity of the text in a largely holistic way, without differentiating specific subtypes of the structures concerned [5]. Several L2 writing studies [8; 21; 37] assess this dimension of complexity using L2 Syntactic Complexity Analyzer. Whereas the latter, fine-grained indices, take into account specific subtypes of phrasal (e.g., prepositions per direct object) or clausal structures (e.g., clause complements per clause, which taps into the use of a particular type of dependent clause). Several fine-grained indices and computational tools that measure these indices have been proposed and developed by different scholars [3; 17] and employed in L2 writing studies. Of these fine-grained syntactic complexity indices operationalized by Kyle (2016) and computed in TAASSC

which have had an impact on recent discussions of clausal and phrasal complexity are then considered in the current study [19] (Li & Yang, 2023) [30].

#### *Syntactic Sophistication*

The second construct, "syntactic sophistication," relates to the acquisitional sequence of specific sets of linguistic construction including syntactic constructions and the mental processes required for producing, and breaking them down [7; 17]. The basic premise is that syntactic construction (e.g., VACs) learned later are more sophisticated [18]. It takes its theoretical cues from the usage-based approach, which maintains that L2 acquisition is a process of contingency learning in which L2 learners create form-meaning pairings through repeated exposure to the target language constructions (e.g., words) using cognitive processes like entrenchment, categorization, association, and generalization [12]. This suggests that grammatical constructions that emerge from the interaction of language input and cognitive processes carry meaning independent of specific lexical items within the sentence. For instance, the ditransitive construction consistently forms (i.e., subject-verb-indirect object-direct object) structure carries the meaning of transferring something from one entity to another as in the sentence "she got/give/kicked/ me a ball," regardless of the verbs (Park & Sung, 2022). These sentence or clause-level pairings between form and meaning are referred to as Verb Argument Constructions (VACs) and are analyzed as independent syntactic constructions in English and other languages [14; 17].

According to the usage-based view, the acquisition of VACs as a type of construction is subject to multiple psycholinguistic factors [12], among which input frequency is the most important driving factor. Specifically, VACs with relatively high frequency in input have a better chance to entrench in the learner's memory and are more likely to be acquired earlier and deemed as less complex, while VACs with low input frequency are considered to be acquired in later stages and thus more complex [13; 33].

Another key factor predicting VAC acquisition is the contingency of mapping between a main verb and the VAC in which it occurs [12]. It refers to the degree of reliability to which a given cue (a particular verb) can predict a certain outcome (VAC) or in reverse, how likely a verb will occur given a certain VAC [18]. Previous research has shown that L2 learners employ verbs that are less strongly associated with a target construction as their proficiency increases [18; 9]. The findings of these studies lent support for the claim that closely associated verb-VAC acquired early, and are thus regarded as less complex, while the more unusual combinations are deemed more difficult to acquire and thus more complex. For example, Kyle and Crossley (2017) found that higher-scoring L2 argumentative essays contained less frequent and more strongly associated verb-VAC combinations (Li & Yang, 2023). In keeping with the taxonomy in recent studies [18; 19; 22], we measured syntactic complexity using traditional absolute complexity indices and fine-grained complexity indices that differentiate structural subtypes of clauses and phrases, and syntactic sophistication using VAC-related features, i.e., frequency and contingency features, ( see Tables 1-4 in the methodology section). In what follows, we provide a systematic review of previous studies that have examined the relationship of L2 writing quality to these different dimensions of syntactic complexity features.

#### *Syntactic complexity and Writing quality*

A number of cross-sectional L2 studies have investigated the relationship between writing quality and different measures of syntactic complexity [1; 8; 36]. Most of these studies have relied heavily on

various large-grained indices that quantify the average length of a certain linguistic unit (e.g., mean length sentences (MLS), mean length T-unit (MLT), mean length clauses (MLC), etc.) and the amount of subordination such as dependent clauses per clause (DC/C) and clauses per T-unit (C/T) [9]. Based on the notion that academic writing derives its complexity from the elaborate use of clausal constructions or clausally complex, with many subordinate clauses (Hyland, 2002).

In their synthesis and meta-analyses, Ortega (2003) and Wolfe-Quintero et al. (1998) noted that the majority of these investigations, as they (1988) reported 23 studies out of 40 studies, discovered a largely consistent positive link between global complexity features and overall assessment of composition quality. To illustrate, among the earliest studies, Homburg (1984) investigated the relationship between a holistic evaluation of ESL writing quality based on the Michigan Test of English Language Proficiency grading scheme (MTELP) and 10 measures of syntactic complexity and found a significant relationship between MLS, MLC, finite clausal subordination (DC/C), and writing score. More recently, to identify key linguistic features on the writing performance of EFL learners, Chuenchaichon, (2022) analyzed a collection of opinion essays written by Thai EFL university students from CU-TEP (The Chulalongkorn University Test of English Proficiency) corpus and found significant correlations between MLT, MLC, MLS and the quality of argumentative essays score. That said, essays with longer T-units and higher amounts of subordination received higher scores and were considered quality text. This finding is also corroborated by a few longitudinal research. For instance, Bulté and Housen (2014) reported longitudinal growth in MLT scores throughout a semester-long English for Academic Purposes (EAP) L2 writing course [20].

In addition, in search of more global indices, Lu (2011) included 14 measures of syntactic complexity recommended by Ortega (2003) and Wolfe-Quintero et al. (1988) and developed an automated analyzing tool, the L2-SCA, based on large-scale data gathered from the Written English Corpus of Chinese Learners (WECCL). He found that seven out of 14 measures progressed linearly across levels of L2 writing quality these were the mean length of MLT, MLC, MLS, and coordinate phrases per clause (CP/C), and T-Unit (CP/T), complex nominals per clause (CN/C) and per T-unit (CN/T). Moreover, Lu found that contextual factors like institution (i.e., the universities the learners attended), timing condition (timed versus untimed), and genre (argumentative versus narrative) have an impact on the linguistic characteristics of written products in addition to the proficiency level of L2 learners. According to his findings (2011), argumentative writing has lengthier units and is more complicated at the phrasal level than narrative essays [37] or application letter writings [38]. Subsequent research has tested the reliability of the 14 measures of absolute complexity, large-grained, indices provided by L2-SCA, with a wide range of variables such as topics [36], instructional setting [28] that affect the production of complexity.

Yang et al. (2015) used seven indices from L2-SCA to analyze a corpus of argumentative essays on two topics written by ESL graduate students, scored using the TOEFL iBT independent writing rating rubric, and found that MLT significantly predicted essay scores on various topics. In another study, using 10 indices from L2-SCA, Ai and Lu (2013) examined the differences in syntactic complexity in the writing of native speakers (NS) and non-native speakers (NNS) of English, showed that NNS produced shorter clauses, sentences, and T-units, less subordination, and fewer noun phrases than do NS. This shows that EFL writers are characterized by distinct syntactic structures from L1 or L2 learners [23]. Taken together the aforementioned L2 writing studies have indicated a positive relationship between large-grained indices of syntactic complexity and writing quality.



Despite the ubiquity and usefulness of such structure-based measures (MLT, and DC/C) recently, however, a number of scholars have questioned their use to investigate syntactic complexity in L2 writing [5; 18]. The first, limitation is that absolute complexity measures mostly capture the degree of elaboration, but give less attention to the degree of variation, in other words, syntactic diversity [27]. Second, [3] argue, with corpus-based evidence, that clausal elaboration assessed by T-unit indices is more characteristic of conversation whereas academic writing is characterized syntactically by the use of noun complex phrase constituents, particularly noun phrase complexity, thus they recommend phrase-level structures as more reliable measures for L2 writing in an academic setting.

Subsequent cross-sectional and longitudinal research (e.g., Biber et al., 2016) [34; 38] has indicated that phrasal complexity is a better predictor of writing quality scores than clausal subordination. For instance, Kyle and Crossley (2018) reported that fine-grained phrasal complexity indices (related to nominal subject, direct object, and prepositional object modifiers) accounted for a larger proportion of variance in essay scores than both holistic syntactic complexity indices and fine-grained clausal complexity indices. Similar results were reported by Zhang and Lu (2022) in their analysis of a collection of application letters and argumentative essays produced by college-level Chinese EFL learners. Similarly, Taguchi et al. (2013) analyzed a collection of argumentative essays written by a number of non-native speakers of English in a university in the United States using ten syntactic complexity indices provided by Biber tagger and found that noun phrase modification features such as attributive adjectives, and post modifying prepositional phrases, contributed to essay quality. Demonstrating that as the academic level increased, so did the use of phrasal complexity features in writing. This finding is also corroborated by some longitudinal research, in which Crossley and McNamara (2014) discovered that high-quality university L2 learners' descriptive texts included more complex noun phrases and fewer embedded clauses.

Although the above reviewed, studies have revealed a positive relationship between different dimensions of syntactic complexity and L2 writing quality, this model is based on considering syntactic structure only and is difficult to link to the frequency of occurrence of form-meaning mappings found in usage-based perspectives on language learning. In complementing this gap, Kyle (2016) by developing and validating TAASSC proposed a range of usage-based VAC indices related to frequency and contingency as a new way of representing syntactic sophistication (M. Abdi Tabari et al., 2023). This proposal inspired a new wave of research on comparing absolute versus relative complexity measures in terms of their predictive power of L2 writing quality [18; Li & Yang, 2023; 30]. This line of research has generated evidence for greater predictive power of VACs-based complexity measures over absolute ones. As an instance, using a corpus of TOEFL independent and integrated essays as well as descriptive essays selected from the Michigan State University corpus, Mostafa and Crossley (2020), found the stronger predictive power of usage-based VAC indices, as these measures explained more variance in L2 writing particularly in the argumentative task than in descriptive and integrated tasks (sources based writing).

Similarly, M. Abdi Tabari et al., (2023) examined the extent to which a number of VAC-based syntactic sophistication indices could predict ESL learners' argumentative writing quality in different writing task conditions and found that only verb-VAC combinations (contingency-related features) were the most significant predictor of syntactic sophistication for all groups. In another study, Li and Yang (2023) compared indices of absolute complexity and VAC sophistication in terms of their ability to predict the writing quality of the overall as well as major sections of English research articles (RA) written by

Chinese Ph.D. students. The finding indicated that VAC-based measures are more useful for indexing the writing quality of RAs at the whole text level while absolute measures have a stronger predictive power at the part-genre level, pointing out that both holistic and fine-grained indices usefully complement each other in capturing the syntactic complexity of L2 production.

Three important observations can be made in the body of research reviewed above. First, the majority of studies of the relationship of syntactic complexity to L2 writing quality have primarily focused on advanced and college-level L2 learners' writing for large-scale standardized test samples (e.g., TOEFL) in the English-dominant context, without taking into account relevant contextual factors including in particular texts written by different EFL learner groups in different socio-cultural contexts. Additionally, few studies on L2 writing have systematically examined the writing produced by these students within the context of their writing classes, though with a few exceptions [31]. Furthermore, the recent conceptualization of syntactic complexity as a multidimensional construct [27], calls for the use of multiple measures that tap into multiple dimensions of complexity, yet as the forgoing review highlighted that there are no studies that explore the syntactic complexity employed a large set of multiple measures in a single study [30]. Specifically, in the Ethiopian context to the best of our knowledge, no single study has systematically investigated the link between writing quality and syntactic complexity features using a large set of measurements including global, fine-grained clausal and phrasal measures and syntactic sophistication thus far.

Despite the paucity of research comparing the relationship of different types of syntactic indices to the L2 writing quality of university-level students in the Ethiopian context, there are at least three reasons for studying it regarding the classroom context. First, Ethiopian students seem to suffer from several difficulties in L2 writing in that expressing complex ideas in words and sentences to form coherent writing is a challenging task (Dawit, 2013; Eskinder, 2018). In particular, at the university level, the researchers, and instructors at university, have observed the texts by university students are generally shorter, not discipline-specific, and less dependent on the use of reliable evidence or sources to defend one's arguments; these are evidenced by their tests, examinations, class works, assignments and senior essay papers. Second, it is clear that a writing sample (i.e., an argumentative essay) produced in a classroom context may produce different syntactical features than writing samples produced in high-stakes assessment settings. Third, the assessment of students' writing samples is generally perceived to be very demanding and time-consuming for writing teachers. It is, therefore, necessary to investigate the relationship between argumentative writing quality and certain syntactic features to inform teachers' manual rating, as well as to lay a foundation for the development of automated programs to supplement manual assessment.

#### *The present study*

Motivated by the afforded mentioned research insight and discussed gaps in the literature the present study aims to investigate the extent to which a number of syntactic complexity and VAC-based syntactic sophistication indices could predict EFL learners' argumentative writing quality. To this end, we focus on two specific questions:

What is the relationship between syntactic complexity (with its different dimensions) and the quality of EFL students' writing?

What is the predictive power of syntactic complexity (with its different dimensions) on EFL students' writing quality?

## Methodology

### *Sample and setting*

Thirty (13 female and 17 male) first-year undergraduate students, studying at Mekedela Amba University (MAU) in Ethiopia, served as the participants of this study. The participants were recruited through stratified random sampling their ages ranged from 19 to 25 years old ( $M = 21.1$ ;  $SD = 1.4$ ) and they had different first language (L1) backgrounds, including Amharic, Awi, and Oromia meanwhile, they had received formal English language education for 12 years in school. The participants were in a context where English is taught as a foreign language and thus have almost no access to producing something outside the classroom. At the time of this study, the participants were studying an Academic English course namely Communicative English II. This course aids first-year students in enhancing their overall communicative competence in particular the writing section striving to develop the higher-level academic writing abilities required to flourish in their disciplinary courses so they have background knowledge and prior experience in writing an argumentative essay. Additionally, the participants reported that they were familiar with writing an argumentative genre and found it easy to do so. The participants' mean rating for their level of familiarity with writing in an argumentative genre was 5.2 on a 9-point Likert scale (1 = least familiar, 9 = extremely familiar). On another 9-point Likert scale (1 = least difficult, 9 = most difficult), the mean rating for perceived difficulty in writing in an argumentative genre was 5.2.

### *Data Collection*

Data for this study were collected through an argumentative writing task adapted from A. Fazilatfar et al. (2020) study for several reasons. Firstly, argumentative writing is a common mode of academic discourse in many L2 writing courses. Secondly, previous studies have reported that argumentative writing tasks elicit more complex syntactic structures than other writing tasks such as narrative. The task prompt required the participants to justify their opinion with the following statement:

*Some people believe that the Internet causes many problems although it has a lot of advantages. To what extent do you agree or disagree? In your writing, use your own ideas and experience, and support them with examples and relevant evidence*

This task was selected to minimize the variability that could have resulted from the topical knowledge; the researchers assumed that the topic of the 'Internet' enabled them to tap into the participants' perspectives regarding a broadly discussed social issue, rather than other personal issues, was probably more related to university level students. The participants were given 30 minutes for the writing activity. In this way, both time and topic constraints were controlled to make results comparable [35].

### *Procedure*

After receiving approval from the Mekedla Amba University community services and research directorate office, we contacted the instructors who had agreed to participate in the study. We explained the study's objectives and procedures to them, and they agreed to conduct the writing experiments as in-class activities and award bonus credits to students who submitted the essays through the course site. They all participated in the research voluntarily, signing an informed consent form. Participants completed the writing task with paper and pencil in a regular writing classroom, and their English teachers administered the procedures. First, the students completed a self-assessment questionnaire adopted from (M. Abdi Tabari et al., 2023). Then, the teacher distributed the writing prompt and the essay sheet in total, 30 argumentative essays were collected. Two-course instructors, who had master's



degrees in English, graded the essays. Intra-class correlation coefficients (ICCs=.870, 95% CI showed 0.671–0.957,  $p=.000$ ; which was high) were calculated to check for the inter-rater reliability, we used a two-way mixed effects model with absolute agreement definition, which assumes that the essay was rated by two or more raters and that these raters were the only raters participating in the study. For computational analyses of syntactic complexity measures, all the essays were transferred from handwriting to typed texts and saved in a .txt file format which is the format required by the TAASSC program.

#### *Computational tools, TAASSC*

The current study employed computational tools namely TAASSC to measure the syntactic features of the writing sample. This is mainly; because TAASSC is a freely available and fully automated text analysis tool (cf Biber Tagger) [3]. Furthermore, previous studies have reported that the fine-grained clausal and phrasal as well as syntactic sophistication indices provided in TAASSC are more detailed than the indices proposed by other tools [18; 30]. TAASSC calculates approximately 370 indices of syntactic variation. Each index is calculated about four sub-corpora written registers in the Corpus of Contemporary American English (COCA) namely, academic, fiction, magazine, and newspaper (Davies, 2015). Since our samples are academic writings, we only used the indices generated from the academic sub-corpus of COCA. After the target, syntactic feature indices were obtained for each writing sample, or the results of TAASSC calculation were analyzed in SPSS (Ver. 21) for significance.

A pilot study was conducted to test the reliability of the POS tagging and TAASSC parsing in learner data. We chose two random essays 5 and 28 and the automated analysis was manually checked; we found high accuracy, with POS tagging at 98% and TAASSC parsing at 93% with minor difficulties encountered in unknown or misspelled words. The precision of the tagging and parsing may be due to the written, punctuated style of the text.

#### *Syntactic complexity indices*

##### *Absolute syntactic complexity indices*

To assess the internal complexity of the text such as syntactic embedding in a largely holistic way, we used the 14 indices reported by Lu's (2011) L2SCA analyzer embedded in TAASSC. These indices calculate syntactic complexity into five dimensions with different measures in each dimension. Table 1 summarizes the complexity measures included in L2SCA. It should be noted that in the L2SCA a clause is defined only as a structure with a subject plus a finite verb. Thus for the subsequent more detailed

Table 1. Absolute syntactic indices

Dimension of measures	Types of measures	Definitions
Length of production unit	The mean length of sentence	# of words per sentence
	The mean length of T-unit	# of words per T-unit
	Mean length of clause	# of words per clause
Sentence complexity	Clauses per sentence	# of clauses per sentence
Amount of subordination	Clauses per T-unit	# of clauses per T-unit
	Complex T-unit ratio	# of complex T-units/by T-units
	Dependent clauses per clause	# of dependent clauses/ clause
	Dependent clauses per T-unit	# of dependent clauses /T-unit

Amount of coordination	Coordinate phrases per clause	# of coordinate phrases/clause
	Coordinate phrases per T-unit	# of coordinate phrases/T-unit
	T-units per sentence	# of T-units per sentence
Degree of phrasal sophistication	Complex nominals per T-unit	# of complex nominals/T-unit
	Complex nominals per clause	# of complex nominals/clause
	Verb phrases per T-unit	# of verb phrases per sentence

analysis, we included a broad range of fine-grained clausal and phrasal measures made use of TAASSC [17].

*Fine-grained clausal complexity indices*

Table 2. Clausal dependent types analyzed by TAASSC [17]

Structures	Abbreviation	Examples of structure
adjective complement	acomp	<i>She [looks]gov [beautiful] acomp</i>
adverbial clause	advcl	<i>The accident [happened]gov [as night fell]advcl</i>
adverbial modifier	advmod	<i>[Accordingly]advmod, I [ate]gov pizza</i>
auxiliary verb	aux	<i>He[is]aux [running]gov</i>
bare noun temporal modifier	tmod	<i>Last [night]tmod, I [swam]gov in the pool</i>
clausal complement	ccomp	<i>I am [certain]gov [that he did it]ccomp</i>
clausal coordination	cc	<i>[Jill runs]gov and [Jack jumps]cc</i>
negation	neg	<i>He did [not]neg [kill]gov them.</i>
clausal prepositional complement	pcomp	<i>They heard [about]gov [you missing classes]pcomp</i>
clausal subject	csubj	<i>[What she said]csubj [is]gov not true</i>
conjunction	conj	<i>He [runs]gov and [jumps]conj</i>
controlling subject	xsubj	<i>[Tom]xsubj likes to [eat]gov fish</i>
direct object	dobj	<i>She [gave]gov me a [raise]dobj</i>
discourse marker	discourse	<i>[Well]discourse, I [like]gov pizza</i>
existential "there"	expl	<i>[There]expl [is]gov a ghost in the room.</i>
indirect object	iobj	<i>She [gave]gov [me]iobj a raise</i>
parataxis	parataxis	<i>The guy, John [said]parataxis, [left]gov early..</i>
modal auxiliary	modal	<i>He [may]modal [be]gov awesome.</i>
nominal complement	ncomp	<i>He [is]gov a [teacher]ncomp</i>
nominal subject	nsubj	<i>The [baby]nsubj [is]gov cute</i>
open clausal complement	xcomp	<i>I am [ready]gov [to leave]xcomp</i>
agent	agent	<i>The man has been [killed]gov by the [police]agent</i>
passive auxiliary verb	auxpass	<i>Kennedy has [been]auxpass [killed]gov</i>
passive clausal subject	csubjpass	<i>[That she lied]csubjpass was [suspected]gov by</i>

passive nominal subject	nsubjpass	<i>[Dole]nsubjpass was defeatedgov by Clinton</i>
phrasal verb particle	prt	<i>They [shut]gov [down]prt the station</i>
prepositional modifier	prep_	<i>They [went]gov [into the store]prep_into</i>
subordinating conjunction	mark	<i>Forces engaged in fighting [after]mark insurgents [attacked]gov</i>
undefined dependent	dep	<i>N/A</i>

Note. “gov” represents the governor of the dependent; \*prepositional modifier representations include the actual preposition

We specifically selected 31 fine-grained clausal complexity indices tap into the amount of usage of specific types of dependent clauses structure from TAASSC, which are seen in Table 2. As seen in the table below twenty-nine indices calculate the average number of particular structures per clause. As well, TAASSC calculates two additional general indices of clausal complexity that take into account the total number of dependents per clause and provide a measure of syntactic variation.

Table 3. Types of phrasal indices used in TAASSC [17]

Index type	Calculated for	Measures
Average number of dependents / each phrase type	All phrase types	16
Occurrence of particular dependent types	All NPs	10
Average occurrence of particular dependent types	Some NPs	40
Total		66

Note that the overall phrasal complexity measures are 132 measures divided into two versions: pronouns and no-pronoun version yet for this study we used the no-pronoun version, as pronouns are not commonly mentioned in academic writing.

*Fine grained phrasal complexity indices*

Given the importance of phrase-level syntactic complexity in academic writing, we used 66 phrase complexity indices from TAASSC. The rationale for this is unlike to L2SCA; TAASSC phrasal complexity measures focuses on noun modifiers (dependent types) and the place that noun modifiers occur (phrase types). The terminology and abbreviation of dependents and phrase types are presented in Appendix A for clarity purposes. In addition, TAASSC considers standard deviation in order to measure syntactic variation.

*Syntactic sophistication indices*

To measure this dimensions of complexity, we also selected indices that calculated verb-argument construction complexity, such as the frequency of VAC, TTR of verb argument constructions, and the strength of association.

Table 4. Summary of usage based VACs indices

Indices types	Verb lemma frequency	VACs frequency	Verb-VAC combination frequency
Mean token score	√	√	√
Mean token score (log-transformed)	√	√	√
Standard deviation (SD) token score	√	√	√
SD token score (log-transformed)	√	√	√
Mean type score	√	√	√
The proportion of items attested in the corpus	√	√	√
<b>Total</b>	6	6	6

#### *Frequency-based indices*

Foremost, we calculated VAC frequency indices using TAASSC which include average frequency scores for main verb lemmas (e.g. to make), VACs (e.g. subject-verb-direct object), and verb-VAC combinations (e.g. subject-to make- direct object) in a target text (in this case the writing samples) based on counts derived from COCA. These measures were calculated based on both the tokens (all instances of a particular structure) and types (unique instances of a particular structure) of the structures extracted from the text and standard deviations for these scores. To account for the potential Zipfian nature, distributional skewness, of the data, we also included the logarithm-transformed versions of these frequency measures. Besides, we also counted type-token ratios (TTR) for main verb lemmas, VACs, and verb-VAC combinations to provide measures of the diversity of the VAC structures in the text. These indices comprise a rough measure of frequency. In the current analysis, in total 18 frequency indices based on the frequency norms derived from the academic section of COCA are used (as seen in Table-4).

#### *Association strength indices*

We also measured the association strength of the VAC structures, i.e. the conditional probability that the main verb lemma and a VAC will occur together. TAASSC utilizes three types of association strength measures to calculate mean association strength scores for both the types and tokens of each VAC structure in a target text: faith, delta P, and colostruational analysis. TAASSC also calculates intra-text standard deviation scores for tokens. Faith calculates the probability of whether the main verb lemma (the outcome) will occur given a particular VAC structure (the cue), and vice versa. Delta P is similar to faith but is a more conservative method. It measures the probability of an outcome given a cue minus the probability of the outcome without the cue. In addition to, Faith *and* Delta P that measure association strength directionally, we used the average, the ratio and standard deviations for colostruational strength analysis computes the joint probability that two items (the main verb lemma and the VAC structure) will co-occur without the concern for directionality (for a total of 17 indices).

#### *Statistical Analysis*

We pruned the 146 syntactic complexity indices before statistical analysis. First, the normal distribution of these indices was verified by assessing their skewness and kurtosis levels, syntactic features that were normally distributed were included in further consideration. After that, Pearson correlations were

performed on the remaining syntactic features to see if there was any real association between them and the holistic essay scores. Any syntactic features that did not achieve an absolute correlation value of  $r \geq 0.100$  with the holistic essay score (which Cohen, 1988 defines as the threshold for a "small" influence) and a significance level of  $p < 0.001$  were eliminated from further consideration. For those variables, we controlled for multicollinearity to ensure we did not include indices that measured similar language features. We did so using VIF (variance inflation factors) test score such that any variables showing a VIF below ten were retained (Levshina, 2015).

Based on the correlational results, (for RQ2) we built five multiple regression models in a stepwise fashion independently for each index type (traditional, fine-grained clausal, fine-grained phrasal, VACs indices and combined all these measures) to investigate the predictive power of the various syntactic features measures (as predictor variables) on human ratings of essay quality (as an outcome variable). Since multiple regression shows combined and independent contributions of predictors, this analytical strategy was used to determine the extent to which the selected syntactic features predict the human-rated holistic writing score. To determine goodness of fit of the multiple regression model we report, Adjusted R-squared ( $Adj.R^2$ ) this is because unlike the R-squared ( $R^2$ ) value assumes all the independent variables considered affect the result of the model, adjusted R-squared only adds new predictors to its model if it improves the model's predicting power.

## Result

The current study systematically investigated the syntactic features in undergraduate EFL students' argumentative essays to identify the unique syntactic features of their academic writing and to analyze the extent to which syntactic features predict writing quality. In the lines that follow the results obtained from correlations analysis and regression analyses for each of the complexity measures included in the current study are displayed.

### *Result of absolute syntactic complexity indices*

First, the potential for the 14 indices in Lu's (2011) SCA to explain the variance in holistic scores of essay quality was investigated. To achieve this, a number of preliminary analyses were performed to confirm that the data was appropriate for correlation and stepwise multiple regression analysis. Per se, eight of the 14 indices violated the assumptions of normality and were excluded from further investigation. Even though the remaining six indices demonstrated normal distributions, five of these indices did not reach the minimum correlation thresholds of  $r \geq 0.100$  and  $p < .001$  and were removed from further consideration. One variable namely the number of complex nominals per T-unit (CN\_T), which measured phrase level elaboration, demonstrates a strong and nearly significant correlation with

Table 5. Correlation between holistic essay score and clausal complexity indices

variables	r	Sig.
nominal complement per clause	.492**	0.02
modal auxiliaries per clause	.459*	0.03
nonfinite clausal complement per clause	-.489**	0.01
adverbial modifiers per clause	-.588**	0.01

Note: \*\* $p < 0.01$ , \* $p < 0.05$



Table 6. Summary of clausal complexity multiple regression model

Entry	variable	r	R <sup>2</sup>	Adj.R <sup>2</sup>	$\beta$	SE	B
1	adverbial modifiers	.588	0.346	0.323	-4.985	1.318	-0.529
2	modal auxiliaries	.696	0.484	0.446	5.104	1.899	0.376

Note. Estimated constant term=3.994;  $\beta$ =unstandardized beta; SE=standard error; B= standardized beta

holistic writing score ( $r=.342$ ,  $P=.065$ ), but was not included in the predictor model. Overall, the result revealed that there was no significant relationship between the traditional syntactic complexity indices and writing quality, implying that these indices had little explanatory power on L2 writing quality [19; 30; 38]

#### *Result of fine-grained clausal complexity*

As for clausal level complexity measures, twenty of the 31 clausal complexity indices violated the assumption of normality (e.g., passive constructions), due to rare occurrences in participants' essays and were excluded from further examination. Seven of the remaining 11 variables did not meet the minimal correlation requirements of  $r \geq 0.100$  and  $p < .001$  and thus were eliminated from further analysis (e.g., direct object). As seen in Table-5 only four variables, namely the frequent usage of adverbial modifiers, modal auxiliaries, nominal complement, and nonfinite clausal complement per clause, which demonstrated meaningful relationship and had a large effect size with writing score, were entered into a stepwise regression analysis to see if the retained indices may explain the variation in L2 argumentative writing scores.

The resulting model, as seen in Table-6, comprised two predictors, namely the use of adverb modifier per clause and modal auxiliary per clause was significant ( $r= .696$ ,  $adj. R^2=.446$ ,  $F_{2,15.83} = 12.64$ ,  $p= .000$ ) and explained 44.6% of the variation in writing scores. To look at the unique contribution of each of the predictors the coefficient model was assessed, the coefficient of adverb modifier per clause was ( $\beta= -.529$ ,  $t= -3.782$ ,  $P < .001$ ) and the coefficient of modal auxiliary per clause was ( $\beta=.376$ ,  $t=2.687$ ,  $P<.005$ ).

Table 7. Correlations between essay score and phrasal complexity variables

Variables	r	Sig.
determiner per nominals	.399*	0.02
determiner per direct object	.383*	0.03
possessive per direct object	-.519**	0.01

Note: \*\* $p<0.01$ , \* $p<0.05$

Table 8. Summary of phrasal complexity multiple regression model

Entry	variable	r	R <sup>2</sup>	Adj.R <sup>2</sup>	$\beta$	SE	B
1	possessive per direct object	0.52	0.27	0.24	-5.49	1.71	-0.52

Note. Estimated constant term=4.285;  $\beta$ =unstandardized beta; SE=standard error; B standardized beta;

Simply put, the model showed a statistically strong negative relationship between adverb modifiers per phrase and a moderately significant positive relationship between the use of modal auxiliary per clause and writing scores implying that higher-quality essays typically contain fewer adverbial modifiers and more modal verbs per clause, the result is in line with [4; 30].

#### *Result of fine-grained phrasal complexity*

As previously mentioned, the current study takes into account 66 phrasal complexity indices of which 16 showed normal distributions. After pruning these 16 variables, three indices (see table-7) specifically, determiner per nominals, determiner per direct object, and possessive per direct object demonstrated a meaningful relationship with writing score and were entered in the stepwise regression model.

The resulting model (Table-8) consisted of one index namely the use of possessives per direct object was significant and explained 24.3% of the variance in holistic essay scores ( $r = .519$ ,  $adj. R^2 = .243$ ,  $F_{22,40} = 10.334$ ,  $P < .005$ ). The coefficient of dependents possessives per direct object was ( $\beta = -.519$ ,  $t = -3.215$ ,  $P = .003$ ). The model demonstrated a strong significant negative correlation between the number of possessive phrases per direct object and writing scores, implying that higher-scoring essays contain fewer dependents, such as possessive and more determiners, as modifiers indirect objects.

Overall, the model showed that indicators of phrasal elaboration were indicators of the quality of the essay.

Table 9. Correlations between essay score and syntactic sophistication variables

Variables	r	sig
average lemma frequency	.416*	0.02
average lemma frequency types	.053**	0.01
main verb lemma type-token ratio	-.472**	0.01
average lemma frequency log transformed	.396*	0.03

Note: \*\* $p < 0.01$ , \* $p < 0.05$

Table 10. Summary of VACs multiple regression model

Entry	variable	r	R <sup>2</sup>	Adj.R <sup>2</sup>	$\beta$	SE	B
1	av lemma frequency type	.503	.253	.227	6.4E-006	0.00	0.503

Note. Estimated constant term=2.488;  $\beta$ =unstandardized beta; SE=standard error; B= standardized beta

Table 11. Summary of combined analysis multiple regression model

Entry	variable	r	R <sup>2</sup>	Adj.R <sup>2</sup>	$\beta$	SE	B
1	adverbial modifiers	.588	.346	.323	-4.002	1.276	-0.425
2	modal auxiliaries	.696	.484	.446	5.206	1.746	0.384
3	lemma type-token ratio	.762	.581	.532	-2.251	0.92	-0.328

Note. Estimated constant term=5.336;  $\beta$ =unstandardized beta; SE=standard error; B= standardized beta

#### *VAC-based syntactic sophistication*

Penultimately, to examine the relationship between syntactic sophistication and writing quality 35 VAC indices (18 frequency related and 17 related to association strength) were employed, of these ten indices violated the premise of normalcy and were excluded from further consideration. Twenty-one indices out of the 25 remaining variables did not meet the minimum correlation standards of  $r \geq 0.100$  and  $p < .001$  and were therefore eliminated from additional analysis. The remaining four variables (presented in Table 9) were included in a stepwise regression analysis.

The final model, as seen in Table-10, comprised a single syntactic diversity measure namely verb lemma frequency type was significant ( $r=.503$ , adj.  $R^2=.227$ ,  $F_{22.9} = 9.51$ ,  $p=.005$ ) and explained 22.7% of the variation in writing score. The coefficient of this index was ( $\beta= 6.5E-006$ ,  $t= 3.08$ ,  $p=.005$ ) revealing that essays with a wider range of frequent VAC structures received a better writing score.

#### *Combined Analysis*

Finally, the ability of the 11 syntactic complexity and sophistication variables included in each preceding regression model to explain variance in essay holistic scores was explored. As all of these variables satisfied normality and minimum correlation with writing score, and none of the indices exhibited collinearity, a stepwise regression was performed on all of them. The resulting model, see Table 11, included two fine-grained clausal complexity and one syntactic sophistication index was significant ( $r=.762$ , Adj. $R^2=.532$ ,  $F_3, 12.87 = 11.99$ ,  $p=.000$ ) and explained 53.2% of the variation in writing scores, suggesting that the best predictive results were achieved by the combination of syntactic complexity and sophistication features. The results indicated that clausal elaboration and lemma type-token ratio contributed to a model of development in a complementary manner.

### **Discussions**

Examining syntactic complexity as a multi-dimensional construct with different levels of sub-constructs, the study revealed complex yet patterned findings about the relationship between syntactic complexity and writing quality, categorized according to the level of syntactic complexity dimensions. The discussion centers on two main areas that our study can illuminate: First, we present explanations for the observed patterns in our study related to their use within argumentation styles followed by measurement issues pertaining to syntactic complexity dimensions.

Regarding syntactic clausal complexity, the fine-grained analysis revealed that the production of two specific clausal structures, namely the use of adverbial modifiers per clause (which accounted for

Table 12. examples of finite and nonfinite adverbial dependent in high-scored essay

Constructions (VAC)	Examples
nsubj-v-advmod-[ccomp]	I <b>agree</b> <i>completely</i> [that the internet <b>causes</b> many problems <b>although it has a lot of advantages.</b> ]< t>
advmod-nsubj-[ rmod]-advmod-modal-v-dobj-[advcl]	<i>Shortly</i> , people [who use internet] <b>frequently may face</b> eyesight problems [as they <b>have</b> seen mobile or laptop screen light for long time].
advmod-nsubj-v-dobj-xcomp-xcomp	<b>Additionally</b> , internet <b>enables</b> us <b>to accomplish our daily activities</b> and <b>to solve problems quickly.</b>

32.3%) and modal auxiliaries per clause (which explained 12.3% of the variation in writing scores), were the best indicators of EFL writing quality and syntactic competence. This finding is consistent with previous studies [30] that found a significant predictive power of these specific clausal features, or others [18; 38] report better predictive power of fine-grained measures of clausal complexity over holistic one and reported development of clausal complexity to plateau at advanced levels [28]. Additionally, this result aligns with Crossley and McNamara (2014) who reported that human rating scores of writing quality are largely predicted by clausal complexity features.

More specifically, the predictive power of adverbial modifier per clause i.e., its inverse relationship with writing score, suggests that excessive use of adverbial modifiers may negatively affect the quality of writing. Yet, technically this inverse relationship could be explained by the inclusion of a range of finite adverbial clauses (e.g., although) or nonfinite complement clauses (e.g., infinitive clause) within another dependent clause in the sentence, as TAASSC counts both finite and non-finite verb phrases as clauses. To demonstrate some of the results we include some example sentences, in Table 12, taken from a high-scored essay extract. To aid identification, these clausal features are demarcated with textual features such as [square brackets], **bolded**, highlighted with broken, single, and doubled lines.

As illustrated in the first extract, the writer employed an adverbial modifier (*completely*) to modify a finite “that” complement clause controlled by the less frequent verb “cause”, surrounded by square brackets, consisting of finite adverbial clause controlled by most frequent verb “have”, allows the writer to acknowledge the oppositions’ arguments while still supporting her argument. This modification adds more information within a single clause, demonstrating writer's advanced mastery of syntactic variety and control. Moreover, as seen in other extracts the use of different adverbial modifier (*in italics*) allows writers to express their stance implicitly entrenching ranges of clausal features such as finite relative clauses, adverbial causatives/conditions, infinitive clauses, and complex nominals. These features highlight an optimal use of adverbial modifiers in clausal elaboration reduces clausal lengths and diversifies sentence structure, resulting in a more sophisticated and cohesive piece of writing. Likewise, frequent use of modal auxiliaries such as “may” in the second extract, can help writer convey a sense of possibility, adding credibility to his claims. Additionally, a closer look at the use of these clausal features in student’s essay at two different score points displayed a remarkable level of syntactic variation in students verbs choices in producing argument constructions to achieve complexity. For example, high-scored essays (see table-12) ostensibly produced by proficient student writers exhibited clausal complements controlled by less frequent verbs (e.g., cause, enables) these features are considered syntactically more complex than clauses controlled by frequent verbs (be). These characteristics show the writer's versatility in sentence construction and the growth of non-finite

subordinate clause complexity in L2 writing [3]. The essay with a lower score had simpler sentence structures and a wide range of dependent clauses, controlled by frequent verbs, resulting in a less engaging and dynamic writing style as in “Although the uses of the internet is greater than that of its disadvantage, people use the internet in a different way and it is useful nowadays. <002>”

Overall, our study found that clausal structures are key techniques of conveying stance in academic writing and have a strong significant predictive value in the writing quality of argumentative essays, which contradicts prior studies [34]. This inconsistency could be explained by differences in writing conditions or student characteristics. For instance, unlike most previous studies, the participants in our study completed the writing task in their classroom context, where students are more likely to plan, reflect, and revise writing through multiple draughts, potentially leading to their ability to make a variety of choices. This could be attributable to different standards of what makes "good" writing in general English classrooms versus high-stakes assessment settings. Put another way, the difficulty of the assignment led them to use more formal, academic verbs, which were not varied. Additionally, participants in the current study might have come from diverse linguistic backgrounds and had different writing experiences, which could have affected their use of subordination, as Lu & Ai, (2015) and Staples & Reppen, (2016) noted.

Alternatively, the difference could also be explained by task-related variables like writing tasks [8], (examined research article quality), scoring rubric [34] and writing prompts (Ryu’s, 2020, academic topic “using animal in experiment”). Thus, more investigation is required to examine these plausible causes and offer a more thorough comprehension of the connection between syntactic clausal patterns and L2 writing.

Regarding syntactic phrasal complexity, the result from absolute complexity indices analysis showed that one variable namely complex nominals per T-unit demonstrated a strong and nearly significant correlation with holistic writing score ( $r=.342$ ,  $R^2 = 0.117$ ). The correlation analysis revealed that the frequency distribution of complex nominals per T-unit increased linearly with writing scores and accounted for 11.7% of the variation in the score, implying that higher-scoring essays contained elaborated complex nominal with more modifiers, which lengthened the clauses and made the essay highly propositional (see the underlined structures in Table-13). Our findings support previous studies [1; 21] found positive correlations between CN\_T and writing quality in various contexts of academic writings and others [19; 37] that revealed proficient writers represent their complex ideas in argumentation through higher phrasal density.

The findings, however, left gaps in our understanding of the specific structures that lead to an increase in the number of complex nominals per T-unit and syntactic function of these structures, as has been widely criticized [7; 2]. Complementing this, the results of the fine-grained phrasal complexity analysis

Table 13. Examples phrase level elaboration from low and high-scored essay

Score	Examples	CN_T
2	I agree that internet has <u>many advantages</u> since it helps to solve <u>many educational problems</u> although it has <u>some disadvantages</u> .	3
5	<u>Shortly internet is the source of knowledge for all people all over the world such that it has a great advantage of development</u> in general and for <u>societal development in particular</u> .	8



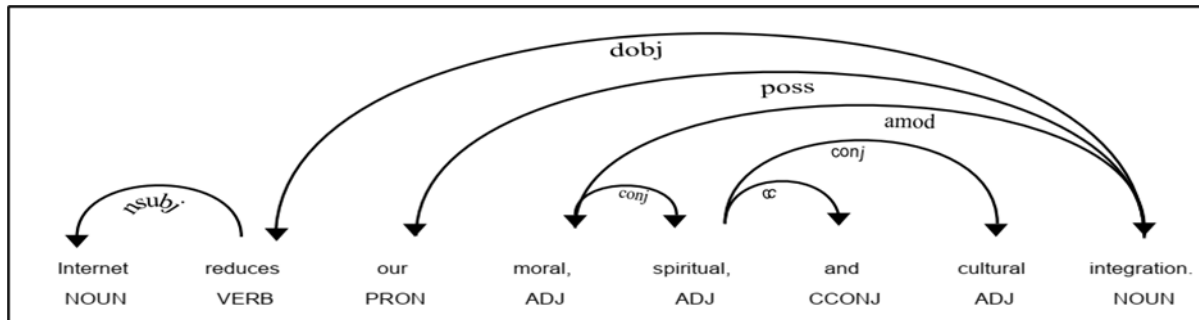


Figure 1. Phrasal complexity: Possessives per direct object

(see Table-7) revealed that the increase in the length of noun phrases points to an increased use of determiners and possessive modifiers in both nominal subjects and direct objects. Specifically, frequent use of possessives in direct objects with other modifier like attributive adjectives is likely to be a trait of good quality. For example, a highly graded essay is likely to include at least one possessive word along with three attributive adjectives to modify the direct object as seen in the example below, see Figure1 for visualization. This finding aligns with previous studies [4; 19] reported that phrasal complexity increases when academic level increases. Example: Internet affects our moral, spiritual, and cultural values”,

Taken together the finding vis-à-vis the syntactic complexity dimensions indicates that the incorporation of different syntactic patterns adds depth and complexity to their writing, making it more engaging for the reader. Particularly, as evidenced in our study clausal elaboration allows academic writers to express complex relationships and abstract ideas more concisely in academic writing. Thus, our study provide evidences that clausal structures are also features of academic writing. Beside, when viewed through Biber et al.'s (2011) and Norris and Ortega's (2009) proposed trajectories for syntactic development in L2 writing (which was never the goal of this study), the trend observed in clausal complexity and phrasal level complexity supports the assertion that writers will switch from writing with finite dependent clauses to writing with nonfinite dependent clauses.

In reference to syntactic sophistication, the results indicate that the relationship between usage-based VACs frequencies indices and holistic scores of writing quality were significant and demonstrated large effect (see table-9). One index, average lemma frequency (type), included in the regression model, explained 22.7% of the variance in essay score, indicating that essays with range of main verb lemmas (e.g., have, use) and with frequent main verb lemma - VAC combinations (e.g., S + V<sub>have/use</sub> +O) tended to earn higher scores. This finding is consistent with previous findings [12; 18] that frequency is an

Table 14. Examples of verb argument construction with verb “have”

VACs	have_subj-v-dobj- adverbial clause
Example	E learning has many advantages ... because students get courses material ...
VACs	have <u>mark</u> -subj-v-dobj- adverbial clause,
Example	I agree <u>that</u> internet has many advantages since it helps to solve ...
VACs	have <u>advmod</u> -subject--v-direct object-adverbial clause
Example	<u>Therefore</u> , it has various advantage in our life when we use the internet ...

important factor in language development, and others [33] reported that L2 learners expand their repertoires of VACs in writing through repeated language experiences with similar combinations. Notably, the positive coefficient of average lemma frequency (type) indices indicated that using most frequent main verb lemma (e.g., be, have) to construct range of VACs is also a feature of higher-quality writings, which runs contrary to previous studies [25; 30] found employing frequent verbs to build an argument construction is relatively downgraded in the writing.

A plausible explanation for this difference might be EFL learners at the university level, regardless of their L2 proficiency (writing score) had similar verb preferences at least for a relatively small set of VACs. For instance, for one of the most frequent verbs “have” with about 61 hits, we recognize more than 10 types of VACs in the students' corpus while most of the sentences (18.3%) belong to the type (subject-verb -direct object: “Internet has many advantages”) and (29.5%) belongs to “subordinator-subject-v-direct object: although it has many advantages... ”; while some of them (4.9%) belongs to complex VACs types (see table-14) embedding such clausal features as adverbial clausal modifiers, clausal complement and range of clausal coordination.

This use of the verb “have” with adverbial modifiers and adverbial clauses is considered as an important addition to proficient students' writing framework suggests the writer's versatility in constructing sentences. Therefore, high-quality essays are expected to demonstrate a flexible use of functional characteristic verb-VAC combinations. Moreover, higher-scored essay samples exhibited similar VAC types controlled by less frequent verbs (e.g., cause/use instead of have) indicating proficient EFL learners were less dependent on certain path-breaking verb to formulate abstract knowledge of less frequent and complex constructions. This claim aligns with the hypothesis of Usage-Based theory that through further language experiences, EFL learners gradually master the way of using more sophisticated and less frequent verbs to build VACs [14; 26]. Thus, we need to be aware that use of high-frequency VAC patterns, which are assumed to be less complex within the usage-based framework is not necessarily, associated with low argumentative writing quality in EFL contexts. Reinforcing this point, the negative coefficient of the lemma type-token ratio index measures the diversity of verb lemma and main verb lemma – VAC combinations, included in the combined complexity model,(see table-11), revealing that the essays that include low-frequency verb-VAC combinations tended to earn higher scores. This suggests from a usage-based perspective that higher proficiency learners have had sufficient language exposure to have learned which verbs are normally used in particular constructions. but it seems that constructing the right constructions, and grammatical and syntactic knowledge of some complex combinations takes more time to settle in for EFL learners. To recap, the findings generally support usage-based perspectives on language learning (e.g., Behrens, 2009; 11] in that indices related to VAC frequency and diversity were indicators of writing development and quality.

### *Implication*

The study provides predictors of writing quality at different dimensions of syntactic complexity, allowing EFL instructors to design classroom activities to promote writing quality. Put simply, as our study indicates that an increased repertoire of certain complex structures along with the flexible use of frequent verb-VAC patterns is likely to improve writing quality, we believe that incorporating and explicit teaching some of the clausal and phrasal complexity structures and VAC-based elements into writing programs will be extremely beneficial. We contend that this kind of instruction could make use of genre-based pedagogy and explicit instruction, which helps students create complex structures that are in line with academic writing conventions and content is particularly promising. For instance, Casal

and Lu (2021) drawn on, Concept-Based Instruction (CBI) approach, propose an instructional activities such as complexity-focused activities targeting essential complexity features (e.g., clausal complement or prototypical VACs) via reflection-oriented group discussions or individual sample text-analysis activities targeting the frequency and functional affordances enable advanced students internalize such features.

Furthermore, Ethiopian teachers should compile a corpus of argumentative essay, identify key verb-VAC expressions, and analyze their distribution across different sections. This can be used for independent analysis and reflection-oriented group discussions on the functions of key verb-VAC patterns. As previously stated, EFL teachers face challenges in manual evaluation of academic genres in classrooms, as they have few criteria or rubrics to rely on. The findings of the study provides a pool of linguistic features predictive of argumentative quality, allowing for the development of a writing rubric especially the coefficients can be used to set weights for different descriptive factors, enhancing construct coverage and assessment accuracy as well the efficiency of EFL teachers [15].

Finally, this study provide valuable insights using an automated assessment system to help EFL teachers provide detailed writing feedback to students at different dimensions of syntactic complexity because we believe that a finer-grained analysis can be rendered possible by developing an automated assessment system that can predict the writing quality of different genres based on carefully pre-selected features.

### Conclusions

The present study used different cohorts of syntactic complexity features to predict the writing quality of argumentative essays. The findings of the study illuminate the best predictive results were achieved by the combination of the fine-grained clausal complexity and VAC-based indices. In a comparison of the predicting power of the two sets of measures, we found that fine-grained clausal complexity followed by VAC-based relative measures are sometimes better at indexing argumentative writing quality in comparison to the absolute measures. This indicates the importance of including argument-based features for assessing the writings of academic genres.

Even though the study was rigorously designed, there are limitations. Firstly, we focused on a specific genre (i.e. argumentation) commonly used in L2 writing classrooms in the EFL context. Future research needs to rigorously test whether syntactic complexity is robust predictor of L2 writing quality across different genres (expository or description) composed by the same learner group using paired writing samples or under different writing conditions (e.g., with or without no time limit) on a larger scale. Secondly, due to the scope of the study, no attempt was made to examine the participants' L1 background. Given that L2 writing complexity seems to be affected by L1 background [23], subsequent research is encouraged to control for participants' L1 background to uncover the effects of linguistic background on VAC-based syntactic sophistication features.

### References

1. 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 Publishing Company.

2. Bi, P., & Jiang, J. (2020). Syntactic complexity in assessing young adolescent EFL learners' writings: Syntactic elaboration and diversity. *System*, 91, Article 102248.
3. 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.
4. Biber, D., Gray, B., & Staples, S. (2014). Predicting Patterns of Grammatical Complexity Across Language Exam Task Types and Proficiency Levels. *Applied Linguistics* 37: 639–668.
5. 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.
6. Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency* (pp. 21–46).
7. Bulte, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing*, 26, 42–65.
8. Casal, J., & Lee, J. (2019). Syntactic complexity and writing quality in assessed first-year L2 writing. *Journal of Second Language Writing*, 44, 51–62.
9. Crossley, S. (2020). Linguistic features in writing quality and development: An overview. *Journal of Writing Research*, 11(3), 415–443.
10. Crossley, S. A., & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing*, 26, 66–79.
11. Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188.
12. Ellis, N. C., and F. Ferreira-Junior. 2009. "Construction Learning as a Function of Frequency, Frequency Distribution, and Function." *The Modern Language Journal* 93 (3): 370–385.
13. Ellis, N. C., Römer, U., & O'Donnell, M. B. (2016). *Language usage, acquisition, and processing: Cognitive and corpus investigations of construction grammar*. Malden, MA: Wiley-Blackwell.
14. Goldberg, A. E., Casenhiser, D., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive Linguistics*, 15(3), 289.
15. Knoch, U. (2011). Rating scales for diagnostic assessment of writing: What should they look like and where should the criteria come from? *Assessing Writing*, 16(2), 81–96.
16. Kuiken, F., and I. Vedder. 2012. "Syntactic Complexity, Lexical Variation, and Accuracy as a Function of Task Complexity and Proficiency Level in L2 Writing and Speaking." In *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy, and Fluency in SLA*, edited by A. Housen, F. Kuiken, and I. Vedder, 143–169. Amsterdam: John Benjamins.
17. Kyle, K. 2016. *Measuring syntactic development in L2 writing: Fine-grained indices of syntactic complexity and usage-based indices of syntactic sophistication*. [Doctoral dissertation, Georgia State University].
18. Kyle, K., & Crossley, S. (2017). *Assessing syntactic sophistication in L2 writing: A usage-based*

- approach. *Language Testing*, 34(4), 513–535.
19. Kyle, K., & Crossley, S. A. (2018). Measuring Syntactic Complexity in L2 Writing Using Fine-Grained Clausal and Phrasal Indices. *The Modern Language Journal*, 102(2), 333–349.
  20. Lahuerta Martinez, A. (2018). Analysis of syntactic complexity in secondary education ELF writers at different proficiency levels. *Assessing Writing*, 35, 1–11.
  21. 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.
  22. Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Language Testing*, 34(4), 493–511.
  23. 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, 16e27
  24. Mahmoud Abdi Tabari, Sima Khezrlou & Yu Tian (2023): Verb argument construction complexity indices and L2 written production: effects of task complexity and task repetition, *Innovation in Language Learning and Teaching*, DOI:10.1080/17501229.2023.2211955
  25. Mostafa, T., and S. Crossley. 2020. “Verb Argument Construction Complexity Indices and L2 Writing Quality: Effects of Writing Tasks and Prompts.” *Journal of Second Language Writing* 49 (1): 1–23.
  26. Ninio, A. 1999. “Pathbreaking Verbs in Syntactic Development and the Question of Prototypical Transitivity.” *Journal of Child Language* 26 (3): 619–653.
  27. 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.
  28. Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24(4), 492e518.
  29. Ortega, L. 2015. “Syntactic Complexity in L2 Writing: Progress and Expansion.” *Journal of Second Language Writing* 29: 82–94. doi:10.1016/j.jslw.2015.06.008.
  30. Ryu, J. (2020). Predicting Second Language Writing Proficiency in the Different
  31. Genres of Writing Using Computational Tool *Korean Journal of Applied Linguistics* 36/1 (2020) pp. 141-170
  32. Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing*, 32, 17–35.
  33. Sung, Min-Chang & Hyunwoo Kim. 2022. Effects of verb–construction association on second language constructional generalizations in production and comprehension. *Second Language Research* 38(2). 233–257.
  34. Taguchi, N., Crawford, B., & Wetzel, D. Z. (2013). What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program. *TESOL Quarterly*, 47, 420–430.
  35. Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*. Honolulu, HI: University of Hawaii Press.