

## Vocabulary Complexity in EFL University Students' Academic Texts

Complejidad del vocabulario en los trabajos de estudiantes universitarios de ILE

WILLIAM CHARPENTIER-JIMÉNEZ  
Escuela de Lenguas Modernas, Universidad de Costa Rica  
[wiliam.charpentier@ucr.ac.cr](mailto:wiliam.charpentier@ucr.ac.cr)  
Orcid ID: [0000-0002-8554-7819](#)

### Abstract

This paper examines students' use of vocabulary in English as a foreign language (EFL) in academic writing. However necessary, the specific vocabulary used by students in academic EFL settings has not received sufficient attention. Thirty-one EFL students at a Costa Rican public university participated in this study, which was conducted over the course of a year. The researcher collected data from participants' final research papers and used specialized software to determine their vocabulary usage. Data analyses indicate that: 1) lexical variety is above average, 2) students' academic vocabulary is high but does not include all the possible subtypes, 3) students range between a C1 and C2 English level in the English Vocabulary Profile (EVP) distribution, and 4) metadiscourse markers are used but highly repetitive in students' papers. These conclusions are consistent with the reviewed literature; they imply a high degree of variation across populations. Furthermore, the analysis suggests that direct instruction may broaden students' lexicon. These findings should serve as a springboard for implementing new teaching strategies that stimulate a curricular evaluation of the study plan.

**Keywords:** academic writing, higher education, language instruction, vocabulary

## Resumen

Este artículo examina el uso del vocabulario de los estudiantes de inglés como lengua extranjera (ILE) en la escritura académica. A pesar de su importancia, el vocabulario específico que los estudiantes utilizan en contextos académicos de ILE no ha recibido la suficiente atención. Este estudio se llevó a cabo con 31 estudiantes ILE de una universidad pública en Costa Rica durante un año. El investigador recolectó datos provenientes del trabajo final de investigación de los estudiantes utilizando programas especializados. Los datos indican que: 1) la variedad del léxico está por encima del promedio, 2) el vocabulario académico de los estudiantes es alto pero no incluye todos los posibles subtipos, 3) los estudiantes se encuentran entre un nivel C1 y C2 en la distribución del perfil de vocabulario en inglés (*EVP*), y 4) los marcadores metadiscursivos son utilizados pero altamente repetitivos en los trabajos de los estudiantes. Estas conclusiones se asemejan a la literatura consultada, pero sugieren una gran variabilidad entre diferentes poblaciones. Además, el análisis sugiere que la instrucción directa puede incrementar el léxico de los estudiantes. Estos resultados deben servir para implementar nuevas estrategias de aprendizaje que estimulen la evaluación curricular del plan de estudios.

**Palabras clave:** escritura académica, educación superior, instrucción del lenguaje, vocabulario

## Introduction

Over the last two decades, there has been a surge of interest in vocabulary acquisition. According to Webb and Nation (2017), more than 30% of vocabulary research has been conducted in the last century. In addition, the advent of the internet, corpora, and vocabulary analysis software has propelled how lexicons are examined in real and classroom settings. This has also improved vocabulary instruction, especially in the context of foreign language learning. However, language instructors often struggle to measure students' vocabulary acquisition despite its importance. On the one hand, teachers can only measure some vocabulary components (form, meaning, or use) through specific means (productive or receptive) and in limited numbers. On the other hand, the field of language instruction lacks "a comprehensive set of tests which allow us to easily and reliably test every aspect of a learner's vocabulary knowledge" (Milton, 2009, p. 17). Thus, it becomes critical to have a broad understanding of students' vocabulary knowledge in order to improve materials, instruction, and the curriculum.

Further, this notion is often underrated in language programs. Therefore, this study examines students' vocabulary usage in written texts. The main lexical features under investigation are lexical diversity, the English Vocabulary Profile (EVP), the Academic Word List (AWL), and metadiscourse. Although research on vocabulary acquisition is abundant, theory and practice have not always complemented each other well. According to Pavićić Takač (2008), evidence suggests that

"psycholinguists have a particular interest in vocabulary development and exploration of the formal models of vocabulary acquisition, and ignore the L2 vocabulary literature" (p. 17). However, the impact on classrooms becomes limited since "applied linguists [...] are mainly concerned with the descriptive aspects of vocabulary and do not draw on existing psycholinguistic models of bilingual lexicon" (p. 17). According to the author, this dichotomy has resulted in the bifurcation of vocabulary research and the creation of a gap between them. Since the present study focuses on non-native English speakers enrolled in an English or English teaching major, it seeks to contribute to the existing literature and the improvement of vocabulary instruction.

In addition, not all instructional settings, such as those in Latin America, can afford to use a series of textbooks that systematically incorporate vocabulary into their content, and syllabi or materials created by professors often do not include expected word knowledge. However, while incidental vocabulary learning is favored over intentional vocabulary learning, it still does not receive proper follow-up treatment. Likewise, while vocabulary may attract more attention at more basic levels, such as high school or elementary school, university education does not focus on general words, idioms, or academic language. Moreover, to this day, the BA in English and BA in English Teaching do not consistently incorporate vocabulary instruction or guidance on which words should be emphasized. Thus, incidental vocabulary learning depends on the readings and materials selected by professors, which vary from professor to professor

or from term to term; however, it remains vague, and proper vocabulary measurement techniques are limited.

The findings of this paper may benefit two groups. First, analyzing vocabulary use will impact students in two ways. Initially, it is hoped that the findings will impact material development and language instruction. Therefore, students will be exposed to a richer and more informed set of vocabulary across the study plan. They may also receive training on how to analyze their writing and vocabulary use. Although the most powerful applications are payware, some downloadable or online freeware allows students to extract valuable and individual data. Second, curriculum developers may use the findings as a starting point for reforming the curricular content of these majors. A more robust language program will prepare students better and also attract more students. Although some of the courses and the context are relatively specific, these results may indirectly motivate other institutions to revise their vocabulary components and use the most appropriate software to conduct their own studies.

This paper is organized into six distinct sections. The introduction discusses the consideration as well as the relevance of the findings. The literature review defines the main vocabulary-related concepts and surveys previous related studies. It also summarizes some of the key ideas from several vocabulary research. Then the methods section explains the methodology employed, as well as the materials, participants, and data collection procedures. The results section summarizes the main findings. The discussion section contains the principal

results of the study and their implications. Finally, the conclusion examines some possible limitations and presents recommendations for future research.

## Literature Review

Existing research has focused on vocabulary acquisition but failed to explore what vocabulary students incorporate into their writing. In addition, while linguists have broadly used the role of computer software to describe vocabulary use, it has eluded practical applications in applied linguistics. The myriad of online and downloadable software has increased, and the options for language teachers and learners have also expanded. This section explores lexical diversity, the English Vocabulary Profile (EVP), the Academic Word List (AWL), and metadiscourse, among the many possibilities to measure vocabulary. It also summarizes the main concepts related to this field.

## Preliminary concepts

The metalanguage used to describe words has widened in recent years. In this sense, Nation (2013) described four possible categories for classifying and thus measuring vocabulary: tokens, types, lemmas, and word families. Tokens fall under the traditional definition of a word and include repeated words. For example, in the sentence, *Her intelligence is her best asset*, the word *her* is repeated, but it accounts as two separate tokens. However, the word *her* would only be counted once in the category of types. When the number of tokens is close to the number of types, we can assume that the writer employed more unique words.

Words are classified using lemmas and word families based on their inflections. A lemma is a headword and its inflected or reduced forms. For example, in the sentence, *I study hard, but she studies harder*, *study* and *studies* are classified as one lemma only, the same as *hard* and *harder*. Word families also include inflectional and derivational affixes. Therefore, it accepts words from different parts of speech. For example, in the sentence *My friends value our friendship*, *friends* and *friendship* belong to the same word family. Thus, the analysis conducted in this paper includes tokens and types only.

### Lexical diversity

McCarthy and Jarvis (2007) described lexical variety as “the range and variety of vocabulary deployed in a text by either a speaker or a writer, as opposed to the potential vocabulary that a speaker or writer may have available but is not currently using” (p. 1). Jarvis and Daller (2013a) defined it simply as “a measure of the variety of words in a text” (p. 113). According to the authors, evidence suggests that more proficient writers and speakers consistently produce a greater variety of words. To calculate lexical diversity, divide the word types by tokens. The greater the number of types compared to tokens, the more lexical variety is present in a text. On the contrary, a higher number of tokens compared to types indicates word repetition (Jarvis & Daller, 2013a). Word variety is vital since it adds richness to the text and makes it less monotonous. Nonetheless, Duran et al. (2004) stress that lexical diversity is “about more than vocabulary range. Alternative terms,

‘flexibility,’ ‘vocabulary richness,’ ‘verbal creativity,’ or ‘lexical range and balance’ indicate that it has to do with how vocabulary is deployed as well as how large the vocabulary might be” (p. 221-222).

Several recent studies have focused on lexical diversity. Some of these studies suggest that lexical diversity does not improve as students advance in their majors. For example, Vidal and Jarvis (2020) found that lexical variety was not significantly better in third-year students than in first-year students. In a similar study, Ha (2019) sought to discover the possible relationship between lexical diversity and writing quality. Ha (2019) found that “unlike previous research that assessed lexical diversity and density, there was no correlation between lexical diversity or lexical density and students' writing [quality]” (p. 21). Akbari (2017) also presented similar results supporting the idea that students incorporated fewer “diverse lexical choices and used fewer academic and lower frequency words in their essays compared to NS students. [...] no difference was observed between the essays written by EFL students in Year 1 and Year 2 in this regard” (p. 16). According to these researchers, lexical diversity does not seem to predict improvement in writing.

On the other hand, other investigations have specifically addressed the importance of lexical diversity when analyzing EFL/ESL students' writing. For instance, lexical diversity seems to correlate with quality in academic writing (Erarslan, 2021) and higher language proficiency (Crossley & McNamara, 2012; McNamara et al., 2010). Lexical diversity indicates bet-

ter written linguistic proficiency compared to other indices, such as depth of word knowledge features or access to core lexical items (Crossley et al., 2011). The environment also seems to affect lexical diversity. For instance, students immersed in second language contexts produce more varied vocabulary than those immersed in foreign contexts. Context also tends to close the gap between native and non-native speakers in terms of lexical diversity (Foster & Tavakoli, 2009). In addition, lexical diversity “is commonly thought to reflect greater linguistic skills, speaker competence, or even a speaker’s socioeconomic status” (Ransdell & Wengelin, 2003). Therefore, lexical variety does not mean word production only. Its relation to writing quality and proficiency also conveys information about the writer’s background.

Although research on lexical variety seems contradictory, its effects on students cannot be denied. As research expands and new data is gathered, language specialists can make more informed decisions about improving language instruction in general and writing instruction in particular.

### **English Vocabulary Profile (EVP)**

Research has also focused on developing databases for the purpose of creating a vocabulary profile. In general terms, an English profile should “throw more light on what learners of English can and can’t do at different CEFR levels, and to assess how well they perform using the linguistic exponents of the language at their disposal” (Milanovic, 2009, p. 5). The vocabulary profile outlines the vocabulary that students at each level should be able

to use. It is directly linked to the bands in the CEFR. This information may help create materials, curricula, and instruction that are more suited to the needs of individual students.

Further, as with lexical diversity, vocabulary profiles offer several advantages. On the one hand, they are related to academic performance and proficiency in a language. For instance, vocabulary profiles may aid in predicting academic performance when used with interviews and other traditional tests (Morris & Cobb, 2004a). Vocabulary profiles also provide a more objective indication of a learner’s proficiency level than other tests (Leńko-Szymańska, 2015) and are helpful in determining students’ vocabulary production at university levels (Sun, 2017a). On the other hand, they include multi-word units like idioms or phrasal verbs (Granger & Larsson, 2021), so a more holistic analysis of language use is possible. Finally, information about EVPs can be used to predict grades in writing (Alfter et al., 2016). Thus, all of this provides a wealth of information for applied linguistics.

Although EVPs have sound advantages, some research has pointed out two main disadvantages. First, they do not seem helpful in preparing materials and developing exams. Second, students may tend to memorize word lists unnaturally. Consequently, this might hinder a more natural language learning process (Sun, 2017b). Nonetheless, the advantages outweigh the possible drawbacks, and EVPs serve as a common reference point for curricular designers and language professors in general.

### The Academic Word List (AWL)

The academic word list is an organized record of recurrent words present in corpora from various disciplines (Coxhead, 2000). Its primary purpose was to aid ESL first-year students in reading academic texts (Nation, 2016). Knowing which words students can use becomes essential at the university level, where they have to read and write academic texts.

Several reasons support the use of the AWL. First, knowing these words allows students to better understand academic texts from a variety of disciplines, write clear academic texts, and be part of the academic community (Coxhead, 2006). This also helps them be more successful in their academic endeavors. On the other hand, a lack of academic vocabulary may hinder students' academic success. Moreover, according to Cardullo et al. (2017), students who lack proper academic vocabulary struggle in an academic setting since they cannot convey their ideas competently in writing. Similarly, other findings suggest that an oral interview alone cannot demonstrate students' vocabulary knowledge, especially academic vocabulary. However, students' formal writing should also be analyzed to assess their readiness for university tasks (Morris & Cobb, 2004b). Finally, the literature suggests that vocabulary can be learned intentionally in EAP courses (McDonough et al., 2018) or even incidentally (Reynolds, 2015), and students can use the AWL to study independently (Coxhead, 2000; Lessard-Clouston, 2013).

Despite these advantages and the potential benefits for English teachers (Lessard-Clouston, 2013) and course

and material designers (Coxhead, 2000), the AWL has received substantial criticism since its inception. In some studies, students' composition grades and their use of academic vocabulary displayed a weak correlation (Alhojailan, 2019). Additionally, the number of academic words used by EFL students decreased significantly between Sublists 1 and 10. However, the list does not indicate what this means or its implications (Charpentier-Jiménez, 2019). Furthermore, these sublists may give students the wrong impression that some words are more valuable or sophisticated than others (Durrant, 2016). Finally, while the AWL asserts that it is derived from and applicable to a variety of academic disciplines, some researchers have argued that no single word list can adequately fit different academic contexts (Durrant, 2016). This idea could ultimately deceive students into believing that learning such words could be beneficial in all fields (Hyland & Tse, 2007).

As is evident, the findings are far from conclusive. Although word lists in general and the AWL, in particular, cannot be used in several contexts, they can guide teachers, students, and language program designers to set specific learning objectives. Knowing what words students produce and how they compare to the AWL will help include particular vocabulary objectives, especially in institutions where they have previously been overlooked.

### Metadiscourse Markers

In terms of language use, linguists often divide words into two main categories: content words (also known as open-class words) and function words

(also known as closed-class words) (Akmaijan et al., 2017; Baker & Hengeveld, 2012; Matthews, 2014; Radford, 2010). In general, content words include nouns, verbs, adjectives, and adverbs. On the other hand, function words deal more with pronouns, conjunctions, and prepositions, among other words. However, metadiscourse analysis is the only analysis that deals exclusively with function words. According to Hyland (2004), it refers to “the linguistic devices writers employ to shape their arguments to the needs and expectations of their target readers [and...] which help relate a text to its context by assisting readers to connect, organise, and interpret material in a way preferred by the writer” (p. 134). Applied linguists, especially those in second language settings, are known in a more functional sense as logical connectors, sequencing items, and hedges, among others (Hyland, 2004). Consequently, they contribute to the sequence and overall structure of discourse.

There are several advantages to using metadiscourse markers. Research has shown a significant relationship between the number of metadiscourse markers and essay writing quality (Sanford, 2012; Sešek, 2016). Similarly, other studies have revealed that they improve writing proficiency in second language learners (Cheng & Steffensen, 1996; Dastjerdi et al., 2010; Kaya & Sofu, 2020; Yaghoubi & Ardestani, 2014). At the cognitive level, metadiscourse markers provide students with a practical understanding of how a text should be processed (Hyland, 2010). Moreover, beyond the direct benefits for students, metadiscourse markers make the text easier

for readers to process (Hyland, 2010) and help writers negotiate meaning with their readers and interact with them, especially at advanced levels of academic writing (Hyland, 2004). In terms of instruction, research suggests that students benefit from the explicit and implicit teaching of metadiscourse markers (Yaghoubi & Ardestani, 2014), making it a perfect complement for academic writing settings.

Further, previous research has not evidenced negative aspects of teaching, analyzing, or even making students aware of the metadiscourse markers. However, several concessions must be granted. First, mastering metadiscourse markers seems difficult for ESL/EFL students (Sanford, 2012). This is particularly worrying when writing courses often fail to include them as part of the curriculum (Mei Hooi et al., 2020). It has also been shown that metadiscourse use improves through exposure and experience (Gu & Xu, 2021; Yüksel & Kavanoz, 2018). Thus, a final possible limitation for the present study is the subdivision of discourse markers (Bax et al., 2019; Hyland, 2004). Discourse markers have been classified in a variety of ways by researchers (Crismore et al., 1993; Hyland, 2019; Vande Kopple, 1985). Hyland (2005) subdivides these markers into interactive and interactional. According to Mat Zali et al. (2021), interactive markers “help the writer to sort out propositional substance to make it clear” (p. 22), whereas interactional markers “permit the author to comment on their messages” (p. 22). The list of metadiscourse markers is shown in tables 1 and 2.

**Table 1**  
*List of Interactive Markers*

Interactive markers	Definition	Examples
Code gloss	Elaborates propositional meaning	for example, for instance, e.g., i.e., that is
Endophoric	Directs to information in other parts of the text	see, noted, discussed below, discussed above, discussed earlier
Evidential	Directs to information in other texts	according to, cite, quote, established, said
Frame*	Refers to discourse acts, sequences, and stages	my purpose, to move on, in regard to, to start with, to conclude
Transition	Expresses relations between main clauses	and, but, therefore, thereby, on the other hand

*Note.* Text Inspector separates *frames* as announce goals, sequencing, label stages, and topic shifts. Transitions are also called logical connectives.

**Table 2**  
*List of Interactional Markers*

Interactive markers	Definition	Example
Attitude marker	Expresses the writer's attitude to a proposition	admittedly, I agree, amazingly, appropriately, correctly
Boosters	Emphasizes certainty and closes the dialogue	actually, always, apparent, I believe, certain that
Engagement	Explicitly builds a relationship with the reader	incidentally, by the way, determine, consider, imagine
Hedge	Avoids commitment and opens the dialogue	almost, apparently, appear to be, approximately, assume
Self-mention	Explicitly refers to authors	I, we, me, my, our

*Note.* Text Inspector does not always use these labels. Self-mentions are also referred to as person markers. Boosters are also called emphatics. Relational markers are also known as engagement markers.

As can be seen from these tables, the possible categories are grouped into 13 subcategories in Text Inspector but only into 10 for Hyland (2005). Nevertheless, direct analysis is still possible since sub-categories are equivalent or just subdivisions of other broader groups.

This literature review presents some of the most important results and concepts regarding vocabulary analysis. Lexical diversity, the EVP, the AWL, and metadiscourse markers have all been explored to provide a theoretical basis for the present study. In summary, previous research has supported the benefits of exploring students' vocabulary use, particularly in ESL/EFL writing contexts. However, while there is a broad agreement that vocabulary should be analyzed and taught, it remains controversial whether the results should be generalized to various contexts and populations.

## Method

### Participants

The population of this study comprises adult, Costa Rican students taking a fourth-year writing course in English as a second language or an English teaching major. The researcher created a personal electronic mailing list of 55 students interested in participating in the study. Of those 55 students, 31 sent their final paper within the indicated time. Therefore, only these 31 students were part of the study. The participants were selected because they had taken all the required writing courses in their majors. The list included every student who agreed to participate in the study. All students speak Spanish as their first language.

### Materials

An electronic, written consent was prepared and sent to invite students to participate. In addition, an anonymous file request system was provided for students to upload their final project anonymously. Students uploaded their integral written work, and the researcher did not intervene or request any format, topic, or criteria. Students' papers include an average of 30 pages per file. To guarantee anonymity, the researcher asked students to eliminate all metadata, including names and document properties. All direct quotations and names in references were deleted since they do not represent the students' writing. Documents were not modified in any other manner, and no final paper was kept from the analysis.

A full version of the online text analysis tool Text Inspector (Bax, 2012) was used to analyze the text. This software offers lexical diversity, an English vocabulary profile, academic vocabulary, metadiscourse, readability, and parts of speech analysis. All of the data was obtained through an in-depth examination of students' writing.

### Procedure

This study used a quantitative, direct needs assessment design. To obtain data, the researcher asked students to voluntarily participate in the study, and the purpose of the study was explained via Zoom. Next, students were sent a link to write their email if they were willing to participate in the study. Afterward, a first email was sent with the written consent and the file request address. Next, students were asked to delete all their personal

information from the document's body and document properties. Before starting the process, the researcher revised the 31 papers to verify this step. The researcher also deleted all direct quotations, names, and any text that the students did not create. Finally, data were obtained and analyzed using the Text Inspector software and central tendency measures or descriptive statistics.

Text Inspector uses two measures to rate lexical diversity. The Measure of Textual Lexical Diversity (MTLD) and vocc-D were analyzed together as they are more reliable than any single measure alone. The MTLD analyses lexical diversity through

a number of operations. Firstly, program separates the text into individual unlemmatized token instances, one token at a time. A TTR score is calculated each time a new type is found. When the TTR score reaches the factor value (default 0.71) the text is cut and a count of the tokens in that factor is recorded. Having cut the text at its factor size, MTLD then resets its TTR value at 1.0 and the process is repeated. (McCarthy, 2005, p. 94-95)

On the other hand, McCarthy and Jarvis (2007) explain that "the vocc program outputs an LD index that is calculated through a series of TTR [type-token ratio] samplings and curve fittings" (p. 460). As mentioned before, it is recommended to run both measures to obtain a better perspective of the writers' lexical diversity.

In principle, the procedure for computing the writers' EVP, AWL vocabulary, and metadiscourse markers is more straightforward. First, students'

texts are compared and contrasted with various lists. A contrast to the Cambridge Learner Corpus (CLC) triggers the EVP results. Next, academic vocabulary was obtained by comparing students' texts to the AWL developed by Coxhead (2000). Finally, metadiscourse markers results were computed by analyzing students' texts and the types identified by Bax et al. (2019). However, manual revision is still necessary as some words may overlap with other categories based on their use.

## Analysis of the Results

The following description and analysis describe students' vocabulary use in their final research papers. Except for lexical diversity, all data were analyzed from three perspectives. First, results are compared to the larger number of tokens used by students. Second, the number of tokens and types are contrasted. As mentioned earlier, a similar number of tokens and types often indicate the use of unique words. Finally, subcategories are compared to elucidate which ones are commonly used by students.

Of the 31 students who submitted their final papers, 24 (77.41%) were females, and seven were males (25.5%). Overall, 23 students (74.19%) reported being between the ages of 18 and 24. Seven students (22.58%) were between the ages of 25 and 34. One student (3.22%) was between 35 and 44. All students are native Spanish speakers and use English as a foreign language. In terms of academic level, the research included 23 students (74.1%) enrolled in the BA in English, six (19.35%) in English Teaching, and two (6.45%)

in both majors. All students were enrolled in the last courses of the study block (fourth year, eighth semester).

As mentioned previously, the most common types of procedures for calculating lexical diversity are MTLD and vocd-D. In both cases, the letter *D* represents lexical diversity. According to Fergadiotis et al. (2015), “for any given sample, MTLD reflects the average number of consecutive words for which a certain TTR is maintained” (p. 3). In contrast, for vocd-D, Duran et al. (2004) explain that an adult ESL user should obtain a measure of 40 to 70 in the vocd-D procedure. On the other hand, texts should receive a score of 80 or above to be considered academic. The maximum reported score was above 100, 105 to be precise. The proficiency median was drawn at 92.5. According to the Text Inspector analysis, students’ vocd-D score was 94.91 (MTLD=69.59), indicating an optimum lexical diversity in academic writing.

To explore the number of academic words used by students in their final research papers, Text Inspector

contrasts the students’ corpus with the AWL. As can be calculated from Table 3, students’ texts included around 11.81% ( $n = 19265$ ) of academic vocabulary. According to Nation (2013), on average, academic words “make up about 9% of the running words in the text” (p. 16). This indicates that students’ use of academic vocabulary is above the expected mean. Additionally, the tokens per type (TTR) ratio in listed academic words was 29.77%. In general terms, these results suggest that the repetition of academic vocabulary is also high. Such repetition is more evident when the subcategories are compared. The minimum TTR was 26.42% in sublist 1, whereas the maximum TTR was 51.08% in sublist 10, with a median of 36.20%. All TTR ratios increased (except for sublists 5 and 6) according to the order of the sublists. The same phenomenon also occurs with tokens alone. In this case, the higher the AWL subclass, the fewer the academic words used by students. For example, sublists 1 and 2 comprised 53.34 % of all academic vocabulary used by students.

**Table 3**  
*Text Coverage Based on the Ten AWL Subclasses*

AWL Sublist	Tokens		Types	
	n	%	n	%
One	5,375	3.30	1,420	5.30
Two	4,887	3.00	1,157	4.32
Three	1,799	1.10	640	2.39

---

Four	2280	1.40	600	2.24
Five	1436	0.88	470	1.75
Six	1,306	0.80	481	1.80
Seven	909	0.56	354	1.32
Eight	641	0.39	315	1.18
Nine	446	0.27	205	0.77
Ten	186	0.11	95	0.35
Unlisted	143,579	88.17	21,057	78.59
Total	162,844	100	26,794	100

---

Note. AWL = Academic Word List

The EVP reflects students' vocabulary production according to the CEFR bands. Table 4 summarizes the overall trend of EVP lexis used by students in their final papers. The Council of Europe (2020) does not provide numerical thresholds to place students at a specific level. Instead, descriptors such as "has a good command of a very broad lexical repertoire" (p. 131) or "can select from several vocabulary options in almost all situations by exploiting synonyms of even words/signs less commonly encountered" (p. 131). However, the Text Inspector provided an approximate word-level band based on the corpus used. As a whole, the results indicate that students' vocabulary use ranges from C1 to C2, with some text falling into the D (academic) level. This data matches the text coverage based on the ten AWL subclasses previously

discussed. Considering the listed number of tokens ( $n = 144,714$ ) and word types ( $n = 22484$ ) analyzed, the TTR was 15.53%. Compared to academic vocabulary use, the EVP displayed a greater (and expected) repetition of words. This occurred since many words, especially in the lower bands, tend to be high-frequency words. In addition, this low versus high-frequency organization correlated with students' vocabulary use. Although the results presented an abnormal shift between the A2 and B1 bands, as the level of the band becomes higher and more low-frequency words are used, the students' number of words decreases.

**Table 4**  
***Text Coverage Based on the Six Bands in the CEFR***

EVP	Tokens		Types	
	n	%	n	%
A1	86,726	53.26	5,044	18.83
A2	14,285	8.77	3,581	13.36
B1	20,448	12.56	5,659	21.12
B2	15,697	9.64	5,406	20.18
C1	5,316	3.26	1,892	7.06
C2	2,242	1.38	902	3.37
Unlisted	18,130	11.13	4,310	16.09
Total	162,844	100	26,794	100

Note. CEFR = Common European Framework of Reference, EVP = English Vocabulary Profile.

Although the Text Inspector listed thirteen categories of metadiscourse markers, some of those categories were grouped and analyzed, following Hyland's (2019) taxonomy. A picture of metadiscourse use is presented in Tables 5 and 6. Owen et al. (2021) highlight that higher-level English users "will typically deploy more metadiscourse markers simply by virtue of producing more words in total" (p. 27). Analyzing the total tokens and the types used in relation to the whole text did not trigger clear results. Therefore, the category of *unlisted* words was not included in the analysis.

Overall, interactive markers were the most used. The TTR of interactive markers was 9.62%. However, this high repetition does not necessarily imply a lack of vocabulary on the part of students. In general, discourse markers are a high frequency but relatively closed category. In addition, some of these words have no synonyms. When comparing the subcategories, it was evident that code gloss and endophoric markers received less attention, which is an abnormal circumstance considering the type and usefulness of these words in academic writing.

On the other hand, although the variety was greater ( $TTR = 23.32\%$ ), interactional markers were not prolific in students' writing. However, personal and engagement markers are unsurprisingly absent.

Some books, professors, and even editors recommend writers avoid personal pronouns such as *I* or *us*, as a more impersonal style is often associated with formal academic writing.

**Table 5**  
*Results of Interactive Marker Analysis from Students' Papers*

Interactive Markers	Tokens		Types	
	n	%	n	%
Code gloss	545	4.30	140	11.48
Endophoric	638	5.03	84	6.89
Evidential	1,873	14.78	250	20.51
Frames	1,292	10.20	289	23.71
Logical connective	8,324	65.69	456	37.41
Total	12,672	100	1,219	100

*Note.* As a reference, the estimated number of metadiscourse markers in the entire text is 9.93%. Interactive markers comprise 7.78% of the whole text and 78.38% of discourse markers used.

**Table 6**  
*Results of Interactional Marker Analysis from Students' Papers*

Interactional Markers	Tokens		Types	
	n	%	n	%
Attitude	529	15.14	122	14.97
Boosters	835	23.89	227	27.85

Hedge	1,419	40.60	328	40.25
Person	428	12.25	60	7.36
Engagement	284	8.13	78	9.57
Total	3,495	100	815	100

*Note.* As a reference, the estimated number of metadiscourse markers in the entire text is 9.93%. Interactional markers comprise 2.15% of the whole text and 21.62% of discourse markers used.

The data provides strong evidence about students' vocabulary use. Four separate categories were used to analyze students' final papers. The metric of lexical diversity indicates a proficiency level suitable for writing academic texts. The AWL measures demonstrate a good command of academic vocabulary, with some categories receiving less attention. In terms of the EVP, the final results provided by Text Inspector reveal that students' vocabulary use is at or above a C level, with data reaching the academic vocabulary use. Finally, students have sufficient knowledge of metadiscourse to construct coherent sentences and establish textual coherence. Nevertheless, some markers are almost omitted. Although these results cannot be generalized, the implications of these data provide significant insights in terms of what students can produce and what they need to improve their writing.

## Limitations

While all measures were taken to guarantee a careful data analysis, three main limitations were acknowledged. First, although powerful, the computational method used has some

inherent limitations. For example, it is impossible to know if students use words with the intended meaning and context. The software can only generate a list of words based on their form. Some authors (Jarvis & Daller, 2013a; Koizumi & In'ami, 2012; Stills, 2016) have suggested that text length should be limited for lexical diversity analysis. However, this study intends to provide a clear but general picture of the students' full production.

Second, no individual data can be obtained from these results. The analysis includes students as a whole; therefore, highly proficient or low proficient English users can skew the data. The study cannot predict how proficient or sophisticated an individual is compared to the rest.

Finally, due to the nature of students' final research paper writing process, other limitations were considered. First, students were able to write and revise their writing at home. Therefore, they had access to dictionaries and other sources to improve their vocabulary. Additionally, considering the nature of a research project, students had to read plenty of academic literature. Thus, some words may have been advertently or inadvertently borrowed from journal articles or books. Although this

type of incidental recognition and vocabulary application is favorable, this knowledge may be spontaneous and temporary, and it does not guarantee that students have mastered the lexicon. Finally, students have access to formulaic vocabulary provided by their instructors, websites, and even books that aid (novice) writers in structuring and including vocabulary suitable for a research paper (Barros, 2016; Howe & Henriksson, 2007). Therefore, some of the vocabularies analyzed here may not be part of students' linguistic repertoire or working vocabulary.

## Discussion

Academic writing encompasses a variety of essential aspects. Vocabulary usage represents one of the main aspects since it helps convey meaning and link ideas in the text. The findings suggest a positive correlation between students' working vocabulary and the metrics adopted here. First, lexical diversity indices demonstrate that students can produce beyond what is expected from an adult English language user. In addition, the text displays an appropriate level of vocabulary for academic writing. The implications of these findings are twofold. First, the findings demonstrate that students can perform adequately in real-life academic scenarios where writing is a fundamental skill in today's world. Concerning the course program, this indicates that, although direct vocabulary instruction is not addressed, incidental vocabulary acquisition permits students to perform at an academic level.

Second, this academic level is also demonstrated in terms of academic

words used. Students' working vocabulary is above average, considering the AWL as a reference. Although adequate, attention should be paid to the sublists used and what this means for acquisition and production. As mentioned earlier, students tend to repeat some vocabulary when other options are also possible. The core courses of the BA in English and English Teaching do not have objectives directed to maximize vocabulary, and its instruction is absent from the course programs. Thus, not having a solid vocabulary goal may hinder students from accessing richer and more varied vocabulary included in other sublists of the AWL.

Third, the EVP from students' writing also places them between the C1 and C2 levels. Overall, students' scorecards show that their vocabulary production also reaches band D, indicating academic vocabulary use. Since English and Spanish share some Latin roots, students may be familiar with C1 and C2 vocabulary. Knowing what words are cognates between both languages could direct students and professors to include those less familiar words in the syllabi. This suggestion deserves particular attention since the EVP revealed more word repetition in all categories included in this study. Equally important is to analyze students' individual production and work with them according to their linguistic needs in vocabulary acquisition.

Finally, some conclusions can be drawn from the metadiscourse analysis. On the one hand, students use a wide array of metadiscourse markers; however, some repetition becomes evident, especially in interactive markers. Also, students used interactional

markers at a significantly lower rate than interactive markers. In terms of linguistic instruction, the core courses of the major do not make this distinction; however, contrary to vocabulary instruction, they do emphasize using metadiscourse markers. This implies that some revision becomes necessary to include other types of markers or discover why students are not using vocabulary from all the categories. As stated previously, some books, professors, and even editors directly instruct writers not to use some markers. Still, no clear guidelines exist in course programs, leaving this issue to the instructor's personal choice.

This discussion also triggers some suggestions, especially for language programs. First, a culture of self-analysis should be created. Ideally, students should often analyze their production. This is not to say that they should always do it or become obsessed with vocabulary use. Still, educated use of the freely available tools online will make students aware of the specific vocabulary they use and where they have to improve. Second, professors should inspect students' texts to make better decisions about what they know and need, even at the initial levels. Additionally, in-class or impromptu writing activities should be carried out recurrently since results may differ when students write at home. Finally, language programs should incorporate direct vocabulary instruction. Accordingly, curricular experts and professors should develop guidelines and materials suitable for students' needs.

The questions raised by this study warrant further investigation. For example, comparisons between populations (i.e., various undergraduate

programs, institutions, or levels) could be made to understand how vocabulary is used in different settings. In addition, longitudinal studies could be carried out to verify students' improvement, especially after new linguistic content is incorporated into the curriculum. Another line of research could also address the limitations of this study by taking into account shorter and spontaneous pieces of writing where students' production is not affected by external factors. For example, the use of the dictionary may sometimes be restricted to guarantee individual production. Finally, further research should be undertaken to determine to what extent direct instruction benefits production. In this sense, professors' and students' opinions should be part of any study that seeks to improve language acquisition at a higher education level.

The researcher invites other professors, institutions, or researchers to replicate the findings of this study. Analysis software development greatly enhances revision costs, speed, and organization. Furthermore, once the specific strengths and weaknesses have been detected, informed language policies and guidelines become easier to create and establish.

## References

- Akbari, N. (2017). Lexical diversity and the use of academic and lower frequency words in the academic writing of EFL students. *Australian Review of Applied Linguistics*, 40(1), 3–18. <https://www.jbe-platform.com/content/journals/10.1075/aral.40.1.02akb>
- Akmajian, A., Farmer, A. K., Bickmore, L., Demers, R. A., & Harnish, R. M. (Eds.). (2017). *Linguistics: an introduction to language and communication* (Seventh edition). The MIT Press.
- Alfter, D., Bizzoni, Y., Agebjörn, A., Volodina, E., & Pilán, I. (2016). From distributions to labels: A lexical proficiency analysis using learner corpora. *Proceedings of the Joint Workshop on NLP for Computer Assisted Language Learning and NLP for Language Acquisition*, 1–7. <https://aclanthology.org/W16-6501>
- Alhojailan, A. I. (2019). The Effect of Academic Vocabulary Use on Graduate Students' Writing Assignment Scores. *English Language Teaching*, 12(9), 33. <https://doi.org/10.5539/elt.v12n9p33>
- Baker, A., & Hengeveld, K. (Eds.). (2012). *Linguistics: the basics*. John Wiley & Sons.
- Barros, L. O. (2016). *The only academic phrasebook you'll ever need: 600 examples of academic language*. Amazon Fulfillment.
- Bax, S., Nakatsuhara, F., & Waller, D. (2019). Researching L2 writers' use of metadiscourse markers at intermediate and advanced levels. *System*, 83, 79–95. <https://doi.org/10.1016/j.system.2019.02.010>
- Bax, S. (2012). Text Inspector. Online text analysis tool. Available at: <https://textinspector.com/>.
- Cardullo, V., Finley, S., Burton, M., & Tripp, L. O. (2017). Attitudes, Perceptions, and Knowledge - Academic Language and Academic Vocabulary of Pre-Service Teachers. *Journal of Higher Education Theory and Practice*, 17(9). <https://doi.org/10.33423/jhet.v17i9.1418>
- Charpentier-Jiménez, W. (2019). University Students Use of Academic Vocabulary in the BA in English and English Teaching. *Revista De Lenguas Modernas*, 30(2).
- Cheng, X., & Steffensen, M. S. (1996). Metadiscourse: A Technique for Improving Student Writing. *Research in the Teaching of English*, 30(2), 149–181. <https://www.jstor.org/stable/40171358>
- Council of Europe (Ed.). (2020). *Common European framework of reference for languages: learning, teaching, assessment ; companion volume*. Council of Europe Publishing.
- Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34(2), 213–238. <https://doi.org/10.2307/3587951>
- Coxhead, A. (2006). *Essentials of teaching academic vocabulary*. Houghton Mifflin Co.
- Crismore, A., Markkanen, R., & Steffensen, M. S. (1993). Metadiscourse in Persuasive Writing: A Study of Texts Written by American and Finnish University Students. *Written Communication*, 10(1), 39–71. <https://doi.org/10.1177/074108839301001002>
- Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: the roles

- of cohesion and linguistic sophistication: PREDICTING L2 WRITING PROFICIENCY. *Journal of Research in Reading*, 35(2), 115–135. <https://doi.org/10.1111/j.1467-9817.2010.01449.x>
- Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2011). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing*, 28(4), 561–580. <https://doi.org/10.1177/0265532210378031>
- Dastjerdi, H. V., Shirzad, M., & Student, M. (2010). *The Impact of Explicit Instruction of Metadiscourse Markers on EFL Learners' Writing Performance*. <https://doi.org/10.22099/JTLS.2012.412>
- Duran, P., Malvern, D., Richards, B., & Chipere, N. (2004). Developmental Trends in Lexical Diversity, *Applied Linguistics*, 25(2), 220-242. <https://doi.org/10.1093/applin/25.2.220>
- Durrant, P. (2016). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes*, 43, 49–61. <https://doi.org/10.1016/j.esp.2016.01.004>
- Erarslan, A. (2021). Correlation between Metadiscourse, Lexical Complexity, Readability and Writing Performance in EFL University Students' Research-based Essays. *Shanlax International Journal of Education*, 9(S1-May), 238–254. <https://doi.org/10.34293/education.v9iS1-May.4017>
- Fergadiotis, G., Wright, H. H., & Green, S. B. (2015). Psychometric Evaluation of Lexical Diversity Indices: Assessing Length Effects. *Journal of Speech, Language, and Hearing Research*, 58(3), 840–852. [https://doi.org/10.1044/2015\\_JSL-HR-L-14-0280](https://doi.org/10.1044/2015_JSL-HR-L-14-0280)
- Foster, P., & Tavakoli, P. (2009). Native Speakers and Task Performance: Comparing Effects on Complexity, Fluency, and Lexical Diversity. *Language Learning*, 59(4), 866–896. <https://doi.org/10.1111/j.1467-9922.2009.00528.x>
- Granger, S., & Larsson, T. (2021). Is core vocabulary a friend or foe of academic writing? Single-word vs multi-word uses of thing. *Journal of English for Academic Purposes*, 52, 100999. <https://doi.org/10.1016/j.jeap.2021.100999>
- Gu, X., & Xu, Z. (2021). Sustainable Development of EFL Learners' Research Writing Competence and Their Identity Construction: Chinese Novice Writer-Researchers' Metadiscourse Use in English Research Articles. *Sustainability*, 13(17), 9523. <https://doi.org/10.3390/su13179523>
- Ha, H. S. (2019). Lexical Richness in EFL Undergraduate Students' Academic Writing. *ENGLISH TEACHING*, 74(3), 3–28. <https://doi.org/10.15858/engtea.74.3.201909.3>
- Howe, S., & Henriksson, K. (2007). *Phrasebook for writing papers and research in english: over 5000 words and phrases to help you write at university and research level in English* (4. ed). Whole World Company Press.
- Hyland, K. (2004). Disciplinary interactions: metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing*, 13(2), 133–151. <https://doi.org/10.1016/j.jslw.2004.02.001>
- Hyland, K. (2005). *Metadiscourse: exploring interaction in writing*. Continuum.

- Hyland, K. (2010). Metadiscourse: Mapping Interactions in Academic Writing. *Nordic Journal of English Studies*, 9(2), 125–143. <https://doi.org/10.35360/njes.220>
- Hyland, K. (2019). *Metadiscourse: exploring interaction in writing*. Bloomsbury Academic.
- Hyland, K., & Tse, P. (2007). Is There an “Academic Vocabulary”? *TESOL Quarterly*, 41(2), 235–253. <https://doi.org/10.1002/j.1545-7249.2007.tb00058.x>
- Jarvis, S., & Daller, H. (Eds.). (2013a). *Vocabulary knowledge: human ratings and automated measures*. John Benjamins Publishing Company.
- Jarvis, S., & Daller, H. (Eds.). (2013b). *Vocabulary knowledge: human ratings and automated measures*. John Benjamins Publishing Company.
- Kaya, F., & Sofu, H. (2020). Exploring Effects of Explicit Teaching of Metadiscourse Markers on EFL Students’ Writing Proficiency. *The Reading Matrix: An International Online Journal*, 20(2).
- Koizumi, R., & In’ami, Y. (2012). Effects of text length on lexical diversity measures: Using short texts with less than 200 tokens. *System*, 40(4), 554–564. <https://doi.org/10.1016/j.system.2012.10.012>
- Leńko-Szymańska, A. (2015). The English Vocabulary Profile as a benchmark for assigning levels to learner corpus data. In M. Callies & S. Götz (Eds.), *Studies in Corpus Linguistics* (Vol. 70, pp. 115–140). John Benjamins Publishing Company. <https://doi.org/10.1075/scl.70.05len>
- Lessard-Clouston, M. (2013). Word Lists for Vocabulary Learning and Teaching. *Undefined*. <https://www.semanticscholar.org/paper/Word-Lists-for-Vocabulary-Learning-and-Teaching.-Lessard-Clouston/8ee32f800d6ae7dd175527c88af4e9625d77f1d3>
- Mat Zali, M., Mohamad, R., Setia, R., Raja Baniamin, R. M., & Mohd Razlan, R. (2021). Comparisons of Interactive and Interactional Metadiscourse among Undergraduates. *Asian Journal of University Education*, 16(4), 21. <https://doi.org/10.24191/ajue.v16i4.11946>
- Matthews, P. H. (2014). *The concise Oxford dictionary of linguistics* (Third edition). Oxford University Press.
- McCarthy, P. M. (2005). An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity (MTLD) [Doctoral dissertation, The University of Memphis]. Retrieved from ProQuest Dissertations & Theses (p. 199485).
- McCarthy, P. M., & Jarvis, S. (2007). vocd: A theoretical and empirical evaluation. *Language Testing*, 24(4), 459–488. <https://doi.org/10.1177/0265532207080767>
- McDonough, K., Neumann, H., & Hubert-Smith, N. (2018). How Accurately do English for Academic Purposes Students use Academic Word List Words? *BC TEAL Journal*, 3(1), 77–89. <https://doi.org/10.14288/bctj.v3i1.293>
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic Features of Writing Quality. *Written Communication*, 27(1), 57–86. <https://doi.org/10.1177/0741088309351547>
- Mei Hooi, C., Tan, H., Lee, G. I., & Victor Danarajan, S. S. (2020). Texts with Metadiscourse Features

- are More Engaging: A Fact or A Myth? *3L The Southeast Asian Journal of English Language Studies*, 26(4), 58–73. <https://doi.org/10.17576/3L-2020-2604-05>
- Milanovic, M. (2009). Cambridge ESOL and the CEFR. *Research Notes*, (37) 2–5. Cambridge: Cambridge ESOL.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters.
- Morris, L., & Cobb, T. (2004a). Vocabulary profiles as predictors of the academic performance of Teaching English as a Second Language trainees. *System*, 32(1), 75–87. <https://doi.org/10.1016/j.system.2003.05.001>
- Morris, L., & Cobb, T. (2004b). Vocabulary profiles as predictors of the academic performance of Teaching English as a Second Language trainees. *System*, 32(1), 75–87. <https://doi.org/10.1016/j.system.2003.05.001>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (Second Edition). Cambridge University Press.
- Nation, I. S. P. (2016). *Making and using word lists for language learning and testing*. John Benjamins Publishing Company.
- Owen, N., Shrestha, P., & Bax, S. (2021). Researching lexical thresholds and lexical profiles across the common European framework of reference for languages (CEFR) levels assessed in the Aptis test. *ARAGs Research Reports Online, AR-G/2021(1)*. <https://www.britishcouncil.org/exam/aptis/research/publications/arags/researching-lexical-thresholds-and-lexical-profiles-across>
- Pavčić Takač, V. (2008). *Vocabulary learning strategies and foreign language acquisition*. Multilingual Matters.
- Radford, A. (Ed.). (2010). *Linguistics: an introduction* (2. ed., 3. print). Cambridge Univ. Press.
- Ransdell, S., & Wengelin, A. (2003). Socioeconomic and sociolinguistic predictors of children's L2 and L1 writing quality. *Arob@se*, 1-2, 22-29
- Reynolds, B. L. (2015). A Mixed-Methods Approach to Investigating First-and Second-Language Incidental Vocabulary Acquisition Through the Reading of Fiction. *Reading Research Quarterly*, 50(1), 111–127. <https://www.jstor.org/stable/43497208>
- Sanford, S. (2012). A Comparison of Metadiscourse Markers and Writing Quality in Adolescent Written Narratives. *Graduate Student Theses, Dissertations, & Professional Papers*. <https://scholarworks.umt.edu/etd/1366>
- Sešek, U. (2016). Revising and Metadiscourse in Advanced EFL/ESL Writing. *International Journal of Applied Linguistics and English Literature*, 5(3), 35–45. <https://doi.org/10.7575/aiac.ijalel.v.5n.3p.35>
- Stills, M. (2016). Language Sample Length Effects on Various Lexical Diversity Measures: An Analysis of Spanish Language Samples from Children. *University Honors Theses*. <https://doi.org/10.15760/honors.250>
- Sun, D. (2017a). A Contrastive Analysis between English Vocabulary Profile and College English Word-list. *Theory and Practice in Language Studies*, 7(9), 729. <https://doi.org/10.17507/tpls.0709.04>

- Sun, D. (2017b). The CEFR Stratification of English Productive Vocabulary of Chinese University Undergraduates Based on DIY Learner English Corpus. *Journal of Language Teaching and Research*, 8(5), 909. <https://doi.org/10.17507/jltr.0805.09>
- Vande Kopple, W. J. (1985). Some explanatory discourse on metadiscourse. *College Composition and Communication*, 36
- Vidal, K., & Jarvis, S. (2020). Effects of English-medium instruction on Spanish students' proficiency and lexical diversity in English. *Language Teaching Research*, 24(5), 568–587. <https://doi.org/10.1177/1362168818817945>
- Webb, S. A., & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford University Press.
- Yaghoubi, A., & Ardestani, S. (2014). *Explicit or Implicit Instruction of Metadiscourse Markers and Writing Skill Improvement*. <https://doi.org/10.7575/AIAC.IJALEL.V.3N.4P.14>
- Yüksel, H. G., & Kavanoz, S. (2018). Dimension of Experience: Metadiscourse in the Texts of Novice Non-Native, Novice Native and Expert Native Speaker. *Advances in Language and Literary Studies*, 9(3), 104–112. <https://doi.org/10.7575/aiac.all.s.v.9n.3p.104>