A Cross-sectional Comparison of First and Second Year Thai EFL Student Writing: Syntactic, Phrasal, and Lexical Features

KIM MCDONOUGH*
Department of Education, Faculty of Arts and Science, Concordia University, Canada
JINDARAT DE VLEESCHAUWER
Department of English, Faculty of Humanities, Chiang Mai University, Thailand
Corresponding author email: kim.mcdonough@concordia.ca

Article information	Abstract
	This cross-sectional study compares the written language of Thai EFL students in their first two years of university study. First- and second-year students (N = 170) wrote opinion paragraphs by hand in response to six prompts. Using automated textual analysis tools, clausal (subordination), phrasal (coordinated phrases and complex nominals), and lexical (AWL use and lexical diversity) measures were obtained. Matched-pairs were created by pairing different first- and second-year students from the same faculty of study who responded to the same writing prompt. The results indicated that second-year students produced significantly more complex nominals and AWL words than the first-year students with effect sizes ranging from small to medium. Implications are discussed in terms of pedagogical approaches and assessment in EFL settings, and suggestions for future research are provided.

INTRODUCTION

Written language development researchers have explored how writers' use of linguistic features, such as accuracy, complexity, fluency, and the frequency or emergence of lexical, morphological, or syntactic forms, changes over time (Polio & Park, 2016). Interpreted broadly, change over time can encompass a wide range of intervals, ranging from a few weeks of targeted instruction to multiple years of English study. Similarly, change may occur in a variety of directions, with features becoming more or less frequent, accurate, complex, or fluent. Furthermore, such changes may have positive, negative, or null effects on judgements about text quality. In other words, writers may change how they use language over time without those changes resulting in "better" texts. Researchers have identified changes in language features by adopting longitudinal designs that track individual writers over time or cross-sectional designs that compare students from different educational or proficiency levels. Although longitudinal research designs provide the greatest insight into how individual writers develop over time, cross-sectional comparisons provide a snapshot of language use by writers at different stages in the learning process. Findings from both research design types can provide instructors and administrators with useful information for making curricular and assessment decisions, such as what types of language forms students can be expected to produce at certain proficiency

levels or in specific classes within a program. To help generate such information, the current study adopts a cross-sectional design to compare the language features of opinion paragraphs written by Thai EFL students enrolled in the first and third English classes in a required four-class sequence.

Previous studies

Turning to previous research that specifically investigated the language features of English L2 argumentative writing, which is similar to the opinion paragraphs investigated here, many studies have used automated text analysis programs to analyze multiple language features associated with syntactic and phrasal complexity along with lexical diversity, breadth, and depth. In an early cross-sectional study using automated text analysis, Lu (2011) analyzed timed argumentative essays written by Chinese EFL students across four years of university study. He used the Syntactic Complexity Analyzer (SCA: Lu, 2010), which includes 14 measures organized into five types: length of production, sentence complexity, subordination, coordination, and specific structures (complex nominals and verb phrases). He found that subordination decreased only from year 2 to year 4, while increased coordinated phrases differentiated year 1 from years 2 and 4. Phrasal complexity, however, increased over time both for adjacent years (1 vs 2 and 2 vs 3) and non-adjacent years. In sum, the cross-sectional comparison revealed that the use of complex nominals (such as nouns modified by prepositional phrases and nouns with adjectives) and coordinated phrases increased with the students' year in university while subordination decreased.

Also adopting a cross-sectional design, Yoon (2017) analyzed morphological, syntactic, and lexical features of argumentative essays written by Chinese university students. However, whereas Lu (2011) used year of study as a marker of development, Yoon used four proficiency levels (A2, B1.1, B1.2, B2) in the Common European Framework of Reference (CEFR) to differentiate among writers. Using a variety of automated text analysis tools to measure morphological, lexical, and syntactic features, Yoon found that length measures (e.g., sentence), phrasal measures (complex nominals per clause and coordinated phrases per clause), verbal morphological complexity, and two lexical features (diversity and word length) significantly increased with proficiency. Post-hoc analysis of five features revealed that significant increases only occurred between non-adjacent levels. As an example, significant differences occurred when A2 was compared to the B1.2 and B2 levels, but not when compared to the B1.1 level. Thus, Yoon confirmed the linear increase for phrasal measures reported by Lu (2011) and provided evidence for the development of lexical features, specifically diversity and sophistication, which was measured as word length.

Adopting a longitudinal design, Yoon and Polio (2017) explored whether ESL students enrolled in an intensive English program showed language development at the end of a 15-week ESL course where they had received approximately 8-10 hours of writing instruction per week. In addition to the SCA measures, they also included lexical measures (word length, word frequency, and lexical diversity), accuracy (errors per 100 words), and fluency (text length). Although the ESL students showed some development over time in narrative texts, no syntactic features exclusively in their argumentative texts showed any significant increases or decreases over

time. There were also no significant changes in the students' accuracy, fluency, or lexical complexity in the argumentative texts. Their findings for syntactic complexity contrast with those of Lu (2011) and Yoon (2017), both of whom reported significant increases in coordinated noun phrases and complex nominals. However, whereas Yoon and Polio examined development across one semester, the other two studies compared students at different years of study or proficiency levels. The time interval for documenting development (years versus months), along with the difference in research design (cross-sectional versus longitudinal), may help account for the conflicting findings.

An additional longitudinal study that similarly included both syntactic and lexical measures compared the argumentative texts written by university students at the beginning and end of an intensive writing course (Mazgutova & Kormos, 2015). Students planning to begin undergraduate studies (referred to as Group 2) used more words from the Academic Word List (AWL, Coxhead, 2000) and less frequent content words while writing sentences with greater lexical diversity and less semantic similarity at the end of the semester. For the syntactic features, although unit length and subordination did not change over time, the students' use of complex nominals and modifiers per noun phrase increased. Their findings provide additional evidence that phrasal features show a developmental trajectory, even in contexts where the interval for assessing development is relatively short (four weeks). However, their lexical findings, including AWL use, contrast with the null findings for lexical features reported previously (Yoon & Polio, 2017). The AWL may be a particularly useful measure in classes with students from different disciplines, as instructors are likely to focus on common underlying academic vocabulary rather than discipline-specific words (Eldridge, 2008).

To sum up, despite differences in context, research design, and specific linguistic measures, previous studies of student argumentative writing have reported consistent increases in phrasal measures (Lu, 2011; Mazgutova & Kormos, 2015; Yoon, 2017) with some exceptions (Yoon & Polio, 2017). Clausal features related to subordination have shown mixed findings, with studies reporting both increases (Lu, 2011) and null effects (Mazgutova & Kormos, 2015; Yoon & Polio, 2017). Similarly, studies that included lexical measures have also reported conflicting findings in terms of both null effects (Yoon & Polio, 2017) as well as increased lexical diversity or AWL use (Mazgutova & Kormos, 2015; Yoon, 2017).

Research question

The goal of this cross-sectional study is to contribute to this line of research by comparing the opinion paragraphs written by first- and second-year university students enrolled in the first and third classes in a required four-class EFL sequence. While our selection of linguistic features was informed by the findings of the research studies described previously, we also considered their potential usefulness for instructors in this Thai EFL university setting. To facilitate positive washback, we selected features used in previous research that bear some resemblance to how instructors teach and assess language in writing in these classes at this university. For example, subordination is relevant to instruction about sentence types (i.e., complex and compound-complex sentences) and punctuation (e.g., sentence fragments and run-on sentences). Similarly, complex nominals are relevant when teaching students how to modify nouns, such as the use

of adjectives, participles, relative clauses, and prepositional phrases to describe nouns. Although many lexical measures exist for capturing potential development, we focused on two measures that are conceptually simple and useful for students and instructors, specifically lexical diversity and AWL words. While they are not referred to by these terms, instructors orient to these measures when encouraging students to use a variety of words in their texts and to chose academic words. Guided by these considerations, our research question was as follows: Are there differences in clausal, phrasal, and lexical features in the English opinion paragraphs written by first- and second-year Thai university students?

METHOD

Participants

The texts in the current study (*n* = 170) came from a larger collection of opinion paragraphs written by first- and second-year students at Chiang Mai University as part of a curriculum evaluation project (De Vleeschauwer, 2020). The first year students were taking the first of four required English classes in their first semester at the university, while the second year students were enrolled in the third class in that sequence. The first three classes in the sequence target foundational English skills and mix students from different disciplines, while the fourth class is tailored for students in specific degree programs (i.e., English for Fine Arts or English for Humanities). Due to variability in the academic profiles of students across faculties, first- and second-year students were paired based on their faculty, which included Agriculture (8), Agricultural Industry (8), Associated Medical Sciences (26), Economics (4), Education (14), Engineering (34), Fine Arts (10), Humanities (26), Mass Communication (14), Political Science (2), Science (22), and Social Science (2). Their background information is provided in Table 1. Descriptively there were few differences in the distribution of gender or their mean scores on Likert-scale questionnaires about writing anxiety (Cheng, 2004) and writing self-efficacy (Abdel-Latif, 2015), both of which were administered for a different research project. While they are not the focus of analysis here, we reported the mean scores to demonstrate that the first- and second-year students had similar levels of anxiety (possible scores ranging from 22 to 110, Cronbach's alpha = .93) and self-efficacy (possible scores ranging from 18 to 90, Cronbach's alpha = .91). The second-year students were a year older and had more English study while the first-year students wrote longer texts.

Table 1
Background information by year of study

	Year 1 (n = 85)	Year 2 (n = 85)
Gender	36 men, 49 women	24 men, 61 women
Age	18.5 (.6)	19.7 (.8)
English study	12.9 (2.3)	14.4 (1.7)
Writing anxiety	60.9 (15.5)	65.0 (14.2)
Writing self-efficacy	49.4 (9.7)	50.1 (9.7)
Text length: Total words	10,821	7,850
Text length: Mean (SD)	127.31 (55.10)	92.35 (34.95)

Writing task

The writing task was a general interest prompt modeled after the opinion essay topic from task two of the IELTS writing test. A general interest prompt that did not require the use of information from sources was selected because students at this university do not write longer, source-based essays until their fourth and final English class unless they enroll in an English minor. Writing is not emphasized in the first two classes, and the third class introduces paragraph writing. An initial list of 15 prompts were compiled through an online search of IELTS practice essay topics and then presented to 18 students at the same university for ranking based on topic interest (1 = *high*, 15 = *low*). Based on their rankings, the six highest-ranked prompts were used to elicit written texts by randomly assigning a topic to classes. Three prompts solicited opinions about changes in society related to traditional values, family relationships, and social media use, while three prompts concerned social problems about pay equity, gender equity, and climate change. All six prompts (Appendix A) presented a position toward the topic and asked whether students agreed (i.e., "To what extent do you agree or disagree with this opinion?"). The prompt was followed by instructions to explain whether they agreed with the position and to give reasons for their opinion. The handout provided space for students to make notes prior to writing if they wished to do so, followed by lined paper for composing their texts. The students wrote texts in response to all six prompts, and first- and second-year students from the same faculty were matched based on writing topic: Social media (n = 50), climate change (n = 32), pay equity (n = 34), family relationships (n = 30), traditional values (n = 16), and gender equity (n = 8).

Procedure

The writing task was administered in the English classes in the middle of a 16-week semester. After completing consent and background information forms, the students filled out questionnaires about writing anxiety and writing self-efficacy, which were administered for a study unrelated to writing development, after which they had 30 minutes to handwrite their texts, which is the same amount of time provided in previous studies (Yoon & Polio, 2017). In this context prior to the global pandemic, students were required to handwrite all in-class writing assignments and timed writing exams to avoid the possibility of internet use. The instructions did not suggest a word length to avoid intimidating first-semester students and to discourage second-year students from stopping as soon as they reached the suggested length. The students were informed that the writing task would not contribute to their class grades, but they would receive feedback about content, organization, and language use. While writing, the students worked individually without access to any resources.

Data coding and analysis

The students' handwritten texts were typed by research assistants (RAs) who ignored all crossed out words and included insertions. They corrected minor spelling and capitalization errors so that automated programs would recognize words and sentence boundaries. For example, if a student wrote the words "transpotation" and "abelity", the RAs typed them correctly as "transportation" and "ability", respectively. Similarly, if a student used capital letters inappropriately

(e.g., "First of all, Everyone"…), the unnecessary capitalization was removed (i.e., "First of all, everyone"…). Each typed text was verified by an additional RA to check for typing accuracy and to ensure that the RAs had not inadvertently corrected any other language issues.

The typed texts were analyzed in terms of clausal and phrasal complexity along with lexical features. For the complexity measures, Lu's (2010) online SCA tool (https://aihaiyang.com/software/) provided measures of clausal complexity (dependent clauses per clause), coordinated phrases (coordinated phrases per clause), and complex nominals (complex nominals per clause), with complex nominals including nouns with prenominal modifiers (adjectives, nouns, and possessive nouns), nouns with postnominal modifiers (prepositional phrases, relative clauses, participles and appositives), nominal clauses, and gerund and infinitive subjects. Table 2 provides examples from the student texts to illustrate the three features.

Table 2
Summary of complexity measures

Feature	Measurement	Example
Clausal complexity	Dependent clauses per clause	*If you share personal information in social media*, it can make problems.
Coordinated phrases	Coordinated phrases per clause	Noun phrases: Other workers like *film actors, company bosses and athletes* might have *the imagination, vision and skills* to do their work. Verb phrases: One factory *spreads carbon dioxide and causes bad water.*
Complex nominals	Complex nominals per clause	(People *who are good at teaching*) aren't working as teachers because of (*low pay*).

For the lexical features, online automated text analysis tools available at https://www.linguisticanalysistools.org/ were used to identify AWL words (calculated as AWL words divided by total words) and lexical diversity operationalized as the original measure of textual lexical diversity (MTLD) for content words.

The language features in the first- and second-year texts were compared using paired-samples t-tests. Although they are commonly used for within-groups designs where the same individual provides two data points, paired-samples t-tests are also appropriate when comparing samples that were purposefully matched for variables that may influence the dependent variables, which in the current study were the variables of faculty and writing topic. To reduce the possibility of Type 1 error caused by running multiple tests on a relatively small data set, alpha was divided by the number of comparisons, which yielded a value of .010 (.05/5).

FINDINGS

The research question asked whether there were differences in the clausal, phrasal, and lexical features of EFL texts written by first- and second-year Thai university students. The mean scores

for the clausal, phrasal, and lexical measures by year of study are provided in Table 3. The students' use of dependent clauses was identical, but descriptively the second-year students produced more complex nominals and coordinated phrases. However, only the increase in complex nominals reached statistical significance with a small effect size based on applied linguistics benchmarks (i.e., .40, Plonsky & Oswald, 2014). For the lexical features, the students had similar lexical diversity values, but the second-year students' greater use of AWL words was statistically significant with a medium effect size (i.e., .70).

Table 3
Written language features by English course

	Year 1 (n = 85)		Year 2 (n = 85)		Comparison		
	M	SD	M	SD	t	p	d
Dependent clauses	.43	.15	.43	.16	.09	.924	.00
Coordinated phrases	.28	.21	.31	.27	.86	.392	.12
Complex nominals	1.04	.43	1.25	.43	3.16	.002	.49
AWL words	.03	.02	.05	.03	5.38	.001	.78
Lexical diversity	44.14	18.02	41.58	18.35	1.08	.283	.14

Thus, the findings indicate that two aspects of written language showed an increase over the first two years of university study: complex nominals and AWL words.

An example of greater use of complex nominals is provided by the two texts in Table 4, from students in the Faculty of Education who wrote about whether traditional values are still important in contemporary society. The first-year text (94 words) has several single-word (e.g., *people, anybody, they*) and two-word noun phrases (e.g., *our social, the little, their personality, a look*). There are few instances of prenominal modification in the form of adjectives or nouns (e.g., *a good people*) or post-nominal modification (e.g., *the outside of others*). In contrast, the second-year text (78 words) has several noun phrases with post-nominal modification in the form of relative clauses (e.g., *that they once knew, that we were born into...*), including a noun phrase with both a prepositional phrase and relative clause (i.e., *the value of all people who are born on this earth*). In addition, there are several prenominal adjectives (e.g., *the above message, their true value, perfect body*)

Table 4
Comparison of complex nominals

Year 1	Year 2
I agree with this statement because in our social today they have change how people look. Why? Because we can't trust anybody now. Some people see only the outside of others. So they didn't know how is he/she or how good their personality is, we see only the little. So if you want to know how to judge people, you need to see deep inside him/her. You can't just take a look and know how good is him/her. If you want to know a good people you need to know how to look people.	From the above message, a person's worth nowadays seems to be judged according to social status, salary, and material possessions. I think that the value of all people who are born on this earth is to see their true value, not the power, money, and reputation as misunderstandings that they once knew. Moreover, regardless of what position we were born into perfect body or not, every race, religion, we tend to misunderstand that we value ourselves from others.

Turning to examples that illustrate the increased use of AWL words, Table 5 provides representative texts. These two students from the Faculty of Agriculture, whose students typically perform poorly in English based on course grades from previous semesters, wrote texts about whether children receive enough attention from parents in modern societies. The actual writing prompt (in Appendix) did not contain any AWL words. Although it is longer (157 words), the first-year text does not have any AWL words. In contrast, the second-year text (108 words) has four AWL types with five tokens (*media* occurred twice). The AWL words come from three different sub lists: Sublist 1(*financial*), Sublist 3 (*technology*), and Sublist 7 (*media, release*).

Table 5
Comparison of AWL use

Year 1	Year 2
I agree with this extent because the family for me is a most important thing for child's future. The love & attention from mom & dad can made their child feel of love and grow up to be a good person. But now you will see some parents or many. Father and mother don't have time for their child because of many lessons like working. But for me if they can manage time. I think they can have a little time for playing with their child. But maybe there a lot of lesson that I didn't know. But I agree with this extent and believe that love from parents and attention is the important thing for every child. No child want to feel lonely or not important with their family. The good modern lifestyles everyone in the family have to love & spend time together that was why I think the family is a most important thing.	The statement show about modern lifestyles that many parents have little time for their children and I agree with the statement. Nowaday, people give important about *financial* more than their children. It cause that many parents have little time for their children and not have time to do activities together. When they go to work, they're usually *release* their children with *technology* and other social *media* and your children will receive bad things from social *media*. And children will be nervous. So they should be take care your children and give time for to do activities with children. Like play sport, watching TV or play game with them.

DISCUSSION

To summarize the findings, this comparison of EFL texts written by first- and second-year university students showed evidence of language development in terms of increased complex nominals and AWL words. Similar to previous studies that compared writers across years of study (Lu, 2011) and proficiency levels (Yoon, 2017) or tracked students longitudinally (Mazgutova & Kormos, 2015), these Thai university students increased their use of nouns with prenominal or postnominal modification, nominal clauses, and gerund or infinitive subjects. Large-scale multi-dimensional comparisons (e.g., Biber et al., 2002) have demonstrated that academic writing (e.g., journal articles, textbooks, course pack materials) is associated with phrasal complexity while oral discourse contains clausal coordination and subordination. Using complex phrases in academic writing, such as a noun with a prepositional phrase or relative clause, reflects information density and helps writers communicate knowledge efficiently. Although these Thai EFL writers are still in the early stages of acquiring academic writing skills, the findings suggest that their use of complex phrases may be beginning to move in an expected direction and might benefit from focused instruction in phrasal complexity as suggested in

prior research (Bowen & Thomas, 2020; Lan et al., 2019).

For AWL words, these writers increased their use of list words by their second year, which confirms prior studies that demonstrated increased AWL usage over time in a university writing course (Mazgutova & Kormos, 2015). Their AWL token use increased from a mean of 3% of total words in year one to 5% in year two, which is similar to the increase shown by lower proficiency pre-academic students at a British university (3.6% to 5.8%, Mazgutova & Kormos, 2015). Based on corpus studies of academic texts across disciplines, AWL words account for approximately 7–12% of total words (see Coxhead, 2011, for a summary). Clearly these EFL writers are still in the process of acquiring AWL words and have not yet achieved usage levels found in academic texts. However, the increase over time indicates improved ability to deploy AWL words when writing without any sources and under time pressure.

In terms of the clausal feature of subordination, our findings are in line with the null findings reported in several previous studies (Yoon, 2017; Yoon & Polio, 2017; Mazgutova & Kormos, 2015). In contrast to the writers in Lu's study (2011) whose subordination rates decreased over four years of university (from .33 to .30), our EFL students' use of subordination was consistently higher (.43 for both years). Our subordination rates were similar to those in the argumentative texts written by ESL students (Yoon & Polio, 2017) at the beginning of an ESL course (.40). Although these EFL writers' total rate of subordination did not change, it is possible that there were more qualitative changes in clausal features. Although a fine-grained analysis of clauses was beyond the scope of the current study, a post-hoc review of the student texts revealed some qualitative differences. For example, nonfinite clauses with "to" occurred after a wide range of verbs in the second-year texts but tended to follow common verbs like "want" and "need" in the first-year texts. Instructors may find Biber's developmental framework for complexity features (Biber et al., 2011) helpful when determining which types of subordinate clauses should be targeted across different levels of instruction.

Finally, the findings for lexical diversity have failed to confirm studies that reported an increase of lexical diversity across proficiency levels (Yoon, 2017) or during an EAP course (Mazgutova & Kormos, 2015). Due to differences in measurement (i.e., D versus MTLD; all words versus content words), it is difficult to compare across studies. It is possible that these EFL writers' lexical development emerged with the increased use of AWL words while lexical diversity remained constant. Increasing the time interval from two to four years of university study might provide greater insight into when their lexical diversity increases and reaches rates achieved by more proficient writers.

As mentioned in the introduction, our focus was on language development in writing as opposed to text quality (see Crossley, 2020 for writing quality versus development). The findings revealed that these writers used more complex nominals and AWL words over time but drawing any inferences about whether these features made their texts better is unwarranted. Nevertheless, the question as to what makes good writing good also has important implications for instruction and assessment. For example, instructors might want to teach students to use features that are associated with good texts and assess their texts in terms of how frequently or how well those features are used. For example, recent studies have found that phrasal features predict

the scores assigned by raters to argumentative essay exams (Thongyoi & Poonpon, 2020) and grades given by instructors on source-based research papers (Casal & Lee, 2019). Additional assessment research is needed to clarify which linguistic features (phrasal, clausal, lexical) are associated with both human judgments of text quality and writing development. It is also important for instructors to consider whether their goal is to assess language development or text quality. In some instructional settings where writing is used as a tool to promote language development, it may be more important to evaluate student writing in terms of how much their written language use develops as opposed to the quality of their texts.

Having demonstrated differences in the use of complex nominals and AWL words by these first- and second-year EFL university students, the findings may be useful for assessment practices at this university and in settings with similar general EFL classes. As Polio (2017) has observed, linking rubric descriptors to observed developmental patterns may help instructors and students recognize changes in their written language use. At this Thai university, the rubrics used to evaluate written paragraphs have three categories (content, organization, and language) and assign more points to content than organization or language. Furthermore, although the number of points available in each category is specified, descriptions of performance at different point levels are not provided other than generic terms such as "good" "average" and "poor." By combining syntactic, phrasal, and lexical features into the single category of "language," the rubric may not be detailed enough for instructors to document change in the students' written language use within or across semesters. Based on our findings, it would be helpful to separate the grammatical and lexical assessment categories. In addition, providing descriptors related to phrasal features for the grammatical category and reference to AWL words in the lexical category may help instructors provide students with more specific information about how their language use is changing, particularly in settings where students write without using sources in response to general interest topics, as opposed to contexts where particular genres of disciplinary writing are taught. However, implementing these changes may be challenging as it could increase the teaching and assessment workload of instructors, many of whom have up to 200 students per semester at this university, and require a redesign of the rubrics along with professional development sessions to socialize instructors into its use.

The findings that AWL use and phrasal complexity were greater in the second-year texts provides some insight into how the English Department at this university might organize and sequence the writing component in the required EFL classes. As mentioned previously, currently the first two classes do not focus on writing while paragraph writing is taught in the third class. It may be helpful to introduce sentence-level writing in the first class, which might emphasize different sentence types and appropriate punctuation. The second class should introduce paragraph length texts focusing on general interest topics that do not require using information from sources. However, the topics should have some relationship to a variety of subject-matter content so that students may begin to understand and use AWL words. This would allow students to develop their skills in organizing information within paragraphs (such as topic sentences and cohesive devices) and continue to work on their language skills and sentence structure without simultaneously learning how to integrate source information. Finally, the third class could then introduce writing from sources so that students could acquire key academic writing skills such as quoting, summarizing, and paraphrasing, and develop their

phrasal complexity, which is as a key feature of academic texts.

Since this study was carried out with undergraduate students who are required to take four EFL classes in their first two years of university study only, it was not possible to adopt a longer time interval for identifying language development. Focusing on English major students would provide a four-year window for observing change over time, but these students often begin university studies at a higher proficiency level and frequently take different classes than non-English major students. To ensure that writing development occurs in the EFL classes required by all university students, it is important to also explore the developmental trajectories of lower proficiency students who may have less incentive to acquire written English skills. In such contexts, instructors should consider how writing serves the overall learning goals, particularly if writing is seen as a vehicle for English language development. Furthermore, unlike English majors, these students write disciplinary texts in their subject-matter classes in Thai, so it is not possible to explore their English writing development as they move from general EFL classes to writing in the disciplines. Greater insight into these students' performance over time also might be gained by examining those who choose to complete an English minor, as they typically take English courses throughout all four years of university study. Although English minor students in their initial studies would produce paragraph-length texts only, as was the case for students in the current study, they acquire the ability to write longer essays as they move through the minor.

The study has several limitations that may impact the generalizability of the findings. Due to the need to control both faculty of study and writing topic, we sampled only 170 texts from a larger collection of texts (De Vleeschauwer, 2020). Carrying out additional rounds of data collection would increase the likelihood of finding more matched pairs but integrating the research task into the existing class schedules is challenging due to time constraints and the need to follow the course outline. To control for potential genre effects, we used writing prompts that elicited only one text type, specifically opinion paragraphs. Although this was an appropriate choice for students at this university, future studies should explore the developmental patterns associated with other types of academic writing, both at the paragraph and essay level. Reflecting these students' overall proficiency level, differences in their academic profiles, and variation in the importance of English in their degree programs, the mean text length was relatively short. However, we would argue that it is important to explore writing development with lower proficiency writers who can only produce short samples to gain insight into a more complete trajectory of development. Since their texts were collected for program evaluation and research purposes, students may have lacked the incentives associated with higher-stakes writing, such as graded assignments and exams. Due to variation in the content of the exams used in the required classes (i.e., some exams do not have writing tasks), it was not possible to compare their exam performance. We examined language features that had been tested in prior studies to build on those findings but also considered their ecological validity for instructors. This resulted in a lower number of target features, so future research that aims to identify measures associated with specific constructs (like complexity or proficiency) would require the inclusion of more measures. Although the target features were relevant for the EFL classes at this university, instructors in other settings may prioritize different language features.

Our analysis focused narrowly on the language features of student writing and did not consider the process by which students came to use those form. For example, recent research has used keystroke logging to identify how students revise phrasal features of their texts (Bowen & Thomas, 2020). Future studies should also elicit writers' perceptions about their texts and how their language choices evolved through various stages in the writing process (e.g., planning, composing, revising). Finally, we compared students in different years of study as a marker of development. For cross-sectional studies it can be challenging to operationalize development in ways that are not tied to text features. For example, having written texts assessed using rubrics and creating groups based on text ratings conflates text quality and development. It is important to use criteria that have some logical relationship to development (e.g., year of study, program level, specific course, beginning/end of a course) but are not based on judgements of writing quality. Our future research aims to further clarify how the written texts produced by EFL undergraduate students in this context change over time, in terms of both language features and other aspects of writing, as they move through all four university years.

THE AUTHORS

Kim McDonough is a Professor and Canada Research Chair in the Department of Education at Concordia University. Her research interests include second language acquisition, task-based interaction, second language writing development, and the role of visual cues in communication breakdowns.
kim.mcdonough@concordia.ca

Jindarat De Vleeschauwer is an Assistant Professor in the English Department at Chiang Mai University, Thailand. Her current research interests involve L2 writing, critical reading, and learning strategies in the Thai EFL setting.
jindarat201cmu@gmail.com

REFERENCES

Abdel-Latif, M. (2015). Sources of L2 writing apprehension: A study of Egyptian university students. *Journal of Research in Reading, 38*(2), 194-212. https://doi.org/10.1111/j.1467-9817.2012.01549.x

Biber, D., Conrad, S., Reppen, R., Byrd, P., & Helt, M. (2002). Speaking and writing in the university: A multi-dimensional comparison. *TESOL Quarterly, 36*(1), 9–48. https://doi.org/10.2307/3588359

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5-35. https://doi.org/10.5054/tq.2011.244483

Bowen, N. E. J. A., & Thomas, N. (2020). Manipulating texture and cohesion in academic writing: A keystroke logging study. *Journal of Second Language Writing, 50*, 100773. https://doi.org/10.1016/j.jslw.2020.100773

Casal, J. E., & Lee, J. (2019). Syntactic complexity and writing quality in assessed first-year L2 writing. *Journal of Second Language Writing, 44*(1), 51-62. https://doi.org/10.1016/j.jslw.2019.03.005

Cheng, Y-S. (2004). A measure of second language writing anxiety: Scale development and preliminary validation. *Journal of Second Language Writing, 13*(4), 313–335. https://doi.org/10.1016/j.jslw.2004.07.001

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(1), 213–238. https://doi.org/10.2307/3587951

Coxhead, A. (2011). The academic word list 10 years on: Research and teaching implications. *TESOL Quarterly, 45*(2), 355–362. https://doi.org/10.5054/tq.2011.254528

Crossley, S. (2020). Linguistic features in writing quality and development: An overview. *Journal of Writing Research, 11*(3), 415-433. https://doi.org/10.17239/jowr-2020.11.03.01

De Vleeschauwer, J. (2020). *A study of the writing development, writing proficiency, anxiety, and self-efficacy of Thai EFL students*. Chiang Mai, Thailand: Chiang Mai University.

Eldridge, J. (2008). "No, there isn't' an 'academic vocabulary,' but…": A reader responds to K. Hyland and P. Tse's "Is there an 'academic vocabulary'?". *TESOL Quarterly, 42*(1), 109-113. https://doi.org/10.1002/j.1545-7249.2008.tb00210.x

Lan, G., Liu, Q., & Staples, S. (2019). Grammatical complexity: 'What Does It Mean' and 'So What' for L2 writing classrooms? *Journal of Second Language Writing, 46*, 100673. https://doi.org/10.1016/j.jslw.2019.100673

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

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing, 29*(1), 3-15. https://doi.org/10.1016/j.jslw.2015.06.004

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878-912. https://doi.org/10.1111/lang.12079

Polio, C. (2017). Second language writing development: A research agenda. *Language Teaching Research, 50*(2), 261-275. https://doi.org/10.1017/S0261444817000015

Polio, C., & J. H. Park (2016). Language development in second language writing. In R. M. Manchón & P. Matsuda (Eds.), *Handbook of second and foreign language writing* (pp. 287-306). New York, NY: Routledge.

Thongyoi, K., & Poonpon, K. (2020). Phrasal complexity measures as predictors of EFL university students' English academic writing proficiency. *rEFLections, 27*(1), 44-61.

Yoon, H. (2017). Linguistic complexity in L2 writing revisited: Issues of topic, proficiency, and construct multidimensionality. *System, 66*, 130-141. https://doi.org/10.1016/j.system.2017.03.007

Yoon, H., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly, 51*(2), 275–301. https://doi.org/10.1002/tesq.296

Appendix: Writing Prompts

Changes in society

Traditional values
A person's worth nowadays seems to be judged according to social status, salary, and material possessions. Traditional values, such as honour, honesty, kindness, and trust, no longer seem important in contemporary society. To what extent do you agree or disagree with this opinion?

Family relationships
Modern lifestyles mean that many parents have little time for their children. Many children suffer because they do not get as much attention from their parents as children did in the past. To what extent do you agree or disagree with this opinion?

Social media
Social media is becoming increasingly popular amongst all age groups, especially the young. However, sharing personal information through social media can create too many problems that put young people at risk. To what extent do you agree or disagree with this opinion?

Addressing social issues

Pay equity
Some people feel that workers like nurses and teachers should be paid more, especially when film actors, company bosses, and athletes are paid huge sums of money for less important work. To what extent do you agree or disagree with this opinion?

Gender equity
Many high-level positions in companies are filled by men even if the workforce is more than 50 per cent female. Companies should be required to allocate a certain percentage of these positions to women. To what extent do you agree or disagree with this opinion?

Climate change
Some people think that individuals can help prevent global climate change by recycling, conserving water, and using public transportation. However, others insist that only governments and large businesses can make real changes. To what extent do you agree or disagree with this opinion?