

Key Defining Linguistic Features in the Writing Performance of First-Year University Students Across Different Language Proficiency Levels

Chalatip Charnchairerk

chalatip.c@chula.ac.th, Chulalongkorn University Language Institute, Thailand

APA Citation:

Charnchairerk, C. (2022). Key defining linguistic features in the writing performance of first-year university students across different language proficiency levels. *LEARN Journal: Language Education and Acquisition Research Network*, 15(2), 858-891.

Received
25/04/2022

Received in
revised form
27/06/2022

Accepted
01/07/2022

Keywords
L2 writing;
proficiency
levels; syntactic
complexity;
lexical
complexity

Abstract

This study sought to investigate the key determining characteristics in the writing performance of first-year Chulalongkorn University students across language proficiency levels as measured by CU-TEP. The focus was on both syntactic and lexical complexity components. The sample comprised the writings from a corpus of 4,812 first-year students divided into four CEFR levels (C1, B2, B1, and A2), using CU-TEP and corresponding CEFR levels as the strata. The sample size of all four groups was identical comprising 50 students each, totaling 200 students. Multiple computational tools were utilized for data analysis. The findings revealed that the distinctive features typifying the most proficient writers include the production of longer as well as more clausally and phrasally complex sentences. They also demonstrated high lexical richness through the use of wide-ranging vocabulary and rare or sophisticated academic words. These features were also discovered in other less proficient groups but to a lesser extent at decreasing proficiency levels. It was also found that the syntactic complexity measures that better differentiated proficiency levels were: mean length of sentence, mean length of T-unit, and mean length of clause while all three lexical complexity indices were proven to be good predictors of L2 writing quality.

Introduction

Effective writing skills have grown in importance within the global community. As such, teaching writing is playing an increasingly important role in second language education (Padgate, 2008; Weigle, 2002), with good writing skills constituting a significant means of effective communication on the part of the writer. Despite its importance, writing is often viewed by learners as the most difficult of all the four skills in English (Fatemi, 2008; Richards & Renandya, 2002) since it requires a high level of language proficiency to generate and organize the writer's ideas in a way that enables the reader to understand their thoughts and ideas clearly and precisely. It is hardly surprising that students at all education levels often find themselves struggling with an academic writing task. As Celce-Murcia and Olshtain (2000) point out, even skilled writers at times find writing no easy task.

In an attempt to explore what makes an L2 learner a proficient writer, much research has focused on the notions of complexity, accuracy, and fluency (CAF), which are used as performance descriptors for both the oral and written assessments of language learners and the indicators of learners' proficiency underlying their performance (Housen & Kuiken, 2009). A number of studies have aimed to explore the distinguishing qualities of learner writing performance at different levels of English proficiency as measured by standardized proficiency tests such as IELTS, TOEFL, and Cambridge English examinations, among others with the focus on various linguistic aspects and diversified groups of learners. In the Thai EFL context, on the other hand, several studies have placed greater emphasis on the error analysis of students' writing (e.g. Kampookaew, 2020; Sermsook et al., 2017), or improving their grammatical accuracy (Tan & Manochphinyo, 2017). Addressing CAF, Thongyoi and Poonpon (2020) looked into phrasal complexity measures that can predict EFL students' writing performance. Although there are a number of studies on lexicon, they largely concentrate on such aspects as the relationship between vocabulary size and reading comprehension or between vocabulary size and years of study.

Very few studies so far have dealt with syntactic complexity in the company of lexical complexity. It is hoped that this present study can contribute to a better understanding of the distinctive characteristics of students' writing performance manifested at each level of proficiency, enabling the teacher to design writing courses and prepare materials that

further improve the writing skills of more proficient students while minimizing less proficient students' weaknesses.

To that effect, in response to the notion that syntactic complexity is an important measure of development and writing proficiency in L2 (Gustin, 2019; Kyle & Crossley, 2018; Larsen-Freeman, 1978; Lu, 2011), the current study intends to investigate both the syntactic and lexical aspects of complexity to provide a richer educational context around students' writing proficiency. Unlike previous studies focusing on the written performance of L2 learners from different backgrounds such as age, nationality, L1, and the like, this study aims to analyze the writing of a relatively homogeneous group of first-year university students for features regarding syntactic and lexical complexity that might typify their levels of English proficiency as determined by CU-TEP.

Literature Review

A growing body of research has adopted the complexity component of the CAF triad, – the other two being accuracy and fluency – as the indices of language proficiency and writing quality. This is in tandem with the views of several scholars that one way of detailing language learners' proficiency is to rely on the indices that measure complexity features (Bulté & Housen, 2015; Kovacevic, 2018; Lu, 2014). The complexity component is further divided into syntactic and lexical complexity.

Syntactic Complexity

Syntactic complexity is commonly characterized as the degree of variation, elaboration, and sophistication of the grammatical structures used in language production (Housen & Kuiken, 2009; Lu, 2011; Ortega, 2003; Wolf-Quintero et al., 1998; Zhang & Lu, 2022). It is considered an important measure for assessing L2 writing proficiency (Larsen-Freeman, 1978; Larsson & Kaatari, 2020; Lu, 2011) since more proficient writers tend to produce more complex syntactic features (Kim, 2014). They also generate syntactically longer texts (Grant & Ginther, 2000; Kim, 2014; Ortega, 2003). Several studies have investigated the syntactic aspects of complexity.

For instance, a study by Kyle and Crossley (2018), which measured syntactic complexity in L2 writing using fine-grained clausal and phrasal indices, found that the fine-grained indices of phrasal complexity were

better predictors of writing quality than traditional or fine-grained clausal indices. The results substantiated Biber et al.'s (2011) findings about complexity and academic L2 writing proficiency. Similarly, Thongyoi and Poonpon (2020) examined phrasal complexity measures as predictors of EFL university academic writing proficiency. Three measures were found to potentially predict the writing proficiency of more proficient students. Unlike these two studies, Chuenchaichon (2011) investigated the syntactic complexity of writing produced by Thai EFL writers, using clauses per T-unit ratio, dependent clauses per clause ratio, and dependent clauses per T-unit ratio. These clausal complexity measures were manifested as feasible measures.

To identify the best indicators of L2 writing proficiency with respect to text length, lexical complexity, and syntactic complexity, Kim (2014) identified different uses of linguistic features across learners' proficiency levels. More proficient writers produced longer texts, used a more diverse range of vocabulary, and wrote more words per sentence and complex nominals.

Empirically, there appears disagreement among syntactic complexity researchers on the validity and reliability of indices of syntactic complexity that can best predict learners' writing proficiency (Zhang & Lu, 2022). Wolf-Quintero et al. (1998) suggest multiple measures for further exploration, some of which include words per T-unit, words per clause, clauses per T-unit, dependent clauses per clause among others. These length-based and subordination-based measures have been adopted out of a common belief that longer production units and higher amounts of subordination reflect higher levels of fluency (Gustin, 2019; Liu & Li, 2016). For example, Chen et al. (2014) reported significant correlations between mean length of sentence and mean length of clauses and the quality of Chinese EFL university students' writing (Zhang & Lu, 2022).

While these assertions have been supported by several scholars (e.g. Banerjee et al., 2007; Chuenchaichon, 2011), the complexity measures have been criticized in terms of their usefulness in the descriptive account of language (Larsson & Kaatari, 2020). For instance, such measures as mean range of T-unit or clausal construct (e.g. subordinate clauses per T-unit) fail to capture any differences between clausal and phrasal complexity (Biber et al., 2020; Larsson & Kaatari, 2020). According to Biber et al. (2011), since clausal complexity is considered a more prominent feature of conversation than academic writing, the use of clausal measures to assess proficiency in novice academic writing could

undermine their reliability. Instead, they recommend phrasal complexity as preferable.

Numerous studies (e.g. Kyle & Crossley, 2018; Lu, 2011) have confirmed Biber et al.'s claims about phrase-level structures as more reliable measures. In a similar vein, many scholars view a reliance on the amount of clausal subordination alone as problematic as research has shown that an increase in the level of noun phrasal complexity is related to higher proficiency writing (e.g. Biber et al., 2016; Kyle & Crossley, 2018; Lu, 2011), leading to a call for more attention to fine-grained features that take into consideration the structural types of clause and phrases (Biber et al., 2016; Kyle & Crossley, 2018; Zhang & Lu, 2022).

Another area that has garnered much research attention is the automatic analysis of syntactic complexity, using wide-ranging computational tools. Kyle (2016) developed the tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC), which features a broad spectrum of fine-grained linguistic measures. Lu (2010), on the other hand, developed the L2 Syntactic Complexity Analyzer (L2SCA), which includes 14 indices of syntactic complexity for analyzing texts written by L2 learners. Several studies (e.g. Gustin, 2019; Kim, 2014) have used L2SCA.

Although research tends to show a shift from clausal complexity to phrasal sophistication or phrasal complexity (Gustin, 2019), a number of L2 writing researchers (e.g. Gustin, 2019; Kyle & Crossley, 2018; Norris & Ortega, 2000) are inclined to embrace the idea that syntactic complexity should be perceived as a multidimensional construct, encompassing complexity at different structural levels (global, clausal, and phrasal). This has led scholars in the field to investigate both length and subordination measures as well as phrasal complexity, nominal complexity, and coordination (Gustin, 2019; Polio, 2017). This view will be justified in the present study.

Lexical Complexity

Lexical complexity is defined as a multidimensional feature of language use. It measures the proportion of relatively rare or sophisticated words in the learners' writings (Zhang et al., 2022) and can predict writing quality as well as the degree of formality (Zhu, 2013). Lexical complexity encompasses diversity, sophistication, and density (Bayazaidi et al., 2020; Read, 2000; Wolf-Quintero et al., 1998).

In essence, lexical diversity refers to the size of the lexicon which has traditionally been measured by means of type-token ratio (TTR) measures (Michel, 2017). However, the measures have been found to be affected by text length or sample size (Bayazaidi et al., 2020; Malvern & Richards, 2002). A more robust method of measuring lexical diversity called D has been proposed by Malvern and Richards (2002) to address the issue of text length (Banerjee et al., 2007; Bayazaidi et al., 2020). According to Skehan (2009), D indicates the extent to which the speaker avoids recycling the same set of words. A text with a lower D suggests the person often relies on the same set of words (Bayazaidi et al., 2020).

Lexical sophistication, on the other hand, refers to the depth of lexis measured by such measures as frequency of rare or academic words (Michel, 2017). Some researchers (e.g. Jarvis et al., 2003) use average word length to gauge lexical sophistication, assuming that longer words tend to be more sophisticated. However, it has been argued that this measure fails to predict writing quality or effectively differentiate proficiency levels (Bayazaidi et al., 2020; Verspoor et al., 2012). In contrast, Lu's (2012) Lexical Complexity Analyzer considered words, lexical words, and verbs not featured on the 2,000 most frequent words list as ranked by the British Nation Corpus (Kim, 2014).

Several studies have identified the use of lexical density, lexical diversity, and lexical sophistication as features that discriminate proficiency levels. To illustrate, Hawkey and Barker (2004) analyzed a corpus of scripts written by candidates at three Cambridge ESOL examination levels to identify the key distinguishing features in learners' writing performance. One interesting finding was that range of vocabulary was possibly a feature determining proficiency levels. Likewise, Banerjee et al. (2007) also confirmed that writing performance at increasing IELTS band levels displayed greater lexical variation and sophistication.

Lexical density is a term often used to describe the proportion of content words (nouns, verbs, adjectives, and adverbs) to the total number of words (Johansson, 2008). By investigating lexical density, Johansson (2008) suggests the notion of information packaging. That is, a text with a high proportion of content words (or lexical words) contains more information than one featuring a high proportion of function words (prepositions, articles, conjunctions, pronouns). A number of studies (e.g. To et al., 2013; Zhang et al., 2021) revealed that the lexical density of EFL students was generally low, which affected the quality of their writing since most of the writings had few content words and were written in a

complex way instead of choosing the exact words to convey the meaning (Syarif & Putri, 2018).

Like syntactic complexity, various automatic tools have been developed to analyze lexical complexity. Chief among them are L2 Syntactic Complexity Analyzer (Lu, 2012), TAALES (Tool for the Automatic Analysis of Lexical Sophistication, Kyle & Crossley, 2014) and LCA-AW (Lexical Complexity Analyzer for Academic Writing, Nasser & Lu, 2020) V. 2.2.

Methodology

Research Questions

The purpose of this study was to investigate the key distinguishing characteristics in the writing performance of first-year Chulalongkorn University students across language proficiency levels as measured by CU-TEP with the focus on the syntactic and lexical dimensions of the complexity component of CAF. The study sought to address the following questions:

1. What are the key defining features in the writing performance of students at four different English proficiency levels with respect to:
 - a. syntactic complexity?
 - b. lexical complexity?
2. Which syntactic complexity measures better differentiate each proficiency level?
3. Which lexical complexity measures better differentiate each proficiency level?

Participants

All of the 200 participants ($N=200$) in this study were selected through stratified random sampling from a cohort of 4,812 students across the university, taking CU-TEP and a compulsory fundamental English course offered by Chulalongkorn University Language Institute (CULI) in 2015, based on their CU-TEP scores and corresponding CEFR proficiency levels (C1, B2, B1 and A2)¹. Despite studying 18 different disciplines, the research population were mostly Thai native speakers, of the same age group (around 17-19), and with the same prior 12-year period of English study. They comprised 1,949 males and 2,863 females.

Corpus Selection

A total of 4,812 students taking CU-TEP and CULI's fundamental English course in 2015 were divided into four CEFR proficiency levels (C1, B2, B1, and A2), using CU-TEP scores and corresponding CEFR levels as strata. Systematic random sampling was then used to sample 50 students from each of the four proficiency lists for consistency in the number for a sample size of 200. The hand-written scripts of the writing task of the 200 students chosen comprising 90 males and 110 females were re-typed into text files, including all errors, for analysis of selected linguistic features, using L2SCA, the Text Inspector lexical profiling tool, and LCA-AW V.2.2. Table 1 shows CU-TEP scores and their corresponding CEFR levels. Table 2 presents the breakdown of the population at each proficiency level while Table 3 displays an overview of the corpora included in this study.

Table 1

CU-TEP Score and CEFR Levels²

CU-TEP (Max 120 points)	CEFR Levels
14-34	A2
35-69	B1
70-98	B2
99-120	C1

Table 2

Breakdown of the Population at each Proficiency Level

Proficiency Level	Number of Students	Percent
C1	67	1.39
B2	1,113	23.13
B1	3,224	67.00
A2	408	8.48
Total	4,812	100

Variables and Computational Tools

CU-TEP

The Chulalongkorn University Test of English Proficiency (CU-TEP) is an in-house standardized test in a multiple-choice format (four choices per question). It consists of 30 listening comprehension questions, 60 reading comprehension questions, and 30 semi-writing questions in the form of error identification for a total of 120 test items. Primarily designed to assess students' academic proficiency, the test is required for admission to graduate programs and undergraduate English programs at Chulalongkorn University (Chulalongkorn University Academic Testing Center, 2007) as well as multiple other educational institutions in Thailand. As all first-year Chulalongkorn University students in the Thai programs take the test prior to the start of their first semester, they can track their language learning progress throughout their time at the university.

Writing Task

During the academic years 2014-2015, as part of the Chulalongkorn University Language Institute (CULI)'s initiative to evaluate first-year students' actual writing ability not directly addressed by CU-TEP, the first-year students were assigned to write an opinion paragraph of a minimum of 80 words on the same general topic generated by the researcher, the then-deputy director for Academic Affairs, within 30 minutes in their first week of their English classes. The students were informed the task was not part of their course grades.

L2SCA

The L2 Syntactic Complexity Analyzer (L2SCA) is a computational tool which generates syntactic complexity measures for analysis of texts written by L2 learners. Developed by Lu (2010), the tool includes 14 indices of syntactic complexity of the sample drawn from Ortega (2003).

Text Inspector

This is a lexical profiling tool that analyzes the vocabulary content of texts, especially for L2 learners. The tool allows users to acquire detailed information regarding complexity and lexical diversity, among others. Its

four corpora include English Vocabulary Profile (EVP), British National Corpus (BNC), Academic Word List (AWL), and Corpus of Contemporary American English (COCA) (Text Inspector, 2022).

LCA-AW V.2.2

A modified version of the LCA (Lexical Complexity Analyzer, Lu, 2012), the tool computes the lexical complexity of texts, using 25 lexical measures that can indicate learners' linguistic proficiency. LCA-AW V. 2.2 is suitable for analyzing the lexical complexity of academic writing. It is used in the present study to complement Text Inspector in measuring lexical density.

Table 3

Overview of the Corpora

	C1	B2	B1	A2	Total
Number of texts in each corpus	50	50	50	50	200
Word count	4547	4,660	4,172	3,177	16,556
Average words per text	90.94	93.20	83.44	63.54	82.78

Syntactic Complexity Measures

The five syntactic complexity measures for the present study (Table 4) were taken from Gustin (2019), who selected from the fourteen indices of syntactic complexity in the L2SCA (Lu, 2010). They include two measures of clausal complexity (sentence complexity ratio and dependent clauses ratio) and three measures of phrasal complexity (mean length of clause³, coordinate phrases per clause, and complex nominals per clause). Another two measures (mean length of sentence and mean length of T-unit⁴) were added by the researcher as influenced by the findings of Kim (2014) since they were deemed to best fit the purpose of this current study which sought to examine syntactic complexity of learner writing at length-based as well as clausal and phrasal levels. Table 4 shows the syntactic complexity measures used in the present study.

Table 4

Syntactic Complexity Measures Used in the Present Study (adapted from Gustin (2019) and Kim (2014))

Type	Measure	Definition
Length of production unit	Mean length of sentence	# of words / # of sentences
	Mean length of T-unit	# of words / # of T-units
	Mean length of clause	# of words / # of clauses
Clausal complexity	Sentence complexity ratio	# of finite clauses / # of sentences
	Dependent clauses ratio	# of dependent clauses / # of clauses
Phrasal complexity	Coordinate phrases per clause	# of coordinate phrases / # of finite clauses
	Complex nominals per clause	# of complex nominals / # of finite clauses

Lexical Complexity Measures

Indices from various computational tools were selected to measure different dimensions of lexical complexity. Combined, these measures could generate a clear lexical profile of the learners at each proficiency level in line with McCarthy and Jarvis (2010) suggesting that more than one single index be used because different indices can capture unique lexical diversity information. Table 5 displays the lexical complexity measures used in this study.

Table 5

Lexical Complexity Measures Used in the Present Study (adapted from Kalantari & Gholami (2017))

Lexical Complexity	Measures	Definition	Tools
Density	Lexical Density	The proportion of content (lexical) words to the total number of words (Johanssen, 2008)	LCA-AW V.2.2

Lexical Complexity	Measures	Definition	Tools
Diversity	Measure of Textual Lexical Diversity (MTLD)	The mean length of sequential word strings in a text that maintain a given TTR (type-token ratio) (McCarthy & Jarvis, 2010)	A module in Text Inspector
	Vocabulary Diversity (vocd-D)	A mathematical algorithm applied to TTR as a solution to the problem of using TTR for language samples of varying sizes (Malvern & Richards (1997)	A module in Text Inspector
Sophistication	Academic Word List (AWL)	A list of 570 head words and their connected sublists that occur frequently in academic texts (Coxhead, 2000)	A module in Text Inspector

Results

The results for each complexity measure adopted in this study include descriptive statistics with a box plot that shows how the data is spread out. Five pieces of information included in the diagram are the minimum, first quartile (the 25% mark), median, third quartile (the 75% mark), and the maximum (Glen, n.d.). The descriptive statistics are accompanied by statistical results that display pairwise comparisons between the four groups under study.

Mean Length of Sentence

Table 6 summarizes the medians for mean length of sentence across different proficiency levels.

Table 6

Medians for Mean Length of Sentence

Level	C1	B2	B1	A2
Median	19.33	18.43	16.50	14.20

Figure 1

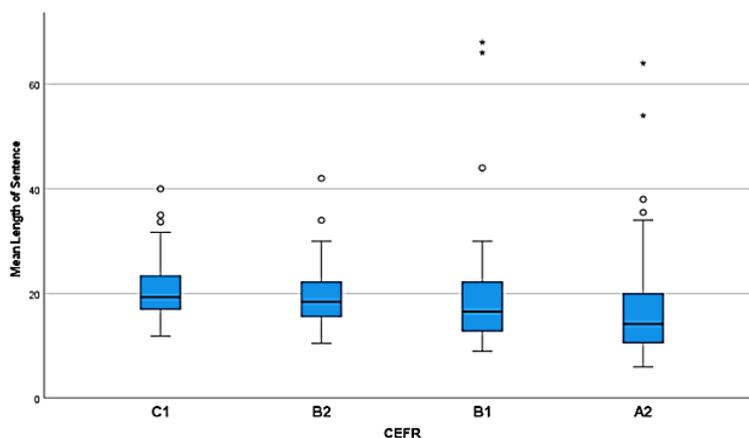
Box Plots for Mean Length of Sentence

Figure 1 shows the C1 group had less variation than the B1 and A2 groups and on average had the most number of words per sentence, followed closely by the B2 group. The A2 plot, in contrast, shifted significantly toward the low values as seen by the longer whisker toward the higher values. This could mean the more proficient students tended to more consistently produce longer sentences.

To determine if there were statistically significant differences between each group of students across all English proficiency levels, Skewness and Kurtosis statistics were used to test the normality of the variable's distribution. As the data in each group was found non-normally distributed, the Kruskal-Wallis H test was chosen.

The adjusted p -value obtained from performing post-hoc pairwise comparisons using Bonferroni Correction shows there was significant difference in the medians of mean length sentence between the A2-B2 pair ($H=31.960$, adjusted $p=0.035$) and A2-C1 pair ($H=44.730$, adjusted $p=0.001$). This means both C1 and B2 groups significantly produced longer sentences than the lower proficient A2 group.

Mean Length of T-Unit

Table 7 exhibits the median values for mean length of T-unit. It is worth noting the medians of the B2 and B1 groups are very similar and close to that of the C1 group with the A2 group trailing far behind.

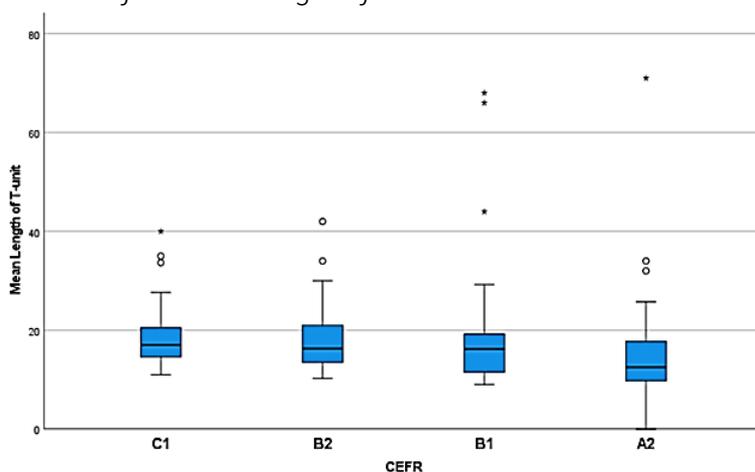
Table 7

Medians for Mean Length of T-unit

Level	C1	B2	B1	A2
Median	17.00	16.30	16.23	12.5

Figure 2

Box Plots for Mean Length of T-unit



The C1 plot shows data values that are the most consistent with its median value as manifested by the smallest box. It can be concluded that on average the C1 group consistently produced more words per T-unit than other groups. While the B2 and B1 group had very similar median values, the A2 plot shifted toward the lower values down to 0.00.

The pairwise comparisons also reveal a significant difference in the medians of mean length of T-unit between the A2-B2 pair ($H=41.500$, adjusted $p=.002$) and A2-C1 pair ($H=51.500$, adjusted $p=.000$). This reflects a similar trend as the mean length of sentence. That is, more proficient writers displayed significantly more syntactic complexity in terms of words per T-unit than writers of lower proficiency.

Mean Length of Clause

Table 8 presents the median values for mean length of clause. While the C1, B2, and B1 groups differed in the number of words per

clause by only 1-2 words, the A2 group trailed behind with only 1.83 words per clause.

Table 8

Mean Length of Clause

Level	C1	B2	B1	A2
Median	10.00	9.00	8.28	1.83

Figure 3

Box Plots for Mean Length of Clause

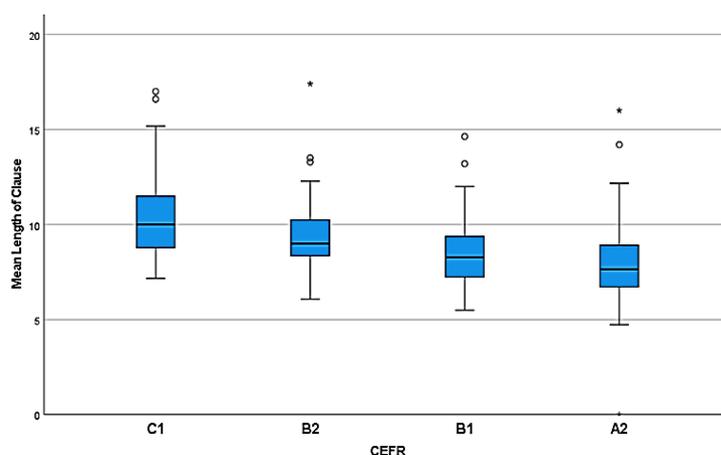


Figure 3 indicates a similar level of variation across the box plots. It also shows that the median values for mean length of clause declined at decreasing proficiency levels. The pairwise comparisons show significant differences in the medians between the A2-B2 pair ($H=45.030$, adjusted $p=.001$), A2-C1 pair ($H=65.340$, adjusted $p=.000$), and B1-C1 pair ($H=48.390$, adjusted $p=.000$). This affirms that more proficient writers had a higher mean length of clause than lower proficient writers. The results reveal that all three length-based measures could significantly discriminate the writing performance of students across proficiency levels.

Sentence Complexity Ratio

Table 9 exhibits the median values for sentence complexity ratio (clauses per sentence). At different proficiency levels, these were very similar. The C1 and B1 groups had the same median value of 2.00.

Table 9

Sentence Complexity Ratio

Level	C1	B2	B1	A2
Median	2.00	1.94	2.00	1.83

Figure 4

Box Plots for Sentence Complexity Ratio

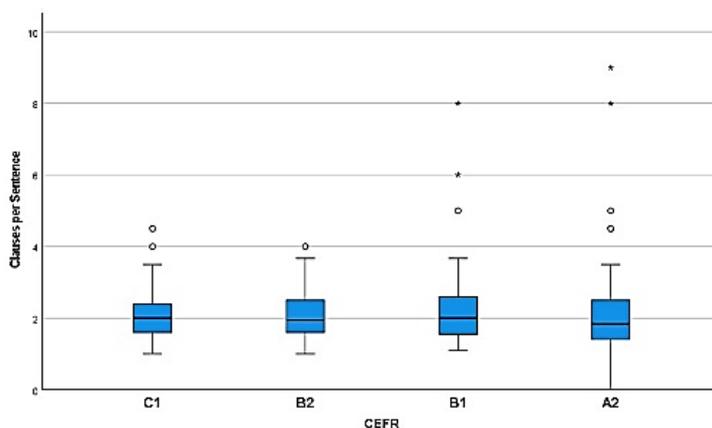


Figure 4 demonstrates similar levels of variation in the C1, B2, and B1 groups. Based on the diagram and the median values in Table 9, it can be concluded that all four groups had nearly the same number of clauses per sentence. Also, the Kruskal-Wallis H test did not find any statistically significant differences as the medians for sentence complexity were equal across all four groups ($H=1.513$, $p=.679$). In other words, sentence complexity ratio was not a feature discriminating the students' writing performance at different proficiency levels.

Dependent Clauses Ratio

Table 10 displays the median values for dependent clauses ratio (dependent clauses per clause). The C1 and B2 groups had nearly the same median values. Likewise, the B2 and A2 groups showed very similar results despite being slightly lower than the previous two groups.

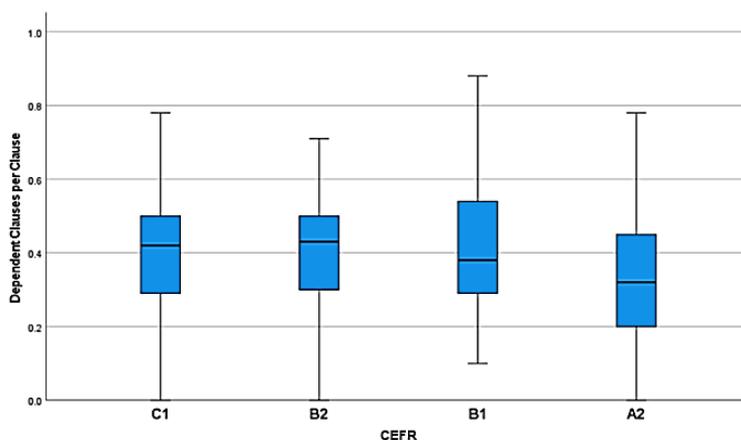
Table 10

Dependent Clauses Ratio

Level	C1	B2	B1	A2
Median	0.42	0.43	0.38	0.32

Figure 5

Dependent Clauses Ratio



All the box plots in Figure 5 indicate a high level of variation. Except for the B1 plot, the other three plots shifted toward the value of 0.00, which could mean some students used no dependent clauses in their writings.

The Kruskal-Wallis H test found a statistically significant difference in the medians for the dependent clauses ratio ($H=8.530$, $p=0.036$). However, the pairwise comparisons reveal no significant difference between each pair.

The results show that sentence complexity ratio and dependent clauses ratio, which are the two measures adopted in this present study

to investigate clausal complexity, could not distinguish the writing quality of students across proficiency levels.

Coordinate Phrases per Clause

Table 11 summarizes the median values for coordinate phrases per clause. The C1 group displayed the highest median value while the B2 group demonstrated the lowest.

Table 11

Coordinate Phrases per Clause

Level	C1	B2	B1	A2
Median	0.29	0.21	0.23	0.22

Figure 6

Coordinate Phrases per Clause

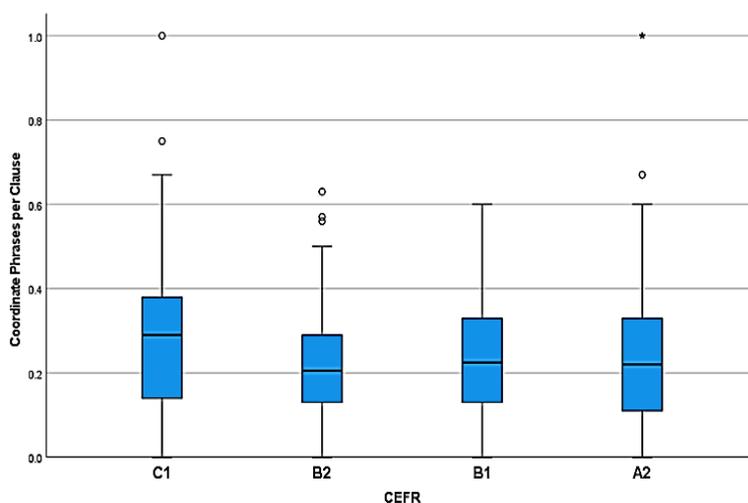


Figure 6 exhibits a high level of variation across the box plots. Interestingly, all the four plots shifted toward the value of 0.00, meaning there were some students in every group who used no coordinate phrases. On average, however, the C1 group had the highest median value for coordinate phrases per clause compared to other groups. The B2, B1, and A2 groups had about the same number of coordinate phrases per clause in their writings.

The Kruskal-Wallis H test found no significant difference in the median values for coordinate phrases per clause ($H=3.834$, $p=.280$). This phrasal complexity measure was, therefore, not a good predictor of students' writing performance.

Complex Nominals per Clause

Table 12 exhibits slightly different median values for complex nominals per clause across the groups.

Table 12

Complex Nominals per Clause

Level	C1	B2	B1	A2
Median	1.29	1.25	1.13	1.10

Figure 7

Complex Nominals per Clause

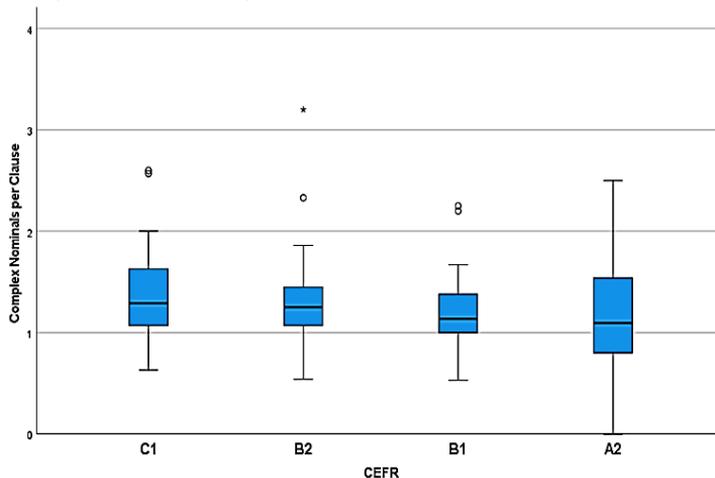


Figure 7 shows the highest variation in the A2 group with the plot shifting toward the value of 0.00. Based on the diagram, on average, there was an increased use of complex nominals per clause with proficiency levels.

The Kruskal-Wallis H test, however, did not find any statistically significant difference in the median values for complex nominals per clause ($H=5.679$, $p=.128$). The results suggest that like coordinate phrases

per clause, complex nominals per clause was not a distinguishing feature of students' writing performance across different proficiency levels.

Lexical Density

Table 13 demonstrates the median values for lexical density analyzed by LCA-AW V.2.2 on the American National Corpus (ANC) and NLTK Python Library. While the A2 group had the highest median value, the C1 and B2 had the lowest.

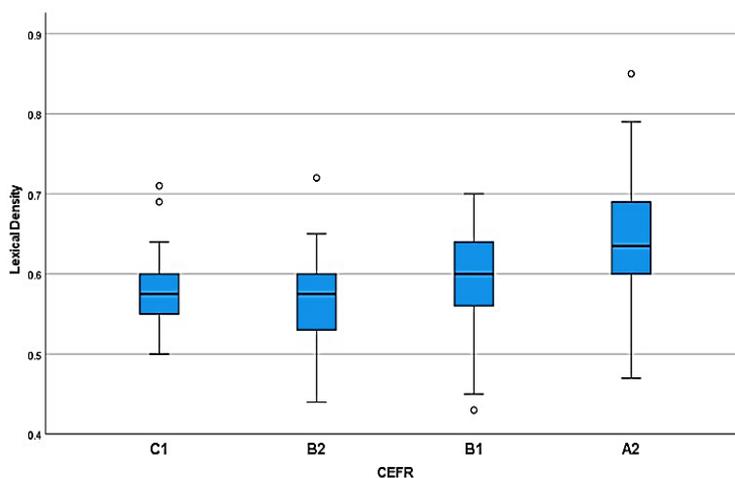
Table 13

Lexical Density

Level	C1	B2	B1	A2
Median	0.58	0.58	0.60	0.64

Figure 8

Lexical Density



The box plots show that the A2 group had the highest level of variation and the highest median value for lexical density. The median values decreased at increasing proficiency levels. Despite having the same median value as the B2 group, the C1 group had less variation, and 50% of the data seemed to be evenly distributed in the first and third quartile.

The pairwise comparisons demonstrate significant differences in the medians for lexical density between the B2-A2 pair ($H=-61.090$, adjusted $p=.000$), the C1-A2 pair ($H=-54.070$, adjusted $p=.000$), and the B1-A2 pair ($H=-32.640$, adjusted $p=.028$). This means the less proficient writers had a higher lexical density than their more proficient counterparts.

Lexical Diversity

Tables 14 and 15 exhibit the median values derived from One-way ANOVA test, using MLTD and vocd-D as measures of lexical diversity. The results in both tables were consistent in that the C1 group had the higher median than the other groups.

Table 14

One-way ANOVA Test Result of MLTD

Level	C1	B2	B1	A2
Median	88.25	71.17	59.50	47.92

Table 15

One-way ANOVA Test Result of vocd-D

Level	C1	B2	B1	A2
Median	90.02	75.67	57.63	53.88

Figure 9

One-way ANOVA Test Result of MLTD

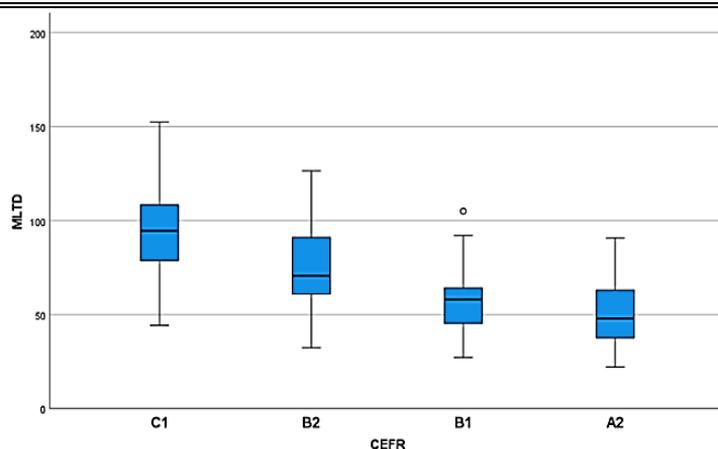
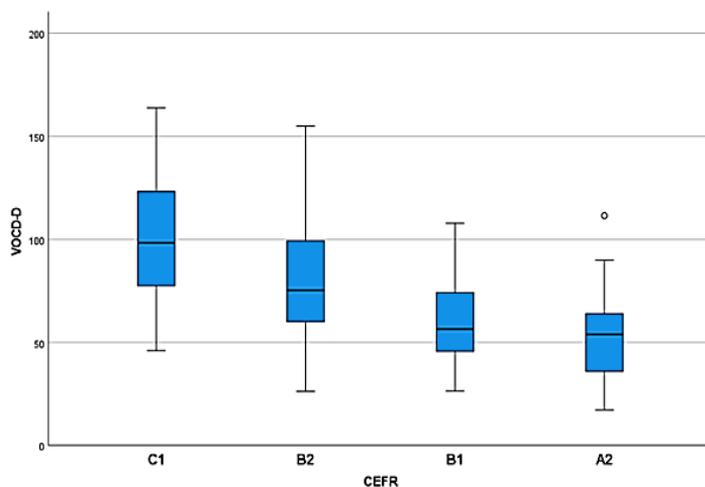


Figure 10

One-way ANOVA Test Result of vocd-D



The box plots in Figures 9 and 10 exhibit a very similar trend in that the C1 group had the highest median value and an almost symmetric data distribution as shown by the equal length of the tails and the median at the center of the box. The medians of the other three groups declined at decreasing proficiency levels.

To determine whether there were statistically significant differences between each group of students, the One-way ANOVA test was performed because the Skewness and Kurtosis statistic found the data to be normally distributed.

The p -value obtained by performing the post-hoc test reveals that the average MLTD score significantly differed in every pair of student groups except for the A2-B1 pair. Similarly, the average vocd-D score exhibited the same results, showing significant differences in all but the A2-B1 pair. This confirms lexical diversity as a distinctive characteristic of students' writing quality.

Lexical Sophistication

To examine lexical sophistication, the AWL tool in Text Inspector analyzed the students' scripts according to the Academic Word List (Coxhead, 2000). The list is classified into 10 sublists (AWL1-AWL10) based on the frequency of occurrence of word in the Academic Corpus. That is, words in AWL1 occur more frequently in academic texts than those in the next sublists in descending order. The K1-K5 sublists, on the other end, include phrasal expressions (e.g. *have to*, *there is/are*, *such as*, *going to*) and are also categorized by their descending order of frequency. Table 16 summarizes the median values for coverage percentage of the words in all the sublists for each group.

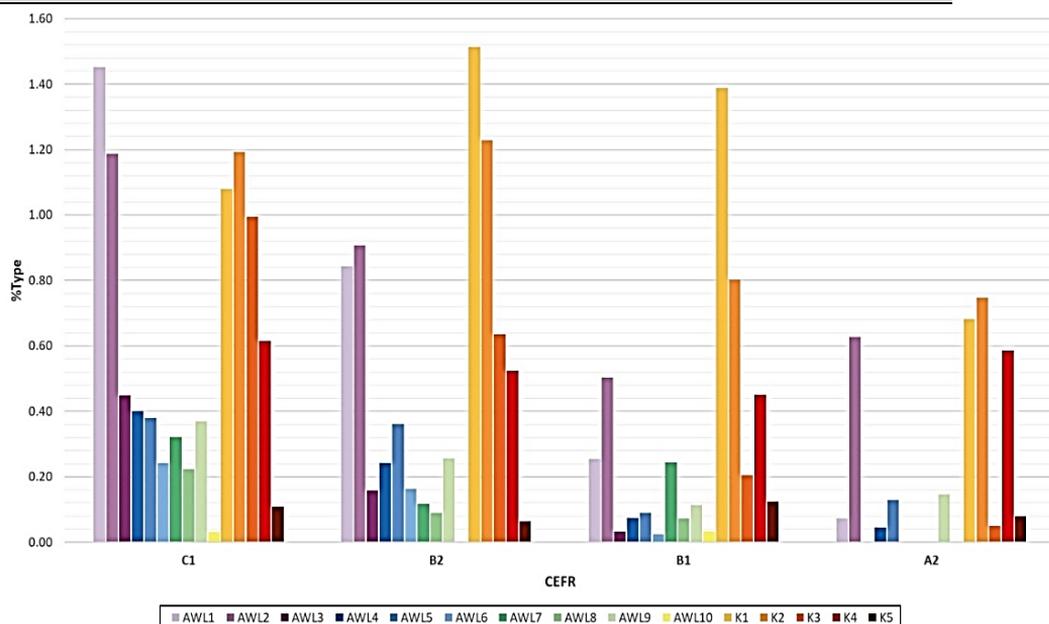
Table 16

Coverage Percentage of Words in All Sublists

Level	C1	B2	B1	A2
Median	4.94	2.86	0.00	0.00

Figure 11

Coverage Percentage of Words in All Sublists



The bar chart illustrates that the C1 group had the most lexical sophistication, using a diverse range of words across the sublists. The B2 group shows the highest use of the most frequently occurring phrasal expressions followed by the B1 group. It is worthy of note that the A2 group exhibited the tendency to use more of the phrasal expressions than words in the Academic Word List. Based on the data, it can be concluded that more proficient writers use more sophisticated and rare words, making lexical sophistication another good predictor of students' writing quality.

Discussion

This study aims to investigate key defining features in the writing performance of students across different English proficiency levels with a focus on syntactic and lexical complexity. The findings reveal that there were statistically significant differences in mean length of sentence, mean length of T-unit, and mean length of clause across all groups, and the mean length of production unit increased with proficiency levels. The results substantiate Chen et al. (2014) reporting significant correlations between mean length of sentence and mean length of clause and learners' writing quality. They are also consistent with the findings of Kim (2014), Grant and Ginther (2000), and Ortega (2003) that more proficient writers are more likely to generate syntactically longer texts. The positive results from

the investigation of all the three length of production unit measures in this study could possibly help justify the usefulness of the length-based syntactic complexity indices.

Regarding clausal complexity, the research found that sentence complexity ratio and dependent clause ratio failed to differentiate the students' writing quality, suggesting that students at all proficiency levels produced relatively the same average number of clausally complex sentences. In this study, for a sentence to be clausally complex, it needs to have at least one main clause and at least one finite dependent clause (Gustin, 2019). For example:

While hazing rituals are considered a rather controversial topic nowadays, it is not necessarily a bad thing.

In this sentence, the underlined part is the subordinate or finite dependent clause while the one in italics is the main clause.

The results of this study contradicted those of Chuenchaichon (2011) which found indices of clausal complexity to be feasible measures. This was probably because less experienced L2 writers tended to use simple sentence structures instead of more complex sentences through subordination. For instance:

I have a new friend. I feel lucky and so happy. I like senior very good take care.

As for phrasal complexity, although multiple scholars have suggested that indices of phrasal complexity are better predictors of L2 writing quality (e.g. Biber et al., 2011; Kyle & Crossley, 2018), the same did not hold true in this present study as the investigation of the two phrasal complexity measures (coordinate phrases per clause and complex nominals per clause) revealed that phrasal complexity was not a feature that could distinguish proficiency levels even though more proficient students appeared to produce more phrasally complex sentences. It is worth pointing out that in this study, a sentence is considered phrasally complex if within a clause, there is at least a coordinate phrase (i.e. adjective, adverb, noun, or verb phrase), and/or at least a complex nominal (i.e. noun plus adjective, possessive, prepositional phrase, relative clause, participle, or appositive, nominal clause, and gerunds and infinitives in the subject position) (Gustin, 2019). Take for example:

University hazing rituals, a major part of freshmen's academic **and** social life is beneficial to students.

This sentence is phrasally complex as it contains a long coordinated noun phrase (underlined).

Another striking result is in the realm of lexical complexity. In contrast with some studies (e.g. To et al., 2013; Zhang et al., 2021) which indicated that the lexical density of EFL students was generally low, the less proficient students in this study were found to have a higher level of lexical density, and the density level decreased at increasing proficiency levels. One possible explanation is that in an attempt to convey information in their writing, some less proficient writers will probably write the way they talk – the so-called conversation writing – resulting in more words used without appropriate punctuation. For example:

I love to enjoy it sometimes I thought I prefer activity than studying Ha Ha Ha just kidding and I will show you some activities such as dance... .

With respect to lexical diversity, the results exhibited statistically significant differences in lexical diversity across proficiency levels, suggesting that the writings of more advanced learners have a much higher lexical diversity or a greater amount of lexical variation, validating lexical diversity as a good predictor in determining writing quality. This could be attributable to the learners' wider range of vocabulary which they could draw upon and use, avoiding frequent reliance on the same set of words.

By the same token, it was found that more proficient writers demonstrated more lexical sophistication, using many of the words in all AWL sublists and beyond, reflecting their large lexical repertoire. Conversely, less proficient writers relied more on common phrasal expressions such as *have to, lead to, there are, used to, a good deal, instead of, as long as*, and the like probably as a result of their repeated use of these expressions in their previous writings. The finding aligns with Banerjee et al. (2007), whereby writing performance at increasing proficiency levels exhibits greater lexical variation and sophistication. As such, it can be concluded that a range of vocabulary is a possible feature discriminating proficiency levels (Hawkey & Barker, 2004).

According to the findings, in answer to the first research question, distinctive syntactic and lexical features typify the most proficient writers, including the production of longer clauses as well as longer and more

clausally and phrasally complex sentences. They also demonstrated high lexical richness through the use of a broad range of vocabulary and diverse, sophisticated academic words. These features were also found in all the other groups but to a lesser extent. In other words, on average, the C1 group displayed the most of these distinguishing traits in their writing. The B2 group exhibited the second most and so on. In response to research questions 2 and 3, the study revealed that of the seven syntactic complexity measures investigated in this study, the indices that could significantly differentiate proficiency levels were mean length of sentence, mean length of T-unit, and mean length of clause, all of which were in the category of length of production unit. In contrast, all the three lexical complexity indices examined were found to be good predictors of L2 writing quality across different proficiency levels.

Conclusion

This study aimed to analyze and acquire detailed insights into what contributes to the writing quality of first-year students at different levels of proficiency with respect to syntactic and lexical complexity. Pedagogically, the results of this study suggest that the writing quality of higher and less proficient students are significantly different in syntactic structures and lexical use. Thus, the practice of using the same writing coursebook for students across proficiency levels without taking into consideration the varying degrees of students' language competence is far from ideal. If need be, it is advisable that teachers prepare supplementary materials specifically tailored to enrich advanced writers' writing proficiency and offset less proficient students' apparent linguistic weaknesses. For instance, teachers may want to introduce some rhetorical devices that advanced writers could use to achieve a specific effect or to assign more challenging writing tasks that command the students' creative writing skills, empowering them to fully maximize their writing proficiency and become successful writers. At the other end of the proficiency scale, beginner writers could benefit from materials that help enhance their grammatical competence and build lexical richness, the two areas in which they seem to falter most. To facilitate such exercises, students should be grouped according to their proficiency levels, allowing them to progress at their own pace. Alternatively, a non-credit prep writing course designed to teach necessary writing skills for less proficient students could be offered to first-year students prior to the start of the semester to enable them to

master the knowledge of grammatical structures and vocabulary needed for academic writing.

Although the findings of this study may contribute meaningfully to the current literature, one issue that was not addressed was the aspect of grammatical accuracy. A future study in this area would be of great help in verifying whether there exists a trade-off between the writers' syntactic and lexical complexity and grammatical accuracy. Also, with each relatively short script in the corpus, the findings might not be transferable to longer scripts. Further investigations of different types of writing with longer texts will allow a greater degree of accuracy on matters of syntactic and lexical complexity to be established.

Acknowledgements

I would like to extend my sincere gratitude to Chulalongkorn University Academic Testing Center for their permission in allowing the use of students' CU-TEP score data. My appreciation also goes to Asst. Prof. Dr. Rananda Rungnaphawet for his meaningful advice and support as well as Mr. Nachanok Potiwattutham, who assisted with data preparation.

Endnotes

¹ The CEFR six proficiency levels are grouped into three broad levels: Basic User (A1, A2), Independent User (B1, B2), and Proficient User (C1, C2) (Council of Europe Portal).

² Based on the research on Mapping CU-TEP to Common European Framework of Reference (CEFR) (Wudthayagorn, 2018)

³ Mean length of clause in the L2SCA and in the present study appears in the category of length of production unit.

⁴ T-unit (Hunt, 1965) refers to "one main clause plus whatever subordinate clauses happen to be attached or embedded within it".

About the Author

Chalatip Charnchairerk: An Assistant Professor at the Chulalongkorn University Language Institute (CULI). She has been involved with CU-TEP over the past several years as a test writer and an editor. Her research interests encompass language assessment, testing, and translation studies. She can be reached at chalatip.c@chula.ac.th

References

- Banerjee, J., franceschina, F., & Smith, A. M. (2007). *5. Documenting features of written language production typical at different IELTS band score levels* (5). British Council. https://www.ielts.org/-/media/research-reports/ielts_rr_volume07_report5.ashx
- Bayazaidi, A., Ansarin, A.-A., & Mohammadnia, Z. (2020). Lexical complexity as a function of task type and proficiency level in the speech monologs of Iranian EFL Learners. *Journal of Modern Research in English Language Studies*, 7(1), 29-44. <https://doi.org/10.30479/jmrels.2019.11406.1418>
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45(1), 5-35. <https://doi.org/https://doi.org/10.5054/tq.2011.244483>
- 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. <https://doi.org/10.1093/applin/amu059>
- 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>
- Bulté, B., & Housen, A. (2015). Evaluating short-term changes in L2 complexity development. *Círculo de Lingüística Aplicada a la Comunicación*, 63, 42-76. http://dx.doi.org/10.5209/rev_CLAC.2015.v63.50169
- Celce-Murcia, M., & Olshtain, E. (2000). *Discourse and context in language teaching: A guide for language teachers*. Cambridge University Press.
- Chen, H., Xu, J., & He, B. (2014). Automated essay scoring by capturing relative writing quality. *The Computer Journal*, 57(9), 1318-1330. <https://doi.org/10.1093/comjnl/bxt117>
- Chuenchaichon, Y. (2011). *The development of paragraph writing for EFL writers through the use of a reading into writing method* [Ph.D. thesis, University of Reading].
- Chulalongkorn University Academic Testing Center. (2007). *CU-TEP*. http://www.atc.chula.ac.th/en_html/en_tep.html

- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238. <https://doi.org/10.2307/3587951>
- Fatemi, M. A. (2008). *The relationship between writing competence, language proficiency and grammatical errors in the writing of Iranian TEFL sophomores* [Ph.D. thesis, Universiti Sains Malaysia, Malaysia]. <https://core.ac.uk/download/pdf/11933914.pdf>
- Glen, S. (n.d.). *Box Plot (Box and Whiskers): How to Read One & How to Make One in Excel, TI-83, SPSS*. StatisticsHowTo.com: Elementary Statistics for the rest of us!
<https://www.statisticshowto.com/probability-and-statistics/descriptive-statistics/box-plot/>
- Grant, L., & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing*, 9(2), 123-145. [https://doi.org/10.1016/S1060-3743\(00\)00019-9](https://doi.org/10.1016/S1060-3743(00)00019-9)
- Gustin, S. (2019). *Differences in syntactic complexity in the writing of EL1 and ELL civil engineering students* [Master thesis, Portland State University].
- Hawkey, R., & Barker, F. (2004). Developing a common scale for the assessment of writing. *Assessing Writing*, 9(2), 122-159. <https://doi.org/10.1016/j.asw.2004.06.001>
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30(4), 461-473. <https://doi.org/10.1093/applin/amp048>
- Jarvis, S., Grant, L., Bikowski, D., & Ferris, D. (2003). Exploring multiple profiles of highly rated learner compositions. *Journal of Second Language Writing*, 12(4), 377-403. <https://doi.org/10.1016/j.jslw.2003.09.001>
- Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Lund University, Department of Linguistics and Phonetics Working Papers*, 53, 61-79. <https://journals.lub.lu.se/LWPL/article/view/2273/1848>
- Kalantari, R. & Gholami, J. (2017). Lexical Complexity Development from Dynamic Systems Theory Perspective: Lexical Density, Diversity, and Sophistication. *International Journal of Instruction*, 10(4), 1-18. <https://doi.org/10.12973/iji.2017.1041a>
- Kampookaew, P. (2020). An analysis of grammatical errors made by Thai EFL university students in an EAP writing class: Issues and recommendations. *rEFlections*, 27(2), 246-273. <https://so05.tci-thaijo.org/index.php/reflections/article/view/248862>

- Kim, J. (2014). Predicting L2 writing proficiency using linguistic complexity measures: A corpus-based study. *English Teaching*, 69(4), 27-51. <https://doi.org/10.15858/engtea.69.4.201412.27>
- Kovacevic, E. (2018). The relationship between lexical and syntactic complexity measures in a learner corpus. In S. Gudurić & B. Radić-Bojanić (Eds.), *Jezici i Kulture u Vremenu i Prostoru VII/2* (pp. 469-479). University of Novi Sad. https://www.academia.edu/37808815/The_Relationship_Between_Lexical_and_Syntactic_Complexity_Measures_in_a_Learner_Corpus
- Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication*. Dissertation, Georgia State University. doi:<https://doi.org/10.57709/8501051>
- 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. <https://doi.org/10.1111/modl.12468>
- Larsen-Freeman, D. (1978). An ESL index of development. *TESOL Quarterly*, 12(4), 439-448. <https://doi.org/10.2307/3586142>
- Larsson, T., & Kaatari, H. (2020). Syntactic complexity across registers: Investigating (in)formality in second-language writing. *Journal of English for Academic Purposes*, 45, Article 100850. <https://doi.org/10.1016/j.jeap.2020.100850>
- Liu, L., & Li, L. (2016). Noun phrase complexity in EFL academic writing: A corpus-based study of postgraduate academic writing. *The Journal of Asia TEFL*, 13(1), 1-71. <https://doi.org/10.18823/asiatefl.2016.13.1.4.48>
- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15(4), 474-496. <https://doi.org/10.1075/ijcl.15.4.02lu>
- 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. <https://doi.org/10.5054/tq.2011.240859>
- Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190-208. https://doi.org/10.1111/j.1540-4781.2011.01232_1.x
- Lu, X. (2014). *Computational methods for corpus annotation analysis*. Springer. <https://doi.org/10.1007/978-94-017-8645-4>

- Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing, 19*(1), 85-104.
<https://doi.org/10.1191/0265532202lt221oa>
- McCarthy, P. M., & Jarvis, S. (2010). MTL, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods, 42*(2), 381-392.
<https://doi.org/10.3758/BRM.42.2.381>
- Michel, M. (2017). Complexity, accuracy, and fluency in L2 production. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 50-68). Routledge.
<https://doi.org/10.4324/9781315676968>
- Nasseri, M., & Lu, X. (2020). *Updated LCA-AW for Python 3, Lexical Complexity Analyzer for Academic Writing, version 2.2*. Zenodo.
<https://doi.org/10.5281/zenodo.4147000>
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417-528. <https://doi.org/10.1111/0023-8333.00136>
- 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.
<https://doi.org/10.1093/applin/24.4.492>
- Padgate, W. (2008). Beliefs and opinions about English writing of students at a Thai university. *PASAA Journal, 42*, 31-53.
<https://www.culi.chula.ac.th/publicationsonline/files/article/IAUy wNh8siTue82813.pdf>
- Polio, C. (2017). Second language writing development: A research agenda. *Language Teaching, 50*(2), 261-275.
<https://doi.org/10.1017/S0261444817000015>
- Read, J. (2000). *Assessing Vocabulary*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511732942>
- Richards, J. C., & Renandya, W. A. (Eds.). (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge University Press.
- Sermsook, K., Liamnimitr, J., & Pochakorn, R. (2017). An analysis of errors in written English sentences: A case study of Thai EFL students. *English Language Teaching, 10*(3), 101-110.
<https://doi.org/10.5539/elt.v10n3p101>

- Skehan, P. (2009). Lexical performance by native and non-native speakers on language-learning tasks. In B. Richards, M. H. Daller, D. D. Malvern, P. Meara, J. Milton, & J. Treffers-Daller (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application* (pp. 107-124). Palgrave Macmillan. https://doi.org/10.1057/9780230242258_7
- Syarif, H., & Putri, R. E. (2018). How lexical density reveals students' ability in writing academic text. *Lingua Didaktika: Jurnal Bahasa dan Pembelajaran Bahasa*, 12(2), 86-94. <https://doi.org/10.24036/ld.v12i1.10408>
- Tan, K. E., & Manochphinyo, A. (2017). Improving grammatical accuracy in Thai learners' writing: Comparing direct and indirect written corrective feedback. *The Journal of Asia TEFL*, 14(3), 430-442. <https://doi.org/10.18823/asiatefl.2017.4.3.4.430>
- Text Inspector. (2022). *Measure lexical diversity*. Weblingua Ltd. <https://textinspector.com/help/lexical-diversity/>
- Thongyoi, K., & Poonpon, K. (2020). Phrasal complexity measures as predictors of EFL university students' English academic writing proficiency. *rEFLECTIONS*, 27(1), 44-61. <https://so05.tci-thaijo.org/index.php/reflections/article/view/241750>
- To, V., Fan, S., & Thomas, D. (2013). Lexical density and readability: A case study of English textbooks. *Internet Journal of Language, Culture and Society*(37), 61-71. <http://aaref.com.au/en/publications/journal/journal-articles/issue-37-2013/>
- Verspoor, M., Schmid, M. S., & Xu, X. (2012). A dynamic usage based perspective on L2 writing. *Journal of Second Language Writing*, 21(3), 239-263. <https://doi.org/10.1016/j.jslw.2012.03.007>
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511732997>
- Wolf-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy & complexity*. University of Hawai'i, Second Language Teaching & Curriculum Center.
- Zhang, H., Chen, M., & Li, Z. (2021). Developmental features of lexical richness in English writings by Chinese beginner learners. *Frontiers in Psychology*, 12, Article 665988. <https://doi.org/10.3389/fpsyg.2021.665988>

- Zhang, S., Zhang, H., & Zhang, C. (2022). A dynamic systems study on complexity, accuracy, and fluency in English writing development by Chinese university students. *Frontiers in Psychology, 13* Article 787710.
- Zhang, X., & Lu, X. (2022). Revisiting the predictive power of traditional vs. fine-grained syntactic complexity indices for L2 writing quality: The case of two genres. *Assessing Writing, 51*, Article 100597. <https://doi.org/10.1016/j.asw.2021.100597>
- Zhu, H. (2013). *Developmental features of lexical richness in English writings by Chinese EFL students*. Shanghai Foreign Language Education Press.