

Analyzing the patterns of lexico-grammatical complexity across Graded Readers levels

Larissa Goulart
Northern Arizona University
United States

Abstract

This study investigates the development of lexico-grammatical complexity in two levels of graded readers, beginner and intermediate. Using the linguistic features described in Biber et al. (2011), 58 graded readers were analyzed. Biber et al. (2011) proposed a lexico-grammatical developmental sequence based on empirical research. In this study, 11 of these grammatical features were counted in these two levels of graded readers. The graded reader corpus was tagged for the lexico-grammatical features, and a Mann-Whitney U test was conducted to account for the differences between graded readers' levels. The results of the statistical analysis suggest that beginner graded readers have more advanced features of complexity than intermediate graded readers. Nevertheless, after close analysis of text excerpts, it is clear that this is a result of the register being investigated, with intermediate graded readers having more features of conversation than beginner graded readers.

Keywords: graded reader, lexico-grammatical complexity, corpus linguistics, phrasal complexity, clausal complexity

Several researchers (Allan, 2008; Hill, 2008; Mohd-Asraf & Ahmad, 2003; Nakanishi, 2015) have addressed the role of Graded Readers (GR) in language learning by discussing the effects of extensive reading in learners' development. The use of GRs is also a common topic in teacher training books. One example is Grabe and Stoller (2011) who argued for the practice of "sustained silent reading of level-appropriate texts" (p. 198) in reading classes, which they argue to be "the single best overall activity that students can engage in to improve their reading abilities" (p. 198). Even though GRs are not the only written input learners will encounter, they tend to be the most extensive input learners have available. Hill (1997) also discussed the advantages of using GRs in language classes as they: a) promote language learning; b) provide motivation and practice; and c) promote literature studies. Allan (2008) suggested that "GR are a useful way of motivating learners to read extensively" (p.23) because they have interesting and complex stories. Taking such recommendations into account, many Intensive English programs (IEPs) incorporate the reading of GRs in their syllabus, making it common for learners to encounter these texts as part of their learning process.

Research on GRs has addressed the use of GRs in the language classroom and their impact on learners' skills development. One such case is Hafiz and Tudor (1989, 1990). In their first study they investigated whether extensive reading could improve students' language skills in an English as a Second Language (ESL) context. They used GRs with one experimental group and had two control groups with students who were not exposed to GRs. Their results show that students in the experimental group improved in their overall language skill, with better results in accuracy of expression. Nevertheless, they noticed that there was little change in the vocabulary used and a preference for simpler forms of syntax. Hafiz and Tudor (1990) conducted a similar study in an English as a Foreign Language (EFL) context. Learners in this context were exposed for longer periods of time to GRs; however, the results showed development in fluency and accuracy but not in the range of structures used.

The studies conducted by Hafiz and Tudor (1989, 1990¹) suggest that extensive reading with GRs can have a positive impact on learners' accuracy and fluency skills, but their study did not find the same results for "syntactic maturity", a term they use to refer to grammatical complexity. This lack of development in grammatical structures might be a reflection of GRs' texts. In other words, while GRs at more advanced levels use more complex vocabulary, it is unclear whether they use more complex grammatical structures. In order to make this claim, it is necessary to investigate GRs' texts to verify if there is development in the grammatical structures used.

Definitions of GRs usually mention vocabulary and grammatical simplification. For instance, David Hill (2008), the former director of the Edinburgh Project for Extensive Reading, defined GRs as "books written for learners of English using limited lexis and syntax, the former determined by frequency and usefulness and the latter by simplicity" (185). Nation and Ming-Tzu (1999) stated that GRs are simplified by "severely restricting the vocabulary that can occur, controlling the grammatical structures that occur and matching the length of text to the vocabulary and grammar controls" (p.363). Nevertheless, most of the research on GR texts to date has focused on lexical development (Nation & Ming-Tzu, 1999; Waring & Takaki, 2003; Wodinsky & Nation, 1988).

Therefore, the goal of this study is to investigate whether there is variation in the grammatical structures used across levels of GRs and if this variation indicates a progression toward greater complexity across levels.

Literature review

Lexico-grammatical complexity

Grammatical complexity, along with variables such as fluency and accuracy, has long been considered an indicator of second language proficiency (Kuiken et al. 2019). Recent studies on complexity development in learners of English as a second language include Lu's (2010, 2011)

¹ It is worth noting, however, that since the studies conducted by Hafiz and Tudor (1989 and 1990), graded readers have improved considerably. Therefore, if said study was replicated with GRs written in the past 5 years, results might be different.

investigations on automated analysis of complexity, and Kuiken and Vedder's (2019) analysis of complexity in learners with different first languages (L1). Research syntheses on grammatical complexity (Bulté & Housen, 2012; Norris & Ortega, 2009) suggest that complexity has been analyzed using measures of length such as T-units and AS-units, and measures of embedded clauses (subordination and coordination). Norris and Ortega (2009) argued that the elements which indicate complexity also vary according to learners' levels. As an example, they mention the fact that for real beginners, coordinated clauses could be a measure of complexity, while the same is not true for advanced levels.

Biber et al. (2011) challenged the use of T-units and embedded clauses as measures of complexity. These authors compared the use of phrasal and clausal grammatical features in conversation and academic writing, showing that embedded clauses and T-units characterize conversations, rather than writing. Their results revealed that writing is more compressed due to the use of phrase elaboration features, such as attributive adjectives, premodifying nouns, and postmodifying prepositional phrases, while conversation uses more clause elaboration, such as verb complement clauses and subordinate clauses. Based on these findings, the authors proposed a sequence for the development of complexity features. This sequence, presented below, suggests progression in two grammatical parameters (Staples et al., 2016, p. 5):

Grammatical form:

Finite dependent clause -> non-finite dependent clause -> dependent phrase.

Syntactic function:

Using clause constituents (e.g., direct and indirect objects, adverbials) -> using noun phrase modifiers.

This progression indicates that when learners start to write in English, they first use patterns of conversations, with finite dependent clauses, such as *the man who was sitting on the bench*, then, move to non-finite dependent clauses as *the man sitting on the bench*, and finally move the structure to the noun phrase, as in *the sitting man on the bench*.

Since Biber et al. (2011) suggested that complexity studies should investigate elements of the noun phrase, researchers in second language acquisition have added features of phrasal elaboration to their research (Bulté & Housen, 2018; Lu & Ai, 2015). These studies include measures such as complex nominals per clause or complex nominals per T-units. Corpus linguistics researchers have adopted an approach closer to that suggested by Biber et al. (2011) using as measures of complexity phrasal features (e.g. attributive adjectives, premodifying nouns, postmodifying prepositional phrases) and clausal features (e.g. verb complement clauses, finite adverbials, clausal coordinating conjunctions). These studies (e.g., Ansarifard et al., 2018; Biber et al., 2016; Taguchi et al., 2013) have tested Biber et al.'s (2011) developmental stages in L1 and second language (L2) writing, supporting the claim that written registers have more linguistic features of information compression in the noun phrase at more advanced levels. Nevertheless, none of these authors has explored complexity in texts students have to read, such as GRs. Therefore, this study seeks to fill this gap by analyzing the use of lexico-grammatical complexity features in two levels of graded readers.

Graded readers and vocabulary

Research on GRs and vocabulary has used corpus tools to investigate issues, such as: learners' vocabulary acquisition based on GRs' progressions (Nation & Ming-Tzu, 1999; Waring & Takaki, 2003; Wodinsky & Nation, 1988;), or comparisons between GRs and naturally occurring language (Allan, 2009; Claridge, 2005). However, to the best of my knowledge, corpus tools have not been used to investigate grammatical progression in GRs.

Wodinsky and Nation (1988) investigated word frequency in two graded readers comparing them to authentic texts. Their goal was to determine whether GRs helped in vocabulary acquisition. The results of their study suggest that reading only one GR in a given level does not offer enough exposure to new words to guarantee acquisition (see Uchihara et al., 2019). Nevertheless, when comparing GRs with naturally occurring texts they conclude that GRs still offer more opportunities for vocabulary acquisition than authentic texts. In a follow-up study of vocabulary in GRs, Nation and Ming-Tzu (1999) investigated the progression of vocabulary in a series of graded readers. These authors analyzed whether at least 95% of the words in a new level had already been presented to the learner in a previous level. Their results show that, in the collection analyzed, every new level recycles the vocabulary of the previous level and adds 5% of new headwords.

While Wodinsky and Nation (1988) compared GRs with naturally occurring texts, Claridge (2005) compared GRs with the original text of the same book. This researcher chose two graded readers to contrast the adapted text with the original passage. Unlike previous research on vocabulary in GRs, she used as a reference a vocabulary list based on leveled GRs. Her results show that the same patterns of vocabulary structures (i.e. collocations, high and low frequency words) as the original passage could be found in the adapted texts, which suggests that GRs are preparing learners to read authentic texts. Allan (2008) also compared GRs with naturally occurring language. Allan (2008) used the British National Corpus (BNC) to compare the lexical chunks of GRs with the ones in authentic texts. Her results show that there are more 4-word bundles in the BNC, and more 2 and 3 words bundles in the corpus of GRs. But, when comparing the types of bundles, she argues that while the corpus of GRs provides less exposure to lexical bundles, it gives learners exposure to all types of bundles that occur in the BNC.

The studies surveyed in this section show that there is progression in vocabulary across GRs levels. Nevertheless, these studies do not give insights into grammatical progression in GRs. Considering our lack of understanding in this area, the study described here seeks to answer two research questions:

- RQ1. Is there linguistic variation between graded readers for beginners and intermediate students?
- RQ2. If so, does it represent the stages of complexity development proposed by Biber et al. (2011)?

Method

In order to investigate the development of grammatical complexity across two levels of graded readers, a corpus of GRs was collected, and the lexico-grammatical features used in the study were identified with the Biber Tagger (Biber, 1988). In the following sections, the corpus, lexico-grammatical features and statistical test used in the analysis are described.

A Graded Readers corpus

Previous studies on GRs have used publishers' corpora of graded readers. For instance, Allan (2008) used Pearson's corpus of GRs, while Nation and Ming-Tzu (1999) used Oxford Bookworm's corpus of GRs. Nevertheless, for this paper, the goal was to use a corpus representative of different publishers, since these are the texts students will encounter in their reading classes. Therefore, a corpus containing 58 GRs from different publishers was compiled. The first step in the corpus compilation was to select the GRs. For this step, I have used the GRs students have access to at my institution's IEP library. Next, the texts were scanned, then converted from pdf to a word document. In this process 12 GRs did not convert to word document, hence they were excluded from this study. The next step was to clean the texts keeping only the text and deleting images, subtitles for the images, sentence fragments, and chapter headings. In some cases, due to fragments of texts not being converted the chapter was excluded from the analysis. After this process, the corpus was ready to be saved in txt files and divided in two subcorpora.

The final version of the corpus contains 58 GRs from eight different publishers (see Appendix II for a complete list of the texts included). Since the corpus contains texts from different publishers, a system to divide them into levels had to be created. The first issue encountered in dividing the corpus in levels was that the publishers do not have a consistent system of levels. For instance, one publisher divides its texts in nine levels, while another divides them in seven. Another problem was that the number of headwords did not match the same level across publishers. For instance, a GR with 800 words is a level 2 book for one publisher, but a level 4 book for another. Therefore, I have decided to use the Common European Framework of Reference for Languages (CEFR) level the publishers associate with the GRs in the corpus. That is, if a publisher stated that a level 4 book was equivalent to level B2 in CEFR, that book was added to the subcorpus of intermediate level texts. Lastly, after dividing the texts this way the corpus was divided into two levels: beginner (A1 and A2) and intermediate (B1 and B2).

Table 1

Description of the Graded Reader Corpus

Level	Number of texts	Types	Tokens	Mean text length
Beginner	36	3,591	123,265	3,166.92 (2,798)
Intermediate	22	8,114	298,244	12,824.5 (6,738)
Total	58	11,705	421,509	6,839.14 (6,624)

Lexico-grammatical features

To analyze the frequency of occurrence for the features investigated it was necessary to tag the texts using the Biber Tagger. This tagger was developed by Biber (1988) to analyze the features of spoken and written discourse. In its current version the tagger accounts for more than 150 linguistic features. According to Staples et al. (2016) the tagger has 95% accuracy for formal written texts and 90% for L2 writing. Previous studies that have used the tagger (Crawford et al. 2019; Staples & Reppen, 2016) suggest that the word *that* shows lower accuracy rates and, based on my own observation of the tagged texts, third person verbs are also misstaged as nouns in the corpus. Therefore, I have sampled the first hundred words of each text to check the use of the tag for (a) *that* as a verb complement clause and relative clause, and (b) all nouns.

The checked tags showed a precision rate above 90% (nouns - ^{^nn*} - 94%, verb complement clauses - ^{^tvt+vcmp} - 100%, relative clauses - ^{^tvt+rel} - 100%). It is worth noting that this is a small sample of the corpus, so the precision figures would probably be different if a larger portion of the corpus had been taken into account. After fixing the erroneous tags, the next step was to run Tag Count. This program gives the normalized count of several features per text; it also combines some of the tags given by the Biber Tagger, giving counts for a larger number of features.

The grammatical complexity analysis was conducted based on a range of grammatical features that have been discussed in previous empirical studies of complexity in writing. Special attention was given to features that were included in studies of language development (e.g. Parkinson & Musgrave, 2014; Staples et al. 2016; Staples & Reppen, 2016). These features include clausal structures (adverbials, verb complement clauses, and clausal coordination), and phrasal structures (premodifying nouns, attributive adjectives, and total nouns). Based on Staples et al. (2016), a set of 'intermediate' features was included. These features are finite clauses functioning as constituents of the noun phrase (*that* noun complement clause, relative clause and *wh*- relative clause). In order to avoid confusion with the intermediate level texts, these features will be referred as modifying clauses.

Statistical Analysis

To investigate the differences in grammatical complexity across GRs levels, means, medians and standard deviations were obtained for all features in both levels. In this study, the independent variable was GR level (beginner and intermediate) and the dependent variable was the features of complexity described on Table 2. Since the data did not meet the assumption of normality on Kolmogorov-Smirnov test (see appendix I), a non-parametric test was conducted to verify if the mean difference between both levels of GRs was statistically significant. Considering that the independent variable consists of two independent groups, a Mann-Whitney U test was conducted using SPSS.

Effect sizes were calculated using Cohen's *d* and interpreted using Plonsky and Oswald (2014) guidelines, which suggest that for independent samples in applied linguistics research, a Cohen's *d* between 0.4 and 0.7 is small, and a Cohen's *d* between 0.7 and 1.0 is medium, and a Cohen's *d* above 1.0 is large. Features with a positive Cohen's *d* value characterize higher frequency in beginner level GRs, while features with a negative Cohen's *d* value characterize higher

frequency in intermediate level GRs. This reflects the order in which the mean scores were entered into the Cohen's *d* calculation.

Table 2

Lexico-grammatical Features

	Linguistic Features	Example
Clausal Features	Finite Adverbial Clauses	(...) the community <u>because they felt neglected and unworthy of a nice, beautiful, functional environment</u>
	Verb + <i>that</i> clauses	mean <u>that 2.6 billion of people have gained access</u> (...)
	Clausal Coordinating Conjunctions	clinical presentation can be confused with infectious diseases or other cancers <u>and at first a simple blood examination is requested</u>
	Total Verbs	Run, walk, think, etc
Phrasal Features	Nouns	Time, system, water, etc
	Attributive Adjectives	the <u>main source of</u> microbial contamination
	Premodifying Nouns	systems have been seen as prominent <u>water source</u> .
Modifying clauses	Wh - Relative Clauses	The person <u>who is using the product</u> will have a good performance...
	That - Relative Clauses	the worker <u>that is sorting the distance to be covered</u> may damage it...

Results

This section presents the results of the Mann-Whitney U test for the phrasal, clausal, and clause modifying features in each GR level. Table 3 summarizes the descriptive statistics of mean, standard deviation, and median and presents the result of the Mann-Whitney U test.

Table 3*Mann-Whitney U test results for complexity features*

	<i>Beginner</i> <i>N = 36</i>		<i>Intermediate</i> <i>N = 22</i>		U	<i>p</i>	<i>d</i>
	Mean (SD)	Median	Mean (SD)	Median			
<i>Adverbial-causative</i>	0.71 (0.89)	0.40	1.35 (0.78)	1.30	195.5	<i>p</i> < .001	-0.76
<i>Adverbial - conditional</i>	0.31 (0.69)	0	1.92 (1.09)	1.85	64.5	<i>p</i> < .001	-1.80
<i>Adverbial – other</i>	1.29 (1.67)	0.55	4.06 (2.01)	3.70	122.0	<i>p</i> < .001	-1.50
<i>Verb + that clauses</i>	1.41 (1.56)	1.25	2.94 (1.57)	2.85	180.5	<i>p</i> < .001	-0.97
<i>Clausal coordination</i>	11.68 (8.18)	10.70	14.64 (5.28)	13.85	267.0	<i>P</i> = .039	-0.43
<i>Total verbs</i>	185.23 (20.94)	188.55	184.32 (19.09)	187.15	359.5	<i>p</i> = .559	0.04
<i>Total nouns</i>	273 (44.24)	277.70	236.54 (31)	231.25	195.5	<i>p</i> < .001	0.97
<i>Attributive adjectives</i>	27.31 (16.58)	24.25	27.1 (7.12)	25.30	336.0	<i>p</i> = .336	0.01
<i>Premodifying nouns</i>	12.65 (9.03)	9.40	8.94 (4.62)	8.15	306.0	<i>p</i> = .149	0.54
<i>Nominalization</i>	15.4 (12.89)	12.90	16.05 (9.63)	14.35	343.5	<i>p</i> = .400	-0.05
<i>Relative clauses (that)</i>	0.46 (0.65)	0.10	1.3 (0.95)	1.00	144.5	<i>p</i> < .001	-1.05
<i>Relative clauses (wh-)</i>	0.46 (0.88)	0	1.82 (0.83)	1.90	95.0	<i>p</i> < .001	-1.59

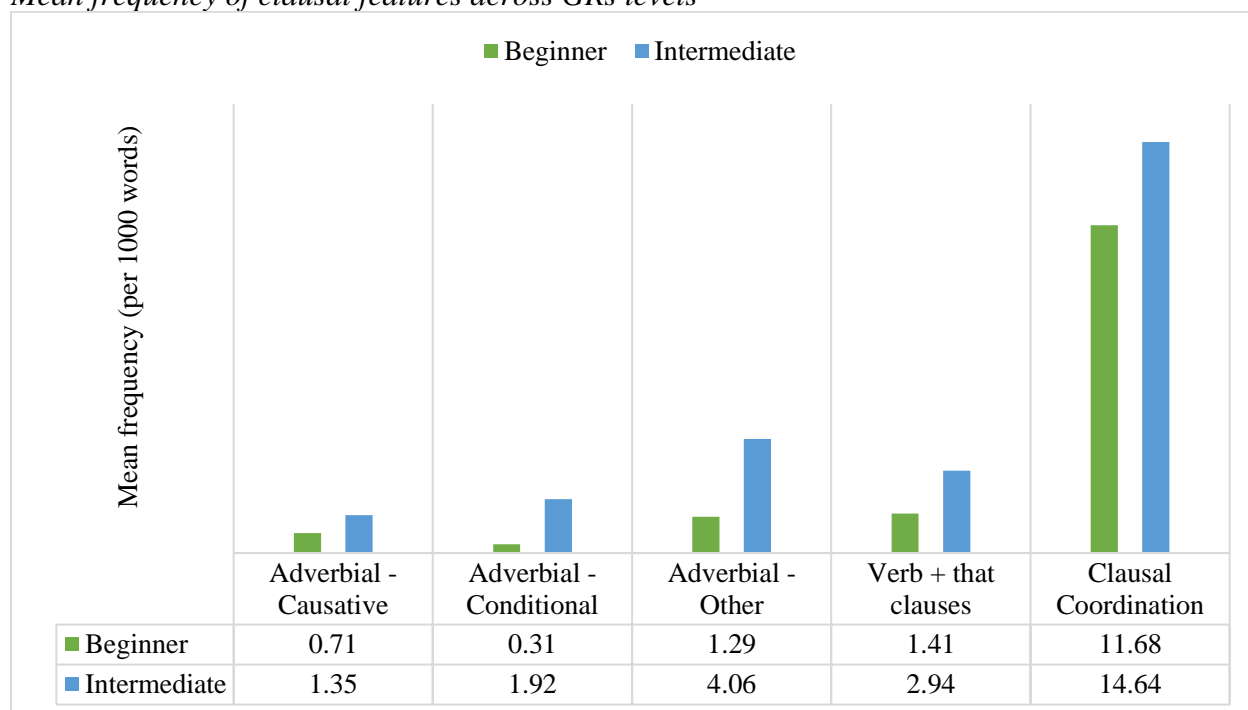
Table 3 shows that for four of the five clausal features investigated, there is variation between beginner and intermediate level. The Mann-Whitney U test indicates that the use of causative adverbials is greater for intermediate GRs (Mdn = 1.3) than for beginner GRs (Mdn = 0.4), $U=195.5$, $p<.001$, $d=-0.76$. The same can be observed for conditional adverbials, $U=64.5$, $p<.001$, $d= -1.80$, other adverbials $U=122$, $p<.001$, $d= -1.50543$, verb + *that* clauses, $U =180.5$, $p<.001$, $d=-0.97764$, and clausal coordinating conjunctions, $U= 267.0$, $p = .039$, $d=-0.43982$. Furthermore, aside from clausal coordinating conjunctions all features have an effect size considered medium or large, with the difference in level explaining a large portion of the difference between conditional adverbials and other adverbials. Figure 1 illustrates the mean differences in clausal features across GR level.

Figure 1 shows the mean frequency difference between GRs' levels for clausal features. We can see that intermediate GRs use more clausal features than beginner GRs, with the mean difference

being larger for conditional adverbials and other adverbials. This graph illustrates the differences in effect size shown in Table 3.

Figure 1

Mean frequency of clausal features across GRs levels



Taking into account the features which were included in its total counts, verbs and nouns, the Mann-Whitney U test shows that the use of verbs in beginner GRs (Mdn = 188.55) does not have a statistically significant difference to that of verbs in intermediate GRs (Mdn = 187.15), $U=359.5$, $p=.559$, $d=0.04$, while nouns in beginner GRs (Mdn = 277.7) are more frequent than in intermediate GRs (Mdn=231.25), $U=195.5$, $p<.001$, with a large effect size $d=0.97$. Figure 2 illustrates this difference in bar charts.

In Figure 2, we see that there is no significant difference in the frequency of verbs across GRs' levels, but that nouns are more frequent in beginner level GRs. None of the other phrasal features included in this study appear to be statistically significant ($p<.05$). Nevertheless, it is worth pointing out that premodifying nouns have a somewhat small effect size, with a Cohen's d of 0.54. Figure 3 illustrates the mean frequency of phrasal features across GRs levels.

Figure 2

Mean frequency of total verb and nouns

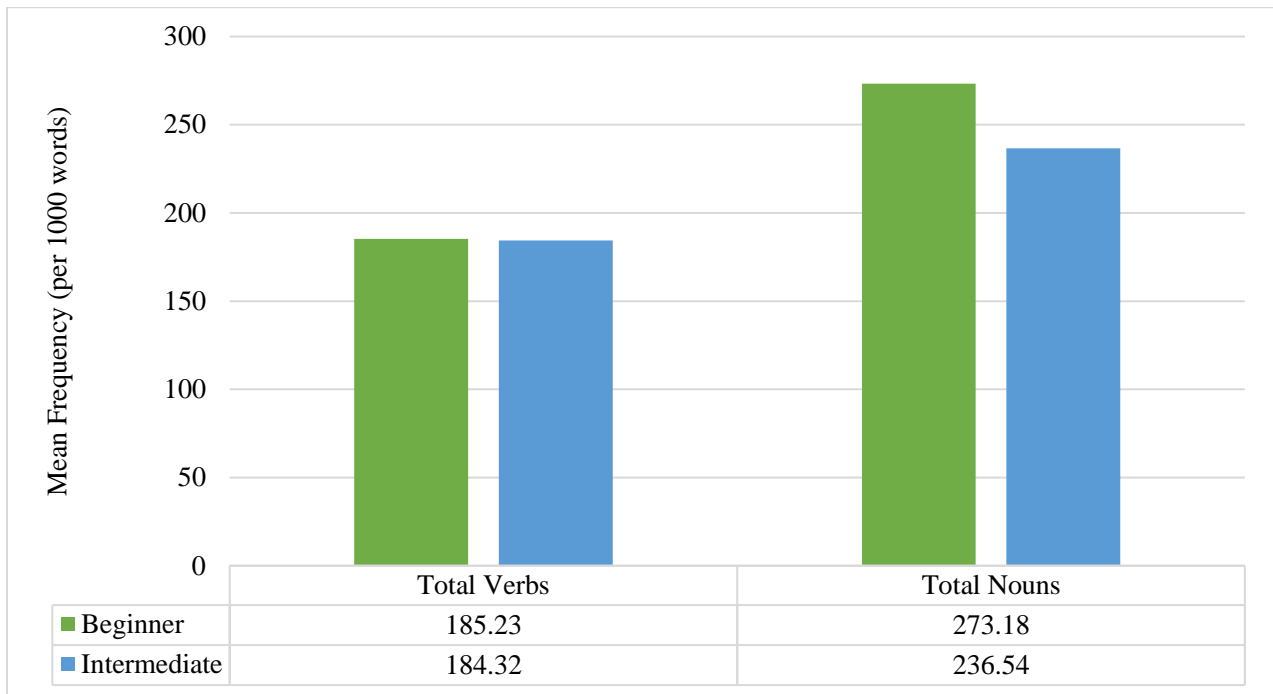
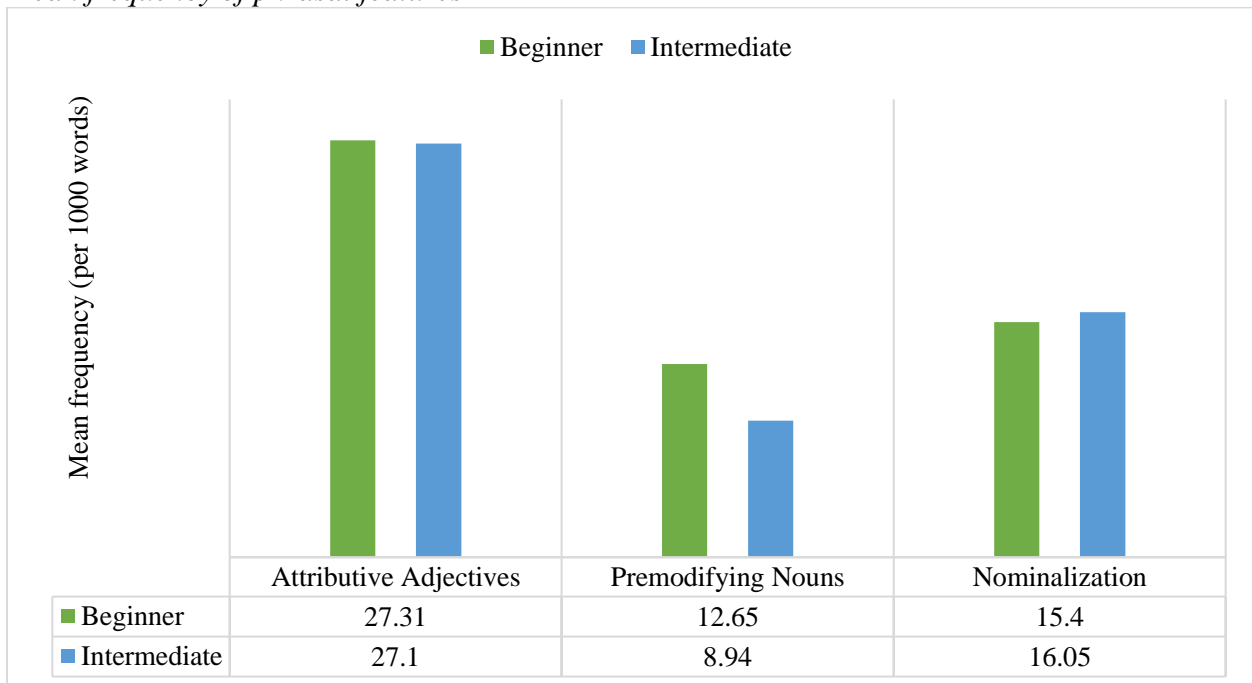


Figure 3

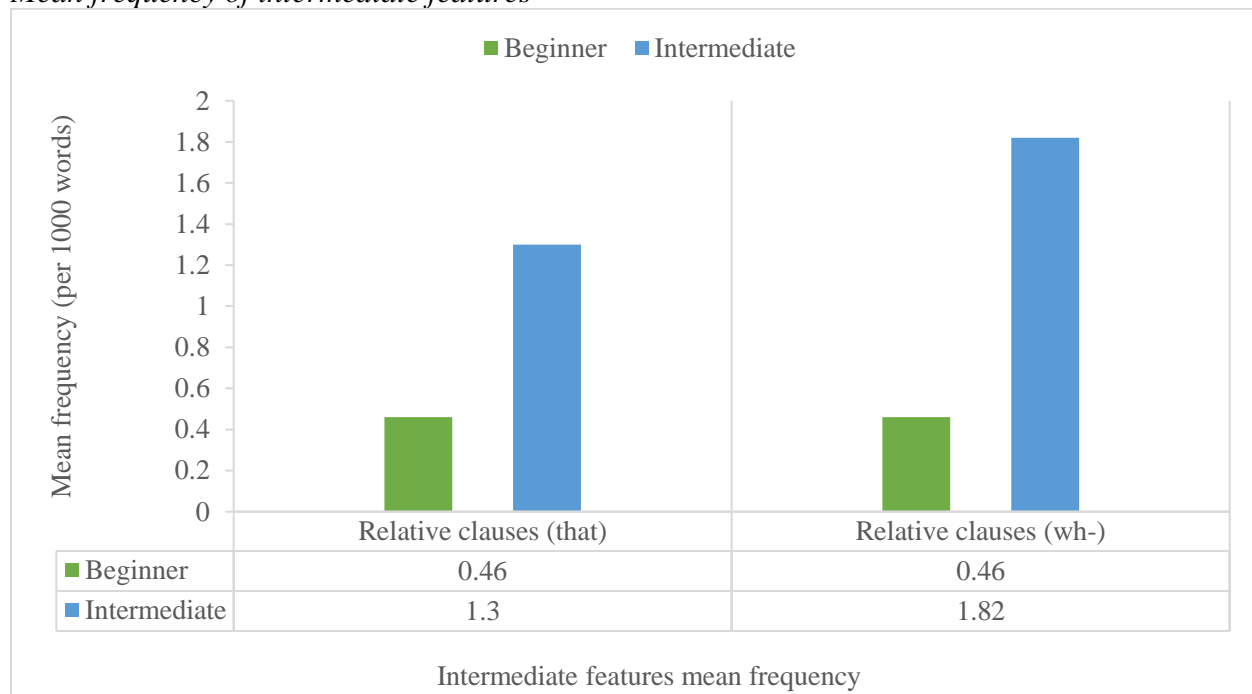
Mean frequency of phrasal features



Finally, taking into account modifying clauses, *wh*-relative clauses and *that*-relative clauses, the results of the Mann-Whitney U test indicate that in intermediate graded readers, relative clauses controlled by *that* (Mdn=1.0), and by *wh*- words (Mdn=1.9) are more frequent than in beginner GRs (Mdn=0.1 and Mdn=0.46) respectively, with the Mann-Whitney test resulting in $U=144.5$, $p<.001$, $d=-1.05$ for *that* relative clauses and $U=95$, $p<.001$, $d=-1.59$ for *wh*-relative clauses. Figure 4 shows the differences in mean frequency for both of these features in beginner and intermediate GRs.

Figure 4

Mean frequency of intermediate features



The results of the Mann-Whitney U test discussed in this section indicate that there is little to no statistically significant variation between GR levels for most of the phrasal complexity features investigated here, with the only exception being the total amount of nouns. It also shows that clausal and intermediate features are more frequent in intermediate level GRs than in beginner level readers. In the next section, these results will be discussed with examples of language use extracted from the GRs.

Discussion

The results of the statistical analysis show that beginner and intermediate graded readers do not follow the patterns of development suggested by Biber et al. (2011). Beginner and intermediate GRs have almost the same mean frequency for the use of phrasal complexity features. And, clausal and modifying clause features are more recurrent in intermediate levels GRs. In this section, text examples from both GRs levels are discussed to explain the patterns of language use found in the statistical test.

In excerpts 1 to 4, the complexity features analyzed are marked as follows: adverbials are marked between parenthesis (because he has enough...), verb complement clauses controlled by *that* are in between brackets (he believed [that the thief was in the house]), clausal coordination is between slashes (opened /and/ closed), verbs are marked in bold (**write, think**), nouns are underlined (desk, sentence), attributive adjectives, premodifying nouns and nominalizations are marked in italics (*detective* story), and relative clauses controlled by *that* or *wh-* words are between tildes (the person ~who was watching the dog~). The first two excerpts are examples of beginner GRs, and the following ones are examples of intermediate GRs.

Excerpt 1

Simon Decker **sat** at his desk, **thinking** of what to **write**. He **wanted** to **write** a detective story, /but/ he could only **think** of the first sentence:
 The *wooden stairs* **made** a *strange noise* when he **walked** on them; that's how he **found** their *secret hiding place*.
Simon **was** only thirty when he **left** the army, /but/ he **didn't like** his job. He **was** a *bomb expert*. He **worked** on a bomb which would **put** people to **sleep**, rather than **kill** them. He **thought** this would **be** the best way to **win** a war without **hurting** other people. Soon, he **disliked** the whole idea of war.
 He **loved reading** *detective stories* when he **was** a boy, and he **had** enough money to **live** without **working** for some time. So, he **decided** to **leave** the army and **try** to **become** a writer.
 His dog, Rocky, **was** at his feet. Suddenly the dog **lifted** his head /and/ **listened**, /then/ **ran** to the kitchen door.
 "What **is** it, Rocky?"
Simon **heard** a woman **scream**, /then/ a gunshot. There **was** another shot and the sound of glass **breaking**. Simon **fell** to the floor and **crawled** towards the kitchen. Rocky **was** in there **barking** and **jumping** at the kitchen door. Simon **held** Rocky in his arms to **keep** him quiet, then he **listened**. A car door **opened** /and/ **closed**, /then/ **drove** quickly away.

Excerpt 1 exemplifies the fact that in beginner GRs, subordinate clauses are almost non-existent, with coordinated clauses occurring more frequently than subordination. Although verbs are not one of the key features in beginner GRs, it is possible to notice that most of the verbs in this excerpt are past tense verbs and action verbs (break, jump, crawl). Both of these are characteristics of narrative texts, as discussed by Biber (1988), which is not surprising since these GRs are fictional books for language learners. Excerpt 2 is part of a different GR, and it illustrates how phrasal complexity features appear at this level of reading.

Excerpt 2

Café Colombo **is** always busy. There **are** always people **sitting** at the tables in Café Colombo. The men and women **are** all well-dressed. They **drink** coffee. They **eat** ice-cream. They **read** newspapers.

Charlie **does not sit** in Café Colombo. Charlie **is** not well-dressed. He **sits** on the pavement outside Café Colombo. He **sits** on a *small* box.

In both excerpts 1 and 2, we see the use of many nouns, attributive adjectives and premodifying nouns. Several of the nouns shown on these excerpts are proper nouns (Charlie, Simon Decker, Rocky). Analyzing the use of the features of phrasal complexity, we notice that they are not being used as features of phrasal compression, but rather as qualifiers and identifiers of the nouns (*small box*, *kitchen door*). Furthermore, excerpt 2 reveals the pattern of language use in beginner graded readers, that is, simple sentence structures composed of a subject, a verb and a complement (Charlie is not well-dressed, Charlie is always busy). In addition, beginner GRs rely heavily on illustrations to explain the story. At this level, it is common to find at least one illustration every two pages. Therefore, the images complement the description of the objects described in the text.

These excerpts indicate that Norris and Ortega's (2009) argument that measures of complexity have to be appropriate to learners' levels might extrapolate to reading texts as well. The Mann-Whitney U test showed that beginner GRs have the same amount of phrasal elaboration as intermediate GRs. Nevertheless, previous research in writing had shown that phrasal elaboration was a characteristic of upper-level writers. These excerpts suggest, however, that beginner graded readers have simple sentence structures, where nouns and attributive adjectives are more frequent. One point to be noticed is that in this analysis, the author did not account for prepositional postmodifying phrases, which might be a better indicator of phrasal complexity (see Bulté & Housen, 2018).

The following excerpts were extracted from two intermediate GRs in the corpus. These excerpts illustrate the common patterns of language use in the intermediate corpus.

Excerpt 3

'Now you're **asking** the *right questions*,' Benjamin **said**. '**Give** us all another drink, will you, Pete?'

Cameron **began mixing** the drinks. A *tall* and *handsome man* of thirty-five, he **moved** quickly /and/ **watched** the others (as he **worked**).

'All right, Doug,' Benjamin **said**. 'We're the *top men* in Granger Shoe. I'm *sales chief*, you're *production*, /and/ Rudy here **is** fashion and design. We're all on the board of directors, /and/ we all **know** what's wrong with the company.'

'What's that?' King **asked**.

'The Old Man.'

'What does he **know** about women's tastes? What does he **know** about women?' Stone **said**. '/But/ he's president of Granger. Year after year, he's president, (because he **has** enough stock **to keep** it that way).'

'/And/ the company **goes** down and down.'

'/And/ my stock **is** worth less and less each year,' Frank Blake **said**.

King **watched**, (as Benjamin **went** quickly to the glass table)
 /and/ **picked up** a red shoe. '**Look** at this!' he **said**. 'This **is**
 what I **mean** by excitement!'
 '**Made up** from my own design,' Rudy Stone **said** proudly.
 '**Take** a good look at it, Doug.'

Excerpt 3 represents a pattern found in most GRs at the intermediate level, the use of dialogues, with few if any excerpts of narrative texts. Considering this it is not surprising that clausal features are significantly more frequent in intermediate GRs. As previously mentioned, the development stages proposed by Biber et al. (2011) suggest that lower level learners use more features of clausal elaboration because these features are more common in conversation. Hence, if intermediate GRs are mainly composed of dialogues, which simulate conversations, it is expected that clausal elaboration will be more frequent at this level.

Excerpt 4

"**wake up!**" she **said**. "We`**re** in the new house, don't you **remember?**"

They **wanted** to **surprise** their mother /and/ **get** the breakfast ready, /but/ first they **went** to **look** outside. The house **seemed** to **stand** in a field near the top of a hill, /and/ they **could see** a *long way*.

"This place **is** much prettier than our house in London" **said** Phyllis.

In both excerpts 3 and 4, we can see that intermediate GRs also have a great amount of phrasal elaboration features, especially nouns. Similar to beginner GRs, these nouns refer to characters in the story (Cameron, Phyllis). Excerpt 4 also has narrative features such as past tense verbs. We can observe in excerpts 3 and 4 that the clausal features most frequent in intermediate GRs are clausal coordinating conjunctions, and in many cases in Excerpt 3 they are used as a way for new characters to join the conversation (And the company goes down and down). Considering intermediate GRs, we can notice that even though relative clauses have large effect size they are not that frequent. Their effect size can be explained by the fact that even though they are uncommon at the intermediate level, they are almost inexistent in beginner GRs.

In this section, the patterns of language use revealed in the Mann-Whitney U test were explored using examples from the text. These excerpts explain the recurrent use of clausal and modifying clause features in intermediate GRs, and they also illustrate the sentence structures encountered in both levels of GRs. In the next section, the research questions will be summarized as well as the limitations of this study.

Conclusion

Research question one derived from Hafiz and Tudor's (1989, 1990) studies indicating that learners did not make syntactic progress in their writing after being exposed to the GRs. Because previous research on GRs did not address the issue of linguistic variation, it was unclear whether GRs gave learners the opportunity to encounter new grammatical forms as levels progress.

Therefore, the first research question asked whether there was linguistic variation in the grammatical features investigated in this study. The results of the Mann-Whitney U test show that there is grammatical variation across levels of GRs. This shows that in addition to increasing vocabulary knowledge, different levels of GRs also introduce students to new grammatical structures.

The second research question emerged from studies of complexity development in learners' writing. Learners' writing research showed that novice writers adopt patterns of conversation when writing (clausal elaboration) and as they become more experienced move to patterns of upper-level writing (phrasal compression). Taking this into account the second research question sought to verify whether GRs also followed this pattern of development. As the Mann-Whitney U test results showed, the beginner GRs have slightly more features of phrasal compression, specifically premodifying nouns, while intermediate GRs have more frequent features of clausal elaboration. However, after examining the GR excerpts presented in the discussion, we can see that the higher mean frequency of phrasal compression features is not equivalent to the phrasal compression mentioned in Biber et al. (2011) or to that observed in research of student writing. The features of phrasal compression are used in beginner GRs as qualifiers of nouns in subjects and predicates. The noun phrases in beginner GRs usually contain only two or three noun sequences. In addition, the recurrent use of clausal elaboration features in intermediate GRs is related to the abundance of dialogues in the intermediate corpus.

The results indicate that lexico-grammatical complexity, as defined by Biber et al. (2011), does not increase across the GR levels investigated here. Nevertheless, the analysis of the excerpts shows that there is an increase in sentence elaboration from the beginner to the intermediate level. While beginner GRs have simple sentences with no elaboration and dependent clauses, GRs at an intermediate level have longer sentences, with more dependent clauses. This supports Norris and Ortega's (2009) argument that complexity measures need to be appropriate to the level being analyzed. That is, in this case, using Biber et al.'s (2011) hypothesized development framework was not appropriate to measure grammatical development in beginner and intermediate GRs. Possibly, other measures of grammatical development could be used to gain a better understanding of linguistic development in GRs. The results of this study indicate that clausal dependent measures of complexity might be appropriate for analysis of complexity between beginner and intermediate levels in GRs.

It is worth highlighting that the goal of this analysis was to measure linguistic variation across levels of GRs. This does not imply, however, that GR writers should be constrained by a list of linguistic features that they must cover in each level. GR writers need to take into consideration character development, complexity of the story, as well as vocabulary use. This paper suggests that publishers might want to include information about grammatical structures present at each level of development in GRs in order to inform teachers and help them better prepare reading activities with these books.

This study is not without its limitations, the main one being the sample of GRs analyzed. The GRs investigated were selected based on their availability at a specific IEP and might not represent the overall population of GRs. In addition, register differences in these GRs were not taken into account in the analysis. That is, narratives, biographies and other registers were

analyzed together based on their level. Future studies might want to investigate linguistic variation across these different registers.

Finally, the results of this study indicate that there is grammatical variation between beginner and intermediate GRs. The fact that publishers do not necessarily consider grammatical development when grading their readers by CEFR does not mean that in the future they should not take these into account when producing a new GR collection.

References

- Allan, R. (2008). Can a graded reader corpus provide ‘authentic’ input? *ELT Journal*, 63, 23-32. <https://doi.org/10.1093/elt/ccn011>
- Ansarifar, A., Shahriari, H., & Pishghadam, R. (2018). Phrasal complexity in academic writing: A comparison of abstracts written by graduate students and expert writers in applied linguistics. *Journal of English for Academic Purposes*, 31, 58–71. <https://doi.org/10.1016/j.jeap.2017.12.008>
- Biber, D. (1988). *Variation across speech and writing*. Cambridge University Press.
- Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly*, 45, 5–35. <https://doi.org/10.5054/tq.2011.244483>
- Biber, D., Gray, B., & Staples, S. (2016). Contrasting the grammatical complexities of conversation and academic writing: Implications for EAP writing development and teaching. *Language in Focus*, 2(1), 1–18.
- Bulté, B. & Housen, A. (2018). Conceptualizing and measuring syntactic diversity. *International Journal of Applied Linguistics*, 28, 147–164. <https://doi.org/10.1111/ijal.12196>
- Claridge, G. (2005). Simplification in graded readers: Measuring the authenticity of graded texts. *Reading in a Foreign Language*, 17, 143–158.
- Crawford, W. J., McDonough, K., & Brun-Mercer, N. (2019). Identifying linguistic markers of collaboration in second language peer interaction: A lexico-grammatical approach. *TESOL Quarterly*, 53, 180–207. <https://doi.org/10.1002/tesq.477>
- Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading*. Routledge.
- Hafiz, F., & Tudor, I. (1989). Extensive reading and the development of language skills. *ELT Journal*, 43, 4–11.
- Hafiz, F., & Tudor, I. (1990). Graded readers as an input medium in L2 learning. *System*, 18, 31–42. [https://doi.org/10.1016/0346-251X\(90\)90026-2](https://doi.org/10.1016/0346-251X(90)90026-2)
- Hill, D. R. (1997). Setting up an extensive reading programme: Practical tips. *The Language Teacher*, 21(5), 17–20
- Hill, D. (2008). Graded readers in English. *ELT Journal*, 62, 184–204. <https://doi.org/10.1093/elt/ccn006>
- Kuiken, F., & Vedder, I. (2019). Syntactic complexity across proficiency and languages: L2 and L1 writing in Dutch, Italian and Spanish. *International Journal of Applied Linguistics*. Special Issue. <https://doi.org/10.1111/ijal.12256>
- Kuiken, F., Vedder, I., Housen, A., & De Clercq, B., (2019). Variation in syntactic complexity: an introduction. *International Journal of Applied Linguistics*, Special Issue, 1–10. <https://doi.org/10.1111/ijal.12255>

- Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, 15, 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, 36–62. <https://doi.org/10.5054/tq.2011.240859>
- Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing*, 29, 16–27. <https://doi.org/10.1016/j.jslw.2015.06.003>
- Mohd-Asraf, R., & Ahmad, I. S. (2003). Promoting English language development and the reading habit among students. *Reading in a Foreign Language*, 15, 83–102.
- Nakanishi, T. (2015). A meta-analysis of extensive reading research. *TESOL Quarterly*, 49, 6–37. <https://doi.org/10.1002/tesq.157>
- Nation, I. S. P., & Ming-Tzu, K. W. (1999). Graded readers and vocabulary. *Reading in a Foreign Language*, 12, 355–380.
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30, 555–578. <https://doi.org/10.1093/applin/amp044>
- Parkinson, J., & Musgrave, J. (2014). Development of noun phrase complexity in the writing of English for Academic Purposes students. *Journal of English for Academic Purposes*, 14, 48–59. <https://doi.org/10.1016/j.jeap.2013.12.001>
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64, 878–912. <https://doi.org/10.1111/lang.12079>
- Staples, S., Egbert, J., Biber, D., & Gray, B. (2016). Academic writing development at the university level: Phrasal and clausal complexity across level of study, discipline, and genre. *Written Communication*, 33, 149–183. <https://doi.org/10.1177/0741088316631527>
- Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing*, 32, 17–35. <https://doi.org/10.1016/j.jslw.2016.02.002>
- Taguchi, N., Crawford, W., & Wetzel, D. Z. (2013). What linguistic features are indicative of writing quality? A case of argumentative essays in a college composition program. *TESOL Quarterly*, 47, 420–430. <https://doi.org/10.1002/tesq.91>
- Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The Effects of Repetition on Incidental Vocabulary Learning: A Meta-Analysis of Correlational Studies. *Language Learning*, 69, 559–599. <https://doi.org/10.1111/lang.12343>
- Waring, R. & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15, 130–163.
- Wodinsky, M. & Nation, P. (1988). Learning from graded readers. *Reading in a Foreign Language*, 5, 155–161.

Appendix A

Appendix A summarizes the descriptive statistics for each complexity feature investigated divided by GRs level. The table also shows the results of the Kolmogorov-Smirnov test of data normality.

Graded Readers	Grammatical feature	Min	Max	M	SD	Skewness	Kurtosis	Kolmogorov-Smirnov	
								Statistics	p
<i>Beginner</i>	Adverbial – Causative	.00	3.8	0.71	0.89	1.59	2.87	.216	<.001
	Adverbial – Conditional	.00	2.5	0.31	0.69	2.15	3.43	.427	<.001
	Adverbial – Others	.00	5.2	1.29	1.67	.004	4.06	.222	<.001
	Verb + <i>that</i> clauses	.00	6.7	1.41	1.56	1.31	2.24	.206	<.001
	coordinating conjunctions	.00	35.7	11.68	8.18	.944	1.12	.135	>.05
	Total verbs	136.1	221.8	185.23	20.94	-.662	-.074	.099	>.05
	Total nouns	188.6	388.8	273.18	44.24	.166	-0.29	.081	>.05
	Attributive Adjectives	9	88.8	27.31	16.58	2.55	7.56	.220	<.001
	Premodifying Nouns	1.7	43.4	12.65	9.03	1.59	2.63	.255	<.001
	Nominalization	.00	68.8	15.4	12.89	2.14	7.39	.167	<.001
	That – relative clauses	.00	2.4	0.46	0.65	1.49	1.51	.243	<.001
Wh-relative clauses	.00	3	0.46	0.88	1.94	2.53	.366	<.001	
<i>Intermediate</i>	Adverbial – Causative	.10	3.5	1.35	0.78	1.37	.31	.115	>.05
	Adverbial – Conditional	.10	4.1	1.92	1.09	.146	-.68	.076	>.05
	Adverbial – Others	1	8.7	4.06	2.01	.849	.357	.176	>.05
	Verb + <i>that</i> clauses	.30	5.9	2.94	1.57	.597	-.233	.181	>.05
	Clausal coordinating conjunctions	6.10	26.9	14.64	5.28	.901	.614	.145	>.05
	Total verbs	144.4	222.3	184.32	19.09	-.138	.519	.100	>.05

Total nouns	188.6	388.8	236.54	31	.166	-.029	.151	>.05
Attributive Adjectives	19	42.1	27.1	7.12	1.02	-.193	.223	<.001
Premodifying Nouns	4.3	23.6	8.94	4.62	1.58	3.58	.158	>.05
Nominalization	3.9	53	16.05	9.63	2.88	10.64	.309	<.001
That – relative clauses	.10	4.2	1.3	0.95	1.88	3.87	.201	<.001
Wh-relative clauses	.00	3	1.82	0.83	.60	3.30	.119	>.05

Appendix B

Graded Readers Corpus

Title	Publisher	CEFR Level
A Ghost in Love	OUP Bookworms	A1
Alissa	Macmillan	A1
Amazon Rally	Penguin Readers	A1
Around the World in 80 days	Macmillan	A1
Between Two Worlds	Longman	A1
Blog Love	Scholastic	A1
Blue Fins	Macmillan	A1
Goodbye, Mr. Hollywood	OUP Bookworms	A1
Gulliver's Travels	Macmillan	A1
In the Frame	Macmillan	A1
L.A Detective	Macmillan	A1
Land of Gold	Nuance Readers	A1
Love or Money	OUP Bookworms	A1
Lucky Number	Macmillan	A1
Photo Finish	Macmillan	A1
Sarah Says No	Macmillan	A1
Simone Decker and the Secret Formula	Express	A1
Ski Race	Macmillan	A1
The Crown	Penguin Readers	A1
The Lost Ship	Macmillan	A1
The Magic Barber	Macmillan	A1
The Umbrella	Macmillan	A1
The Well	Macmillan	A1
A Christmas Carol	Express	A2
A Picture to Remember	Cambridge English Readers	A2
Another World	Penguin Readers	A2
Dante's Peak	Pearson	A2
Dead Man's Island	OUP Bookworms	A2
Death in the Freezer	OUP Bookworms	A2
Earrings from Frankfurt	OUP Bookworms	A2
Five Children and it	OUP Bookworms	A2
Hampton House	Express	A2
Logan's Choice	Cambridge English Readers	A2
Pirates of the Caribbean	Penguin Readers	A2
The Earthquake	Penguin Readers	A2
Wonders of the Past	OUP Bookworms	A2
Black Cat	Macmillan	B1
Count of Monte Cristo	Penguin Readers	B1
David Copperfield	Express	B1

Frankenstein	OUP Bookworms	B1
Happy Christmas	Penguin Readers	B1
Life Exchange	Express	B1
Mysteries of the Unexplained	Penguin Readers	B1
Secret Codes	Penguin Readers	B1
Stories of Courage	Penguin Readers	B1
The Ironing Man	Cambridge English Readers	B1
The Portrait of Dorian Gray	Express	B1
The Railway Children	OUP Bookworms	B1
The Royal Family	Penguin Readers	B1
Two Lives	Cambridge English Readers	B1
Dorian Grey	Penguin Readers	B2
Jungle Love	Cambridge English Readers	B2
King`s Ransom	OUP Bookworms	B2
LA Movie Park	Macmillan	B2
Mosquito Coast	Penguin Readers	B2
Murder Maker	Cambridge English Readers	B2
Staying Together	Cambridge English Readers	B2
The Bride Price	OUP Bookworms	B2

About the Author

Larissa Goulart is a PhD candidate in Applied Linguistics at Northern Arizona University. She teaches English as a Second Language at the Program in Intensive English at the same university. Her areas of interest include using corpus tools in English teaching, writing in a second language and register variation. E-mail: goulart@nau.edu