Abstract
Syntactic complexity is a crucial aspect of linguistic proficiency and thus understanding and supporting such development in learners is a keen concern for language teachers. Research conducted has shown growing sophistication of noun phrase structures by writers of different abilities in academic writing (Biber & Gray, 2010; Liu & Li, 2016; Parkinson & Musgrave, 2014). In comparison, relatively less is known about the development of syntactic sophistication among student technical writers. Adopting a discipline-specific approach, this study compared the complexity of noun phrase structures in student texts of three performance levels from high, mid and low. The data taken from 45 technical reports of university engineering students were analyzed quantitatively using automatic syntactic analyzers and qualitatively by manual coding and text examination. The results show that noun phrase complexity is a differentiating factor for different performance levels with the stronger texts exemplifying more varied modifying structures and in greater numbers. Complex and lengthy structures, such as multiple use of prepositional phrases and combined use of other modifiers such as -ed or -ing clauses, are common in postmodification, whereas premodifying structures are simpler in structure and shorter. The study also explored the contribution of complex noun phrases to expression of meaning, showing that complex syntactic structures are commonly used to perform a diverse range of language functions essential to technical communication, such as explaining a scientific mechanism. The close connection between syntactic complexity and expression of meaning suggests that structures for pre- and postmodification should be learned as a meaning-making resource.

Keywords: syntactic complexity, noun phrase complexity, postmodification, technical writing, corpus analysis
Introduction

Syntactic development is an important indicator of writing ability. Syntactic range and sophistication typically feature in English proficiency assessment such as IELTS and TOEFL. It is expected that learner writers would increase in syntactic sophistication as their language proficiency improves, allowing them to express ideas in greater sophistication and compactness. Understanding and supporting the development of syntactic complexity in learners is therefore a keen concern for language teachers. One linguistic structure that has received much attention in recent years has been the noun phrase. In their large-scale corpus-based studies, Biber and Gray (2010) provide convincing evidence that complex nominals, rather than clausal subordination as in some popular beliefs, are a defining characteristic of academic writing. In another paper, Biber et al. (2011, p. 30) hypothesize that the development of noun phrase complexity might follow a process with the use of “extensive phrasal embedding in the noun phrase” representing the most advanced stage. Many other scholars have contributed to understanding of noun phrase complexity development among learners of different levels as well as with different registers including research papers, argumentative essays, TOEFL essays, etc. (Liu & Li, 2016; Parkinson & Musgrave, 2014; Shadloo et al., 2019; Thongyoi & Poonpon, 2020).

While there is increasing understanding of noun phrase complexity development in academic writing in general, relatively little is known about the topic against the realm of technical writing (Guillerit, 2020). Given that technical writing aims to convey very specific information and is also associated with distinct registers (such as technical reports), it may pose different requirements for syntactic complexity compared with general academic writing. Technical communication courses are a staple component for engineering curriculums. The purpose of this study is to investigate the use of complex noun phrases, especially the combined use of different pre- and postmodifiers, by students of three performance levels, high, mid and low. It is hoped that this understanding would shed light on the development of noun phrase complexity among writers of different abilities and contribute to better support for their linguistic development.

Review of Literature

The noun phrase has been the focus of attention regarding the study of linguistic complexity in academic writing. Using large corpora of professional academic writing and natural conversation, Biber and Gray (2010) provided strong evidence that academic writing and conversation have dramatically different linguistic characteristics. They assert that “academic writing is not structurally ‘elaborated’ . . . [but] structurally ‘compressed’, with phrasal modifiers embedded in noun phrase” (p. 2). Biber et al. (2011) further investigated the occurrence of 28 grammatical features associated with structural complexity in their corpora, arriving at a hypothesis for the developmental stages for complexity features starting with finite complement clauses for conversational competence to non-finite clauses and extensive phrasal embedding in noun phrases for competence in academic writing in the final stage. Such extensive and thorough studies have attracted much subsequent interest in features contributing
to complexity in academic writing. For example, Kyle and Crossley (2018) studied 480 argumentative essays written as part of the TOEFL examination. Their Tool for the Automatic Assessment of Syntactic Sophistication and Complexity (TAASSC) returned 66 fine-grained indices for various grammatical structures. They concluded that phrasal complexity are better predictors of writing quality than clausal indices, thus lending support to Biber et al.’s (2011) assertion about phrasal complexity and academic writing proficiency.

Other scholars have likewise contributed to the growing body of literature on complex noun phrases in different kinds of writing as well as among a variety of learner groups. In their 2014 study, Parkinson and Musgrave give support to Biber et al.’s hypothesis. Comparing the writing of less proficient pre-graduate students to writing by those who had been enrolled in graduate programs, they confirmed Biber et al.’s proposed developmental model in the sense that the less proficient group relied heavily on attributive adjectives, a modifier hypothesized as being acquired early. In addition, use of noun modifiers by the more proficient group was much closer to published frequencies for academic prose than was used by the less proficient group. Casal and Lee (2019) explored the relationship between syntactic complexity and writing quality in assessed source-based research papers produced by ESL undergraduate writers in a first-year writing course in three grade tiers: high, mid and low. They found that low level texts have significantly lower complex nominal densities whereas the highest densities of complex nominal types are present in high-rated papers. Liu and Li (2016) compared master theses of Chinese students in applied linguistics in the People’s Republic of China and research papers written by subject experts, making the observation that, even at the more advanced postgraduate level, learner texts are less complex in terms of mean length of clauses, frequency of complex nominals per T-unit and frequency of complex nominals per clause. They further identified features of one aspect of complexity, noun phrase modification, and found significant underdevelopment of postmodification in student writing relative to published texts.

While research is being built up to illuminate on syntactic development, especially in terms of noun phrase complexity, among learners, it cannot be taken for granted that similar developmental stages hold true across different registers or disciplines. For example, Shadloo et al. (2019) found that the pattern of noun phrase complexity predicted by Biber et al. (2011) was not observed across their argumentative essay corpora at different performance levels, prompting them to query if such discrepancy might stem from register differences between argumentative essays and academic subgenres such as the research article. Currently, studies that specify different academic disciplines remain relatively few. In this connection, Conrad (2017)’s work on the writing of engineers and student writers in civil engineering, which shows much difference in sentence structures between student and practitioner writing, provides highly needed insight into the needs of technical disciplines and the performance of students aspiring to become professional engineers. A study growing out of Conrad’s research is that by Guillerit (2020), who specifically studied noun phrases in civil engineering writing. This line of research, however, has tended to concern engineering students native in English. Much
work is needed to understand how second language engineering students develop sophistication in syntactic complexity in their writing.

The Study

This study responds to the need to investigate the role of syntactic complexity in student technical writing and to explore the relationship between syntactic complexity and holistic writing quality assessed by teachers. Three research questions guided the investigation:

RQ1. Are the linguistic structures selected for this study useful indicators of writing performance? RQ2. How do technical student writers use complex NP in terms of syntactic complexity? RQ3. How do technical student writers use complex NP in terms of expression of meaning?

Corpus Description

The data for this study come from an archive corpus of proposal report assignments written by second-year engineering majors taking a technical communication course at a Hong Kong-based university. This assignment required students to propose a technical innovation, review similar designs in the literature, produce a technical description of the innovation, and evaluate its feasibility and benefits. Forty-five scripts were randomly selected from the corpus as representative samples from three grade bands: high, mid, and low (i.e., assignments scored at the 75th, 50th and 25th percentile). The scripts were collected from multiple classes and were graded by the teachers of the students using a standardized rubric for the assignments. Each script was about 1,000–1,500 words, giving a total of 56,600 words. Table 1 presents a summary of the corpus compiled for this study.

<table>
<thead>
<tr>
<th>Grade Band</th>
<th>No. of Scripts</th>
<th>Word Total</th>
<th>Word Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>15</td>
<td>22,315</td>
<td>1,487.7</td>
</tr>
<tr>
<td>Mid</td>
<td>15</td>
<td>17,841</td>
<td>1,189.4</td>
</tr>
<tr>
<td>Low</td>
<td>15</td>
<td>16,344</td>
<td>1,089.6</td>
</tr>
<tr>
<td>Total</td>
<td>45</td>
<td>56,500</td>
<td>1,255.6</td>
</tr>
</tbody>
</table>

The scripts were manually cleaned to remove irrelevant information such as names, script titles, footers and the bibliography. Then the corpus was tagged using the CLAWS part-of-speech tagger developed by UCREL at Lancaster University (Garside, 1987).
Data Analysis

The analysis covered a total of six syntactic complexity features. The first three are global measures of syntactic complexity computed by Lu’s (2010) Syntactic Complexity Analyzer (L2SCA):

- mean length of clause (MLC),
- complex nominals per T-unit (CN/T), and
- complex nominals per clause (CN/C).

A one-way ANOVA was then conducted on these indices to compare the effect of syntactic complexity on grades. The other three syntactic features are more fine-grained, referring to the following structures when they function as postmodifiers of nouns:

- the nonfinite -ing clause,
- the nonfinite -ed clause, and
- the postmodifier complex (i.e., a structure consisting of two or more postmodifiers; see Biber et al., 1999, p. 641).

Instances of these three features were identified from the corpus with AntConc, a freely available concordancing program (Anthony, 2020). The pre- and postmodifiers identified by the concordancer were then manually coded by the three teacher-researchers for this study to ensure a high degree of accuracy. Any discrepancies of analysis were resolved by discussion and consensus.

Findings and Discussion

Evaluating syntactic complexity by automatic tools.

Automatic tools were used in the first stage of the investigation to provide an overview of the holistic characteristics of the three levels of writing featured in this study. The results returned by L2SCA are presented in Table 2, which shows a notable correlation between the three measures across grade bands. The H scripts have the longest mean length of clause, the most complex nominals by T-unit as well as by clause, followed by the M scripts and then the L scripts. The differences between the levels are statistically significant.

<table>
<thead>
<tr>
<th>Grade Band</th>
<th>High (H)</th>
<th>Mid (M)</th>
<th>Low (L)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean length / clause (MLC)</td>
<td>14.88</td>
<td>13.38</td>
<td>12.06</td>
<td>0.0011</td>
</tr>
<tr>
<td>complex nominals / T-unit (CN/T)</td>
<td>3.09</td>
<td>2.75</td>
<td>1.82</td>
<td>0.0151</td>
</tr>
<tr>
<td>complex nominals / clause (CN/C)</td>
<td>2.05</td>
<td>1.82</td>
<td>1.54</td>
<td>0.0069</td>
</tr>
</tbody>
</table>

sig. level = 0.05

Note. The three measures refer to the following respectively: mean number of words per clause, number of complex nominals per T-unit, and number of complex nominals per clause.
Apart from L2SCA, the concordancer AntConc was used to identify the target structures in the data. Figure 1 shows the frequency of -ed and -ing clausal modifiers found in the sub-corpora, where the x-axis shows the number of occurrences of the target structure in each level. Again, a clear and distinct pattern emerges. A total of 82 -ed clauses are found in the H texts, giving on average 5 to 6 instances in each text (normalized for 1,000 words) for 15 texts in all. Use of the structure is progressively fewer in the M texts and the fewest in the L texts. This overall trend is also found with the -ing clauses, even though on the whole, they are less frequently used in the postmodifier position compared to the -ed clauses.

**Figure 1**

*Frequency of -ed and -ing Clausal Modifiers in Three Levels of Student Technical Writing*

![Bar chart showing frequency of -ed and -ing clauses in different levels of writing]

*Note.* Frequency counts are normalized for 1,000 words.

**Pre- and postmodification in complex noun phrases.**

With some evidence for the differences between the three levels of texts, manual coding was carried out to probe into the combined use of various pre- and postmodifying structures in the entire noun phrase. Four structures can serve as premodifiers in a noun phrase and six structures as postmodifiers. A notable feature of the postmodifiers is that four of these are clausal, being relative clauses, -ed clauses, -ing clauses and to clauses. The noun phrase structure in Figure 2 posited in Parkinson & Musgrave (2014) was employed in this study.

Figure 3 shows the frequency of various modifiers used with -ed modified noun phrases in the sub-corpora. In line with the literature (e.g. Quirk et al., 1985), prepositional phrases are by far the most common modifiers in the -ed modified noun phrases that we identified. The number of prepositional phrases is higher than the -ed clauses both for the H texts and the M texts, showing that some of the
noun phrases are modified by more than one prepositional phrase. Other postmodifying structures are used but with lower frequency.

Figure 2
Structure of the Noun Phrase

Figure 3
Frequency of Types of Modifiers Used with -ed Noun Phrases in Three Levels of Student Technical Writing

Note. att adj = attributive adjective; part adj = participial adjective; poss = possessive adjective; prep phr = prepositional phrase; rel cla = relative clause; ed cla = -ed clause; ing cla = -ing clause; app phr = appositive phrase; to cla = to clause
Here is a sentence taken from an H script that uses 3 prepositional phrases (underlined) in a row to modify the head noun (bolded):

**Example 1:**

*TL patients have an opening called stoma on the lower part of the neck for breathing purpose.*  
(H5)  
\[ed\text{ cla} \ + \ \text{[prep phr]} x 3\]

Figure 3 also shows that premodifiers are rather commonly used. Most of these are attributive adjectives but there are also participial adjectives and nouns used as adjectives. Combined use of premodifiers is common, but in our data, combined use of premodifiers mostly involves two modifiers but rarely three. Here is an example showing two different premodifiers (double-underlined) used together:

**Example 2:**

... *the application will receive a digitalized sound signal recorded by ...*  
(H5)  
\[\text{[part adj]} + \text{[noun]}\]

This preference for no more than two premodifiers by the student writers could be seen as a restriction in the use of premodifiers. Three or more attributes stacked up before a head noun could pose a burden to reader comprehension while the lack of structural variety in the premodifying position may be a limitation. In contrast with the lack of structural variety in the premodifying position, postmodification is more flexible syntactically. In our corpus, flexible combination of different postmodifying structures can be found in all three levels of texts, with greater sophistication in terms of length and variety found in the H scripts. Here is an example where three different clausal postmodifiers are used together with a prepositional phrase in a complex noun phrase:

**Example 3:**

*A method finding the time required to reach the head of the queue is ...*  
(M8)  
\[\text{[ing cla]} \ + \ \text{[ed cla]} \ + \ \text{[to cla]} \ + \ \text{[prep phr]}\]

Figure 4 presents the frequency of various modifiers used with *-ing* modified noun phrases in the sub-corpora. As discussed before, *-ing* clauses are fewer in number compared to *-ed* clauses but the general pattern is similar: advanced writers use more modifying structures than writers of other levels. In a similar way, prepositional phrases are the most common structures with *-ing* clauses in noun phrases. The relatively frequent use of *to* clauses can also be noted here. In our data, *to* clauses are fairly commonly used with *-ing* clauses as postmodifiers. This is understandable given that the *-ing* clauses often carry an active meaning whereas *to* clauses are typically associated with the purpose of an action.
These two modifiers are used together because of their close association in meaning, such as in the following:

Example 4:
... a type of **photosensor** converting photons into current to sense the amount of ambient light present (H9)

![Figure 4](image)

**Figure 4**

*Frequency of Types of Modifiers Used with -ing Noun Phrases in Three Levels of Student Technical Writing*

On the whole, -ing modified noun phrases are simpler in structure than -ed modified noun phrases. Use of more than two postmodifiers is not common. The following is a typical example showing the use of an -ing modifier together with a prepositional phrase.

Example 5:
... provides them with a **number** representing their place at the queue (M8)

![Figure 4](image)

There is yet another observation that can be made of the -ing clausal postmodifiers. In our corpora, H writers do not use any relative clauses whereas a very small number in the other two groups of writers do. This poses a point of interest given that the advanced writers use other modifying
structures more often. A similar pattern can also be seen with the -ed clauses in Figure 3 where lower ability writers use more relative clauses in the full form with the relative pronoun which. While our corpus is small and the number of instances involved is limited, it would be interesting to further investigate if advanced writers tend to prefer more compact structures and therefore might be inclined to use relative clauses in the reduced form without “which”.

The foregoing analysis revolves around use of syntactic complexity in three levels of student writing. In line with the statistical measures presented in Table 2, there is increasing sophistication in the noun phrases from the lowest to the highest writing levels in terms of the number of different constituent modifiers and their combined use. The following three examples from the sub-corpora illustrate these differences:

Grade band L

In order to reduce the death caused by stroke, many technical innovations were invented to detect stroke more accuracy with a simple design in recent years. (L1)

Grade band M

Once the drones tray sensor detected all required items are ready, it follows the designated route calculated by the cloud server at an altitude of 100m above ground. (M3)

Grade band H

Inspired by the GPS SmartSole project, the Semi-self-powered GPS Tracking Shoe is a smart shoe designed to provide the instant location of people with Alzheimers to their caregivers over a significantly longer period. (H4)

Noun phrase complexity and expression of meaning.

Our data show that the complex noun phrase, with all the possibility made available by different types of modifiers, offers a potential structure for various meaning units to be attached to the head noun or to sub-heads. For example, the -ing clause often carries an active meaning whereas the -ed clause carries a passive meaning thus adding attributes to the noun that precedes it. By building up various structures, both before the head noun as premodifiers or after it as postmodifiers, a writer can specify all the meaning units intended in a flexible way. In the following example, the head noun “emitter” is the topic in focus in the proposal report. This head noun is premodified but also most remarkably postmodified by a long string of postmodifiers, consisting of five clausal structures including two -ed clauses, two -ing clauses and a to clause, with some of these further modified by prepositional phrases. Together, these modifiers inform the reader about the mechanism, the purpose, the operation, together with the action associated with the head noun. While premodifiers are usually restricted to a few
adjectives, post modification can be infinite in length theoretically and used either linearly or with embedding.

Example 6:
The proposed innovation, Ultrabreak, is a **toothbrush-sized ultrasound emitter** attached to an **oven-sized power converter and control panel loaded onto a medical trolley**, **providing therapy for lipoma utilizing the principle of ultrasonic cavitation to decompose lipoma cells.** (H11)

It has been mentioned that prepositional phrases are the most productive structure for noun phrase postmodification. Prepositional structures are simple to use and powerful in meaning expression. It is also a structure that is acquired early in Biber et al.’s (2011) model, employed readily by all three levels of writers. Writers who are in command of correct preposition usage have a powerful tool to convey different meaning units with relative ease. In the example below, four different prepositions, namely by, at, of and above, are used in a row to bring out information about agent, location, distance and position respectively.

Example 7:
*Once the drones tray sensor detected all required items are ready, it follows the designated route calculated by the cloud server at an altitude of 100m above ground.* (M7)

Another notable feature in complex noun phrases is the embedding structure, which enables information to be given retrospectively about the noun that precedes it. In the following example, the head noun of the sentence is the (traffic control) “system”. At this point readers are told that this system is centralized and co-ordinated, and it is only after this that readers know there is a computer system that is responsible for the action of centralization and co-ordination, and then that this computer system is located at Transport Department control centers. Embedded postmodifiers allow information to be given only after a sub-head is identified. The effect is the readers are not inundated with information but fed the information only when they are ready for it.

Example 8:
The **traffic control system in Hong Kong, which is called Area Traffic Control System, is basically a traffic control system centralized and coordinated by a computer system located at Transport Department control centers, on an area-wide basis.** (M10)

**Conclusion**

This study has examined the relationship between syntactic complexity and quality of assessed students’ technical writing in an attempt to answer the following questions:
RQ1. Are the linguistic structures selected for this study useful indicators of writing performance?

RQ2. How do technical student writers use complex NP in terms of syntactic complexity?

RQ3. How do technical student writers use complex NP in terms of expression of meaning?

Our quantitative evidence shows that, for RQ1, the holistic measures of MLC, CN/C and CN/T are good indicators of linguistic writing performance. Specifically, our ANOVA results point to the conclusion that the three measures affect assignment grades. Advanced writers tend to write longer clauses and use a higher number of complex nominals.

For RQ2, results of our quantitative and qualitative analysis reveal differences in the student writers’ use of pre- and postmodifiers. A major finding is that the students flexibly combined different postmodifying structures (see examples 1 and 3 to 8) in their technical reports, such as by using a nonfinite -ed clause together with one or more prepositional phrases to modify a head noun. In contrast, complex premodifiers are relatively less common (see example 2). Earlier on, we have referred to studies investigating noun phrase complexity by focusing on discrete pre- and postmodifying structures (e.g., Parkinson & Musgrave, 2014; Shadloo et al., 2019; Thongyoi & Poonpon, 2020). Just like Liu & Li (2016), we consider it fruitful to study the whole noun phrase, especially the postmodifier complex. For this, we had to resort to manual coding of the whole noun phrase, but we also gained a new level of understanding for how students use modifiers in the entire noun phrase.

For RQ3, our examples demonstrate how complex syntactic structures are commonly used to perform a diverse range of language functions, e.g., explaining a scientific mechanism, stating the purpose of a technical solution, and describing the operation of an invention. Other functions include providing information about agent, location, distance, and so forth. This study focused on -ed and -ing clauses used as postmodifiers in noun phrases, with our data revealing that our learners, especially advanced ones, combined these non-finite clauses frequently with one or more prepositional phrases (example 1) and also fairly frequently with other postmodifying structures (examples 3, 4, 6 and 7). A command of flexible combination of different postmodifying structures is evident.

Our study has pedagogical implications for language teachers in higher education. First, complex syntactic features should not be introduced to students as form-focused entities but as meaning-making devices. In other words, students should not equate the use of syntactically complex structures with a demonstration of their linguistic skills. Rather, they should view and employ syntactic complexity as a meaning-making resource that allows them, for example, to achieve a compact writing style. Second, when students attempt to incorporate syntactically complex features in their language, they are likely to make mistakes, as was revealed in the manual coding stage of our research. Instead of becoming apprehensive about student mistakes, teachers should emphasize that taking risks in language use is one of the ways that students learn and build their confidence. Teachers should also plan appropriate instructional activities to help advanced students with difficult-to-master syntactic structures, e.g., the correct prepositions to use in long prepositional phrases and appropriate use of defining and non-defining relative clauses.
A possible limitation of this study is the relatively small sample size. It is important to point out that we do not intend to make generalizable claims; rather, we have attempted to provide a detailed account of our research process so that the readers, armed with this knowledge, can make connections between elements of our study and their own experience (Barnes et al., 2005). As our research environment and our students are not atypical, it can be expected that aspects of our study are transferable to other contexts. Future studies can be conducted in different contexts with larger groups of students from either engineering or other disciplines.

References
Anthony, L. (2020). AntConc (4.0.3) [Computer software]. Waseda University. https://www.laurenceanthony.net/software


**About the Authors**

**Joyce Lee** is a lecturer in the Center for Language Education, the Hong Kong University of Science and Technology. She has been involved in the writing of ESP courses for several engineering disciplines. She teaches at both the undergraduate and postgraduate levels, and her research interests include curriculum development and technical communication.

**Kam Yin Wu** works in the Center for Language Education, the Hong Kong University of Science and Technology. He teaches EAP and ESP courses to undergraduate and postgraduate students, and lectures on the MA in International Language Education programme. His interests include academic writing and genre analysis.

**Eric Lee** works in the Center for Language Education, the Hong Kong University of Science and Technology. He teaches EAP and ESP courses to undergraduate students. His academic interests include course design and materials writing.