

Lexical bundles in academic writing among ESL Law students in Ghana

Ivy Jones-Mensah¹, Michael Owusu Tabiri^{2*}, Patience Calys-Tagoe³,
Joseph Kwame Ocloo⁴, Dickson Amexo⁵

¹Department of Communication Studies

University of Professional Studies-Accra & Department of Applied Linguistics,
University of Education, Winneba, Ghana

²Language Centre/Department of General Studies

Ghana Communication Technology University & Department of Applied Linguistics,
University of Education, Winneba, Ghana

^{3,4}Department of Communication Studies, University of Professional Studies, Accra, Ghana

⁵Department of Language and Communication Skills, University of Media,
Arts and Communication, Accra, Ghana

ABSTRACT

The purpose of this study is to explore the complexity of the lexical bundles and/or formulaic patterns in Law texts, create a corpus of authentic formulaic patterns of verb forms in law and propose a workable method for identifying and teaching the accumulated specialised registers of formulaic patterns of law, the types and function of law lexical and formulaic bundles in the academic writing classroom. The study adopts corpus linguistic approach to investigate the patterns in the Law students' texts. The study used 792,237 corpus size. The routine of the types and functions of formulaic patterns and the multiword in Law subject areas found on a virtual learning platform were analysed. This procedure revealed brief prevailing six-word lexical formulaic patterns from Law texts. The discussion from the outcome of this search for extensively used formulaic patterns in Law leads to a consideration of challenges in formulaic pattern or multiword patterns data into English for academic writing in the law discipline in tertiary institutions. The challenges lead to suggestions for the instructors in the law faculties to engage their students in working with word lists made up of formulaic categorisations in law. The study revises literature building on prior study to contribute to more understanding of law bundles. It introduces a novel approach using Sketch Engine to analyse multiword patterns in legal texts, offering theoretical and methodological insights that support specialised vocabulary development and future research in legal language registers. The study provides formulaic patterns that are particularly useful for lecturers and Law students in the academic writing classroom in the Law faculties.

Keywords: corpus linguistics, lexical bundles, formulaic patterns, multi word patterns

INTRODUCTION

The study generally aims to explore lexical bundles from Law students' texts. The study has specifically explored the complexity of the lexical bundles in Law texts, it creates an authentic formulaic patterns in law corpus with the Sketch Engine and proposes a workable method for identifying and teaching the accumulated specialised vocabulary, registers of formulaic patterns of law in the academic writing classroom.

A text, according to Bryd and Coxhead (2010, p. 32) is defined by its recurrent combinations. These combinations constitute formulaic patterns, and the patterns could occur anywhere in the text from a single word to several words. At least in every formulaic pattern, frames like words, phrases, clauses, expressions and other groups of neighboring words that collocate are included. In fact, the formulaic patterns are prevalent in law (Breeze 2013). When words combine in two or more sequence of words in a repeated chain without the interruption of any adverbial in the set of the number of times

that word occurs in a specific corpus, the combined words are known as lexical bundle or bundles, according to Hajizaded et al., (2018), and Smith (2021).

According to Smith (2021), lexical bundles must appear frequently throughout the corpus of texts that make up

Corresponding Author e-mail: mtabiri@gctu.edu.gh

https://orcid.org/0000-0002-4202-5156

How to cite this article: Jones-Mensah I, Tabiri MO, Calys-Tagoe P, Ocloo JK, Amexo D. Lexical bundles in academic writing among ESL law students in Ghana, Pegem Journal of Education and Instruction, Vol. 15, No. 2, 2025, 17-29

Source of support: Nil

Conflicts of Interest: None.

DOI: 10.47750/pegegog.15.02.03

Received: 03.08.2024

Accepted: 18.12.2024

Published: 01.04.2025

the corpus for them to exist. Smith (2021) explained that the recurrence of sets of words used by a specific speech community must be encouraged rather than the use of sets of words that are characterised in a particular book by a particular writer. The high frequency of occurrence of the sets of bundles is more indicative of formulaic patterns rather than only lexical patterns, according to Hajizaded et al., (2018). It means that rather than finding and reporting on the statistical significance of bundle occurrences, the examination of the formulaic pattern necessitates that the lexical bundles report high frequency of co-occurrences. In the case of formulaic patterns, Biber et al. (1999) and Leech and Rayson (2014) preferred a range of 10 repeated occurrences of words, but Cortes, (2004) and Hyland, (2008) advocated 20 to 40 times repeated occurrences in every million words in the Law academic articles. In this study, 50 times occurrences per million words is preferred.

Ellis (2012) and Hinkel (2005) have made distinctions between different types of texts, including newspaper discourse, conversational language, creative writing, academic writing, including language for individual variances. Granger (2018); Meunier (2012); Yakut et al. (2021) did a comparative work of bundles in published literature and students' project works and dissertations in other disciplines, which is another area of distinctive bundle's analysis in corpus search. The findings from Ellis (2012); Granger (2018); and Meunier (2012) showed that vocabulary has a major role in defining the distinctions in discourse patterns among the fields.

Researchers, Choi and Zhang, (2021); Liu (2013) and Shintani (2012) found that vocabulary play a large role in defining the differences in discourse types than grammar. They came to understand that certain research papers have different introduction chapters because they use varied and distinctive vocabulary depending on the discipline. Two writings could have some similarities even while their distinctions are clearly distinct. Nevertheless, one can tell if they belong to the same discipline or to another by looking for recurring word patterns. The goal of the introduction sections of texts in one field for example Law, Marketing, and the introduction sections of texts in other disciplines (i. e. History, Science, Management, Communication, Business Communication), is something that is discovered throughout the vocabulary search journey because these disciplines distinguish themselves in the use of vocabulary in their writings.

According to Lee and Macaro (2013), the commonality arises from the shared objectives of the entrance point level, student fundamentals, and language targeted at a specific class or age level. An important task for the applied corpus linguist is to gather and pair language patterns that are common to

various academic disciplines as well as pairs that are restricted to use in particular penal contexts (Byrd & Coxhead, 2010; Rees, 2018). For this reason, this study investigates the paired vocabulary in the texts of level 200 Law students.

It is necessary to investigate lexical bundles in discipline specifics because Ganger and Meunier (2008); Durrant (2017) declare that there is little research done to investigate lexical bundles in the various disciplines, especially, texts of undergraduate students pursuing Law. Martínez (2024) declare that the discipline lexical bundles occur in complex syntactic patterns and there is the need to study the complexity of the phraseological patterns of these bundles. Martínez (2024, p.252) posits that "there is currently little research into L2 writing in ... contexts". According to Ganger and Meunier (2008), Golparvar and Barabadi, E. (2020). it is crucial to investigate the phraseological approach to teaching and learning English in the many disciplines because it is the main problem ESL lecturers and students encounter. Therefore, more empirical research is required to determine the precise impact of lexical patterns (phraseological patterns) in vocabulary teaching and learning. According to Granger and Meunier (2008), the ability to identify and produce such patterns has become unrelenting challenge for tertiary students in the law disciplines. These students find it difficult to develop both fluency usage and appropriate usage of these patterns in their discipline (Breeze, 2013; Gardner, 2021; Gilquin et al., 2007).

Further, the argument is that lexical bundles have been used to investigate varieties of texts and speeches but not much of these investigations pay much attention to formulaic patterns and multi-word patterns in Law. Castellano-Risco (2024) declares that students enroll in programs frequently to encounter difficulties comprehending lectures primarily due to their restricted vocabulary knowledge they have in the second language. To her, to address the issue it is intensifying knowledge in lexical study that can help solve the problem. Many Law students in Ghana have problems comprehending law lexical structure (Law 2022). Castellano-Risco (2024); Vuković-Stamatović and Čarapić (2022) confirmed that the problem with lexical knowledge is evident in development of writing skills, they continued that students need lexical density because the amount of specialised academic and scientific vocabulary the students need for composing academic and discipline specific text is enormous. Vuković-Stamatović and Čarapić (2022) agreed that students find vocabulary knowledge appealing and motivating therefore investigating vocabulary will increase the teaching and learning of lexical density. They noted, "the measures of lexical density are generally well-established for writing (Vuković-Stamatović & Čarapić, 2022).

Research shows that identifying the effectiveness of the phraseological approach to teaching and acquiring vocabulary contextually in English for specific purposes is difficult because there are only a limited number of bundles in the literature of ESL studies. Often, researchers look at high-frequency multiword sets rather than word bundles (Hajizaded, et al., 2018; Smith, 2021). Modern research often looks at word bundles rather than frequency occurrences. This investigation of formulaic patterns and multiword patterns in the writings of Law students were motivated by the current fashion of using the phraseological approach. Vuković-Stamatović and Čarapić (2022, p. 210) confirm that the use of phraseological patterns enhances “cognitive engagement and higher-order thinking” in students and could lead to improved teaching and learning of “scientific English” for the purpose of this study discipline subjects.

Additionally, there is no room for deciding how word patterns must be repeated as a guiding principle for teaching and learning discipline specific vocabulary. The idea of finding shorter bundles from longer ones and extracting the shorter ones out of the longer ones becomes cumbersome (Bryd & Coxhead, 2000). If a shorter phrase such as “it is based on”, occurs in some of the repeated phrases, what kind of classifications will be given to such bundles? Coxhead (2010) and Akmajian et al. (2017) enquired if such repeated forms will be classified as spoken languages than written texts. Students and their instructors need information on the use of formulaic patterns from their discipline texts to be able to compose accurate academic texts. Further, it is important to identify that some word patterns are not necessary for such texts; rather replacing these bundles with a word might be more formal and acceptable in the academic environment. Martínez (2024) noted that “native speakers expand their vocabulary voluntarily by about 1,000 words” (p. 219) but this is not restricted to special scientific vocabulary. However, the ESL student in a special discipline like law may need much time to build his lexical density.

The issue is complicated for the ESL tertiary institutions because there is complete silence regarding the necessity of exposing students to discipline-specific vocabulary studies and vocabulary understanding in general. It is common for some professors to insist on knowing the type of vocabulary to teach their students and to find out whether teaching vocabulary and how the various words mix are crucial in tertiary institutions (Authors, 2022). There is still more work to be done to improve vocabulary and vocabulary knowledge among teachers and their students. The study of instruction of these complex discipline-specific repetitive forms are crucial. The approach to teaching these patterns, as well as teaching repeated forms, are still difficult, according to

(Coxhead, 2008; Tripto et al., 2018). If determining bundles in undergraduate students’ texts continues to be difficult, then analysing the issue and determining the accuracy of the patterns is very important.

What is the significance of this study? The study contributes to our understanding of how discipline texts accumulate specialised vocabulary, registers of formulaic patterns, and multi-word phrases. It also shows that the authors are familiar with the textual styles of particular disciplines’ patterns. This paper offers information on how to use corpus techniques to learn more about the mysterious word patterns that recur frequently in legal essays. It demonstrates the intricacy of formulaic language used in some different legal contexts like International Law, Human Rights Law, Business Law, and Administrative Law. The most frequent word combinations, formulaic patterns, and multiword patterns are noted, and their functions within the text are investigated. It is important to identify the issues with bundles and create a lasting solution to the problem. The study considers revelations about types of lexical bundles among ESL Law students’ writings, identify the functions and proposes how one can work with the bundles in this discipline. Specifically, the objectives are broken down into the following units.

- i. to investigate the types of formulaic patterns in the law students’ texts in new methodological perspectives.
- ii. to identify the authentic functions of the formulaic patterns and multi-patterns and how these contribute to literature.

To achieve these objectives, the authors pose the following questions:

- i. What types of formulaic patterns are found in the law students’ texts in a new methodological way?
- ii. What are the authentic functions of the formulaic patterns and multi-patterns and how do these contribute to literature?

LITERATURE REVIEW

English as the language of Law is studied either professionally or academically. Thus, there is the pervasive use of English in Law. Though English is universal, certain forms of the vocabulary and expressions used in the study of law are specific to this field, making the language technical. What many academics refer to as English for Specific Purpose (ESP), which is need-oriented, topic-oriented with purpose-driven attention and use, include Chalikandy (2013); Kim (2013), and Lockwood (2012). This suggests that several things play a role in the way this type of English is used because ESP differs from other English usage in a number of ways.

Various scholars across different fields and with various titles have recognised certain elements of English tailored for Specific Purposes. It is called English for Academic Purpose (EAP) since it is a subset of ESP, according to Nartey and Dorgbetor (2014); Castellano-Risco (2024). While Chalikandy (2013) asserts that ESP has grown to be a significant component of English for students who learn English as a second language. According to Chalikandy (2013), ESP is not driven by any theory, ideals, or philosophy; rather, it begins with the necessity of the time. It is merely because it is pertinent to Applied Linguistics and the numerous academic fields of law, commerce, information technology, and other contemporary issues.

Hutchinson and Waters (1987); James (2013) and Véliz (2023) agree on the emergence of English used for discipline purposes such as Law, is a call for a heroic new world and an upheaval period in linguistics. The primary focus of the ESL teachers evolves around addressing students' requirements and guiding them in effectively employing language to fulfil those needs. They contend that because English was the language of knowledge, Applied Linguistics must adapt its strategy, resources, and methods towards teaching specific English for the disciplines' sake to meet the needs and requirements of its stakeholders, including the teacher, students, and other parties. English has evolved into a topic that addresses the desires of all and not just those in academia (Hall et al., 2017).

The existence of special English for academic purposes is about the fundamental presumption that academic writing before entering a university is fundamentally different from writing at the university level, according to Afful (2007); Alfes and Dison (2000); Jones-Mensah et al. (2022). For example, Evans and Green (2007); Hadley (2014) posited that the English language demands of tertiary students go beyond just requiring language help at universities and should instead focus on academic disciplines. Hyland and Milton (1997) observed that students in general across all faculties view their ability in English courses as a key factor in determining their academic achievement in a range of academic settings. In reality, to become proficient in this type of English, a learner must be aware of the vocabulary knowledge in that discipline.

Research into various features of English language use in various fields is prompted by the awareness that various faculties place a high priority on success in English-language courses. English for special disciplines is marketed under variety of names across the world. Afful (2007) claims that special English for disciplines writing is used everywhere including native speakers in the United States, the United Kingdom, and Canada. In Ghana, Academic Writing, or Advanced Academic Writing, is the norm. The promotion

of English for disciplines writing classroom resulted from the increased creativity in establishing new English courses for tertiary institutions to meet the current trend of constant change in educational systems (Jordan, 2002). It is impossible to overstate the value of English in various professions. It is impossible to overstate English's significance for some subjects because it fosters central, seamless, and catalytic academic communication for students in tertiary institutions.

According to Jordan (2002) and Rose et al. (2022), the promotion of disciplines' English in the academic classroom has resulted in the realisation and the internalisation of advanced communication rules in English in the tertiary education. This is unquestionably crucial for enhancing and expanding vocabulary knowledge in the many discipline-specific studies. Adika (2019) and Fadda (2011) contend that undergraduate writing programs must be implemented to enhance the internalisation of corpus knowledge in tertiary education. The use of lexical bundles, formulaic phrases, and many terms that are common in the writing of a specific academic field rather than the use of specialised language or technical terminology reveals vocabulary competence, according to research (Baker, 2014 & Biber et al., 1999; Rodrigo, 2023). Rodrigo (2023, p. 329) continued that these bundles do not only reveal vocabulary competence but "highly frequent bundles frame semantic meanings" and show "discourse functions that they perform". These packages may be essential for gaining acceptance into a certain discipline's community.

Several institutions in Ghana and elsewhere consider the instruction and the exploration of specialised discipline writing especially, at tertiary institutions to be highly significant. Writing for a particular discipline is necessary for the establishment of a shared academic language for that discipline. This is "academic lingua franca", as Duszak (1997, p. 21) recognised it. Because it ensures both graduate and undergraduate students' easy transition into the academic community, the significance of these lexical, formulaic, and multi-word patterns in vocabulary study in EAP cannot be overstated (Afful, 2007).

Afful (2007) and Afzal (2019) complained that the lack of attention given to study skills programmes that provide university students with various vocabulary knowledge in the form of lexis, formulaic, and multiword in the various disciplines in academic writing skills is alarming. The courses are advertised as being quite inconsequential; some people refer to them as remedial, adjunct, and periphery courses (Afful, 2007; Afzal, 2019). According to Zwiers (2014), a significant portion of the common core element for teaching academic study skills is equipping the student with a general academic register in English, formal academic writing style,

and academic language proficiency. Writing programmes at all levels serves special and specific needs for tertiary students, especially it serves the need of equipping students with vocabulary skills.

MATERIALS AND METHODS

Research approach and design

This study uses a corpus approach. The ESL Law Corpus was developed using written academic texts from 100 students in second year in the Law School. In all **792, 237 corpus** size of was used. Repetition of formulaic and lengthy word groups in the corpus size were gathered using the Sketch Engine and compared against British English. The corpus Sketch Engine's tool was able to find all the bundles of a specific length in a selection of texts. The software reported on the set of frequencies of 2, 3, 4, 5, and 6 bundles found in the beginning Law students' texts. After that, a cut of point of 50 pieces texts for each of the levels were found in the *N-grams*. These is manually determined based on the manageable amount of data being used. Coxhead (2010) asserts that to create valid data; a cutoff threshold must be selected based on what previous studies found in a related corpus search. As a result, the cutoff point of 50 bundles is calculated per million thresholds that was established in the *N-gram* (Paquot & Granger, 2007; Breeze, 2013). Each *N gram* consists of roughly 13,000-20,000 items derived from corpus 792,237 words which was normalized to 1 million (total) words. Per each search

Each of the N-gram is assigned a cutoff value of 50 bundles. There were 100 texts in the corpus. All disciplinary sub-areas of the entire corpus' topic areas are listed in Table 1. As much as feasible, a balance between the numbers of

texts (13,000 to 20,000 running items) were first attempted when choosing the texts for the corpus. The second-year law students' essays in the academic writing class as well as other similar legal educational systems who study academic writing were selected because they represent the main fields of study for undergraduate law students in Ghana. For instance, subjects like equity and succession law, administration of justice, and constitutional law were considered aside the academic writing texts. The total tokens, words, sentences and document counts for the Law corpora are stated in the table that follows in Table 1.

The *N-grams* as in the language bundles were collected from the general corpora as demonstrated in table 2. The percentage of the count information from the texts per *N-gram* is distributed in the table 2.

The corpus contained only texts from students from the Law faculty in the chosen university. The corpus shows how Law students use English. Some sections on the British English revised edition were used to ascertain the value of use of Law diction among ESL students. The British corpus was included to have a true corpus of the ESL students' English. Hyland and Tse (2012) recommended that the true value of the bundles occur in sequences that are formulaic, data reveals that these bundles are just prolonged collocations that happen to occur often than predictable by causal occurrence. The sequence of bundles helps the writer to present meanings in particular situations. Additionally, the sequence adds to our perception of the coherency in text. ConcGram with lexical and situational varieties of words have received more attention in recent years. Together, these lexical patterns show distinction between beginning users and proficient users across a variety of genres and these two types of varieties are widely used in academic language development.

About 6% of the entire written Law academic embodiment consisted of portions from Law academic corpora. The word counts in the study were developed using the N-gram count from the Sketch Engine (authors' [Dashboard Sketch Engine](#)). Coxhead (2000) and Heatley et al. (2002), developed the original academic lists of words using the ranges in which the words occur. They used word ranges of words in a bundle

Table 1: Level 200 counts information by structure.

Tokens	475,290
Words	426,642
Sentences	15,282
Documents	100

Table 2: Level 200 count information by N-gram

Level 200	Tokens	Words	Frequency per N-gram	Total frequency count info	Percentage frequency
N-gram 2	10552	19887	325072	475,290	68.3944
N-gram 3	10552	24751	219048	475,290	46.087
N-gram 4	10552	21161	147045	475,290	30.937
N-gram 5	10552	17124	108144	475,290	22.753
N-gram 6	10552	14262	85895	475,290	18.072

instead of grams in a bundle. This study is different in terms of extracting the bundles in that it uses N-gram bundles from the Sketch Engine than extracting ranges of words in a text as it occurred in the study by Coxhead (2000) and Heatley et al (2002).

Population

The population is 300 students from the Faculty of Law in one of the universities in Ghana. Only level 200 students are considered. Only level 200 students were considered because they have been exposed to much academic writing exercises and most of their subjects are electives (Law courses) so it is likely these students would be using discipline-specific language sufficiently for the researchers to derive accurate discipline bundle from their texts. Each class in level 200 has 100 texts only which were selected from each discipline purposively. As a result, only scripts that contain detailed write-ups were collected purposively and conveniently used for this research.

Sampling technique

A purposive sampling technique was used to identify the Faculty of Law. Texts were collected conveniently. Thus, only texts with copious law prose were chosen. Table 3 below illustrates the text data by frequency and word-count.

There are different areas of courses that represent several different crucial areas of studies for pre-tertiary students in the university. Different levels were considered for the ESL to have vast corpora made up of seven subjects contained in about 792,237 running words.

RESULTS AND DISCUSSIONS

The analysis in this study begins with the creation and the presentation of a list of statistical counts in the graphs. A description was drawn below the statistical representation from each graph. Additionally, analysis continued with the presentation of a graph showing the use of the lexical bundles for corpus of Law students' texts from some subject areas in law. The regular occurrence of each word bundle in law texts of 6 different *N-grams* (pairs) was collected classified, and analysed, trying to seek patterns of similarity and difference

across the *N-grams*. The data for the frequency of the lexical bundles were standardised for a million words because of the small variations in the size of the texts chosen for this study. The results were discussed in the next sub-heading.

The 50 pieces of bundles and multi-words were transferred into an excel sheet. The hits of distribution of the bundles were drawn from the excel sheet in a hit chart. In different bundle groupings. Each bundle distribution on the bar graphs has different values. In each graph, there are different values based on the total value in each bundle. The first bundle worth discussing is the 2 *N-gram* bundle.

The N-bale / N-gram bundles

N-gram is a kind of tokenization. It is a kind of **compendium suitable for building N-(tokenizing)**, as well as to produce original content by leveraging the *N-gram* structure of a provided text for text producing new texts grounded in the n-gram structure to create a new text content by utilising structures inherent in the input texts thus (babbling). The compendium has used the analysis for creating “bots” for specific subject language learning. In this study, the n-gram is used to check the probability of the adjoining categorisation of *N- bits items* of the Law corpora. The *N-bit* items here are words or base duos. This is according to the application used for this study. The findings here are in line with the revelation identified by Byrd and Coxhead (2010). The figure below illustrates the collection of 2 *N-gram*.

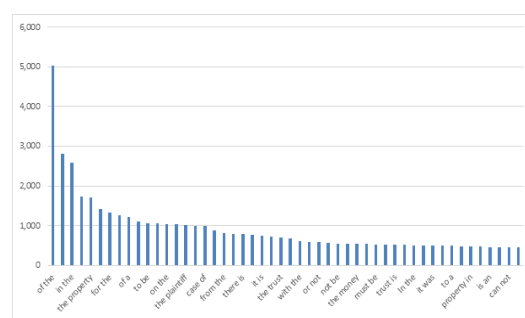


Fig. 1: Percentage of total words in lexical bundles by 2 *N-grams*.

https://app.sketchengine.eu/#ngrams?corpname=user%2FIvyjones%2Flaw_level_200&tab=advanced&wlimit=freq=2&ngrams_n=2&ngrams_max

Source: Authors' construct 2024.

Table 3: Level 200 text data by frequency and word-count.

Year Groups	Items	Total Frequency	Reference Item	Reference Frequency	Words
Level 200 A	7,368	75,383	17,124	108,144	210,869
Level 200 B	9,085	97,937	21,105	137,892	328,238
Level 200 C	7,620	67,909	24,956	284,886	253,130
Total	24,073	241,229	63,185	530,922	792,237

In the graph above, the bundle for the 2 *N-gram* for the bundle, *of the*, is the highest used bundle in the texts of Law. It occurred 83.3 % times while the next bundle is, *to the*, which occurred 46.8%. The next bundle obtained 42.8% for the bundle, *in the*, and 12 successive bundles after the bundle *in the* have a distribution of 28% to 16%. The rest have an insignificant percentage of occurrences, thus from 15% to 7.4%.

In this *N-gram*, the bundles are of universal occurrences in other texts however, the bundle that truly depicts law bundles are *the property*, *the plaintiff* and, *property in*. Unfortunately, these bundles have the lowest occurrences, and these have not even shown any even distribution patterns rather it shows the pattern of a decline in use. The probability of selecting one over the other is 4. Thus, after every four words one of the 2 *N-gram* law patterns is likely to occur. Another possibility found in the text is the occurrence of 3 *N-grams* as in figure 2.

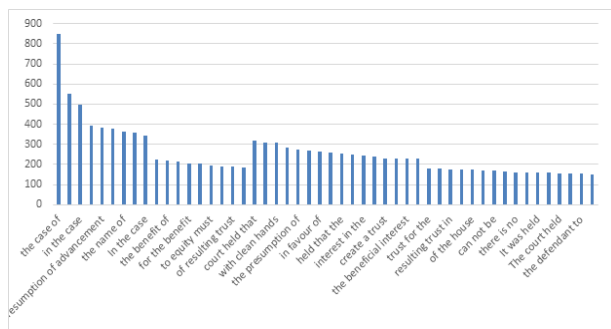


Fig. 2: Percentage of total words in lexical bundles by 3 *N-grams*.

https://app.sketchengine.eu/#ngrams?corpname=user%2Fivyjones%2Flaw_level_200&tab=advanced&wminfreq=2&ngrams_max_n=3&showres

Source: Authors' construct, 2024.

In the graph above, the bundle for the 3-*Ngram* for the bundle *the case of* is the highest used bundle in the texts of Law. It occurred 94.5 % time while the next bundle is *a resulting trust*, which occurred 61.3 %. The next bundle obtained 55.2% for the bundle, *in the case*, and about 10 successive bundles after the bundle *in the case* has a distribution of 43.3 % to 31.3%. The rest have occurred between the percentages of 25.1% to 16.7 %.

The 3 *N-gram* carries a chunk of bundles suitable for Law texts. Only 3 pieces of the bundles appear universal. For example, *the name of*, *cannot be* and *there is no*. The rest are very relevant to the area of study. Beside the three pieces of bundles, four pieces were found.

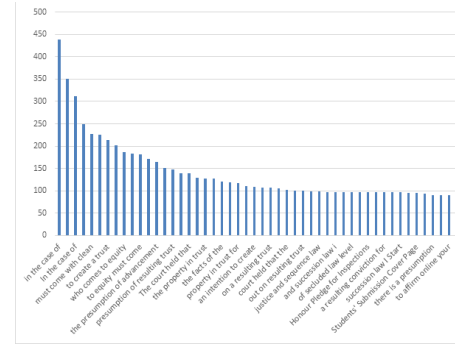


Fig. 3: Percentage of total words in lexical bundles by 4 *N-grams*.

https://app.sketchengine.eu/#ngrams?corpname=user%2Fivyjones%2Flaw_level_200&tab=advanced&wminfreq=2&ngrams_n=6&ngrams_max
Source: Authors' construct 2024.

In the graph above, the bundle for the 4 *N-gram* for the bundle *in the case of* is the highest-used bundle in the texts of Law. It occurred 87.8 % time while the next bundle is *in the name of* occurred 70.2 %. The next bundle obtained 62.2% for the bundle, *in the case of*, and about 10 successive bundles after the bundle *in the case of* have a distribution of 49.8 % to 30.4%. The rest have occurred between the percentages of 28.1% to 18 %. The 4 *N-gram* pattern is strictly Law-patterned language.

However, the lexical bundles that are shared across the levels do not exist in equal occurrences of the disciplines. The graph reduces the list to those bundles in a reasonable distribution across 3 levels by choosing the bundles that consist of at least 10% of each level. An example is, *in case of*, occurs many times in repeated succession in almost all the *N-grams*. This procedure decreased the number of communal bundles to a minimum in *N-gram* 2 and maximum of 6 *N-grams*, but it is exceedingly likely to find in all four levels in the ESL Law corpus as the graph indicates. In the graph below, the illustration shows the occurrence of five pieces.

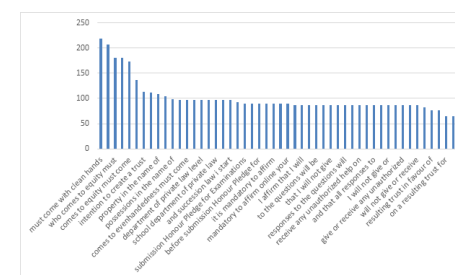


Fig. 4: Percentage of total words in lexical bundles by 5 *N-grams*.

https://app.sketchengine.eu/#ngrams?corpname=user%2Fivyjones%2Flaw_level_200&tab=advanced&wminfreq=2&ngrams_n=5&ngrams_max1/1
Source: Authors' construct 2024.

In the graph above, the bundle for the 5 *N*-grams for the bundle *must come with clean hands* is the highest used bundle in the texts of Law. It occurred 87.6 % time while the next bundle is *equity must come with, to equity must come with, comes to equity must come, and he who comes to equity* which occurred 82.2 %. The next bundle obtained 54.2% for the bundle, *intention to create a trust*, and about 10 successive bundles after the bundle *an intention to create* had a distribution of 34.8 % to 31.3%. The rest have occurred between the percentages of 30% to 16%

The analysis supports the finding that limited use of verbs is shown in sequence in lexical bundles. Dominantly, the lexical verb found on the slant is the verb *come*: for example, *comes to equity*, *must come*, and *he who comes to equity*. Additionally, future markers of the modal verb, *will* is worth studying among Law students because this has also dominated the formulaic patterns for the law diction in *N*-gram 5. There is further revelation in *N*-grams 6 as the graph below indicates.

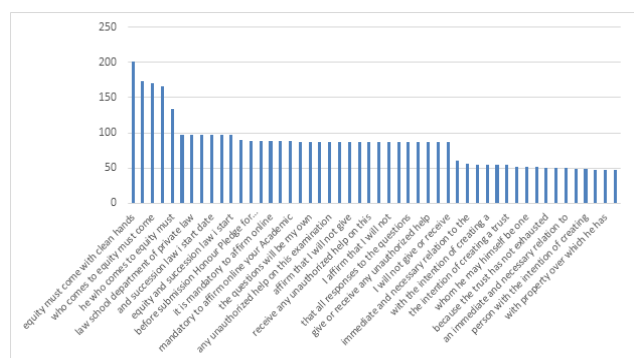


Fig. 5: Percentage of total words in lexical bundles by 6 *N*-grams

https://app.sketchengine.eu/#ngrams?corpname=us-er%2Fivjones%2Faw_level_200&tab=advanced&wlmfref-q=2&ngrams_n=6&ngrams_max1/1

Source: Authors' construct, 2024.

In the graph above, the bundle for the 6 *N*-grams for the bundle *equity must come with clean hands* is the highest used bundle in the texts of Law. It occurred 80 % time while the next bundle is *who comes to equity must come* which occurred 69.2 %. The next bundle obtained 35.6% for the bundle, *to affirm online your academic integrity*, and about 10 successive bundles after the bundle *in the case* has a distribution of 32 % to 19.2%.

Importance of Lexical Bundles in Law

The use of the active voice is heavily present in this bundle, because, there are verb forms that depict the use of active voice in the legal language. No doubt Hyland (2008) and Byrd and

Coxhead (2008) agreed that passive voice is more of scientific language than academic-specific disciplinary language. This also supports what Conrad (2008) revelation that passive voice is automatically not exceedingly occurring form in all-Law texts in among the tertiary academic writings. A shared list that combines lexical bundles of all the levels in academic Law study is possible to lessen the significance of operational structures, which are more distinguishing feature of one of the disciplines' levels than the others are. The accuracy, objectivity, and formal tone necessary that are necessary for legal communication (Cravens, 2024). He confirms that these helps enhance by the key linguistic tool of passive voice usage in ESL legal English. Cravens (2024) further reveals that in legal language, passive constructions are popular, especially in lexical bundles, which are predetermined word sequences that are frequently, used together in particular circumstances. It is important to pay attention to the passive voice in teaching academic writing to Law students to enable them to learn to comprehend the use and significance of passive voice in these lexical bundles.

Instead of emphasising the subject of the verb in legal constructions, passive voice highlights the action or the object of the action. This change in emphasis is very helpful in legal documents because when the action being performed, or the condition being described must be the focus rather than the person carrying out the action (Frese & Sabini, 2021). For instance, when a sentence like "an unauthorised help on this examination affirm ..." shows how the passive voice draws the focus not on the parties involved but rather on the action's completion and legal meaning (see *N*-gram 5, line 11). This implies that an action-oriented approach is consistent with the legal writing for accuracy and lucidity.

Furthermore, the formal and impersonal tone of legal language is enhanced using the passive voice (Garner, 2023). This is the reason in the book; *Law and objectivity*, Greenawalt (1995) confirms that legal documents frequently seek to convey facts in an impartial and objective way, avoiding the appearance of subjectivity or subjective opinions. Legal writers might continue to use this impersonal tone by employing passive constructions as a means of achieving formality in law. As an illustration, the sentence "property in the name of possessing ... comes even handedness ..." (see *N*-gram 4 lines 10 and 12) is more neutral and formal than the active voice substitute "He brings property in the name of ..." This show how impartiality and objectivity are critical for establishing legal circumstances and neutrality.

Moreover, the graphs show how passive voice in lexical bundles can enhance the clarity and readability of legal texts by structuring information in a way that aligns with legal conventions. For example, that all responses on the

question “*has been given*” or “*is receiving* any unauthorised help” (see N-gram 5 lines 15 and 16). At times, it is interesting to see how the passive marker *is*, is hidden in the process of communication. However, the passive voice is clearly used in the context of Law as shown in the illustration. This agrees with Garner’s (2023) revelation that passive voice is used in disguise to align with the legal conventions.

Legal documents are designed to be precise and unambiguous, often following established patterns that are easily recognizable to those familiar with the law. Passive constructions fit well into these patterns, making the documents more predictable and easier to navigate. For instance, “it must be” and “be resulting” (see N-gram 2 and 3) are phrases that lawyers and judges expect to see, helping them quickly understand the document’s intent as indicated by Kalen (2022).

Furthermore, Eliazarian (2024) observes that the inclusion of the deviant passive voice allows for the inclusion of necessary legal details without overloading the sentence with information about the agent performing the action. This can be particularly useful in complex legal scenarios where multiple parties and actions are involved. By using the passive voice, legal writers can focus on the legal principles and outcomes without the need to specify every agent involved, thus avoiding unnecessary complexity. For example, “the trust has not *been* exhausted” (N-gram 5 line 22) succinctly conveys the legal outcome without needing to specify who found the defendant guilty in that clause.

Consequently, it is revealed that the passive voice is instrumental in maintaining consistency and cohesion in legal documents. This kind of writing often involves referring to previous actions, conditions, or states in a consistent manner throughout the text which is shown in all the extracts in the grams. Passive constructions probably maintain this consistency by allowing actions and conditions to be described in a liveried way. For example, using “has been with clean hands, must come with clean hands” throughout grams ensures that the focus remains on the established fact, irrespective of the varying subjects involved. This consistency aids in the text’s overall coherence, making it easier for readers to follow and understand the legal arguments presented. Passive voice plays a significant role in the lexical bundles of ESL legal English by emphasising actions over agents, maintaining a formal and impersonal tone, enhancing clarity and readability simplifying complex information, and ensuring consistency. These attributes are essential for the precision, objectivity, and effectiveness required in legal communication. As such, we recommend that Law students must have mastery of passive constructions indispensable for legal professionals.

Importance of *to-infinitive* in the Lexical Bundles of Law

The integration of English as a Second Language (ESL) instruction tailored specifically for law is vital for non-native English speakers pursuing legal studies and careers. This specialized focus not only enhances general language proficiency but also equips students with the specific terminology and linguistic structures necessary for understanding and practicing law. One critical aspect of this instruction is the use of the *to-infinitive* form, which serves as a foundational tool for developing the linguistic patterns essential in legal contexts. The *to-infinitive* form in English, which consists of “to” followed by the base form of a verb, is instrumental in constructing clear and precise legal language. In legal writing and speaking, precision and clarity are paramount, as they help avoid ambiguities that could lead to misunderstandings or misinterpretations. The *to-infinitive* allows law students to construct sentences that clearly express intentions, obligations, and stipulations, which are frequently encountered in legal documents and discourse. By mastering this form, ESL law students can enhance their ability to draft legal texts that are both accurate and effective.

The construction of legal arguments and the presentation of cases both heavily rely on the *to-infinitive* form. *To-infinitive* constructions can be a useful tool for lawyers who frequently need to articulate their goals, lay out their cases, and define necessary activities. For instance, terms like “to argue the case,” “to establish the facts,” and “to comply with regulations” are frequently used in legal terminology. Students who receive ESL training that places a strong emphasis on the *to-infinitive* become more skilled at creating these and related expressions, which enhances their capacity to explain complicated legal concepts clearly.

Furthermore, the *to-infinitive* is integral to understanding and using formulaic patterns in legal language. Legal English is replete with set phrases and standardized expressions that convey specific meanings and legal concepts. These patterns often rely on the *to-infinitive* for their structure, such as “to have and to hold” or “to cease and desist.” By familiarizing themselves with these patterns, ESL law students can better comprehend legal texts and produce language that adheres to the conventions of legal discourse. This familiarity not only aids in comprehension but also in the practical application of legal principles in written and spoken communication.

In fact, incorporating the *to-infinitive* in ESL instruction for law students supports their transition from academic settings to professional environments. Transitional grades, where students move from theoretical studies to practical applications, require a solid grasp of the language tools that will be used in their professional careers. The *to-infinitive*,

with its versatility and frequent usage in legal contexts, is one such tool. By mastering its use, students are better prepared to engage in legal practices, whether in drafting documents, presenting cases, or negotiating terms. This preparation is crucial for their success as legal professionals in English-speaking environments.

The bales of bundles have *to infinitive* in most cases as in the prepositional phrase *mandatory to come*. This suggests that in the study of ESL English for Law, the *to-infinitive* is an important tool for developing the Law bundles and formulaic patterns at transitional grades to the advanced levels and the formulaic bundles move beyond the existing adverbial form that are old-style forms in law text of second language users. Lessons in the academic writing classroom of law students could focus on the teaching of the use of adverbs and adverbials like ‘*in the case of* or *on the basis of*’ which are possible forms that can occur in the disciplinary areas like Law

The ESL English for Law that emphasizes the *to-infinitive* form plays a crucial role in developing the linguistic competence of law students. It enhances their ability to construct precise and clear legal language, formulate arguments, understand and use formulaic patterns, and transition smoothly from academic to professional settings. As such, it is an indispensable component of legal education for non-native English speakers, the nested bundles help the ESL students to achieve fluency and proficiency in the specialised language of law.

The Bundle Nesting Distribution

Nested Bundle distribution in the N-gram feature of the Sketch Engine is made possible by an advanced process that makes it possible to create and examine intricate word patterns and their connections. N-grams are used by the Sketch Engine to recognize and evaluate word sequences, which can include single words, bigrams, trigrams, or larger grams. The N-grams are initially extracted from the corpus by the Sketch Engine. This entails segmenting texts into groups of N words. For instance, under a bigram model, the N-gram produces the string of two consecutive words and trigram, three consecutive times and many others. The tool finds patterns in these N-grams where word sequences (or

bundles) occur inside other sequences. A trigram such as “the quick brown” for example, may be nested inside longer phrases or patterns.

Next, the distributions of these nested bundles are examined using the Sketch Engine. It looks at the frequency with which these nested sequences occur in various text segments and circumstances. In order to do this, the frequency, distribution, and nested occurrences of these N-grams within bigger N-gram bundles must be calculated. The tool applies statistical methods to understand the significance and patterns of these bundles. This might include measures of co-occurrence, dispersion, and collocational strength. The results are presented in a user-friendly format, often through visualization tools of cluster analyses, helping the researchers understand the distribution and significance of nested N-gram bundles in their texts. The statistical distribution is presented in the tables below.

Bundles across the N-gram

Table 4: Percentage of total bundles used across the text.

Number of N-grams	Widely used bundle per N-grams	Occurrences	Percentage
2 N-grams	of the	5,032	83.8%
3 N-grams	the case of	851	94.5%
4 N-grams	in the case of	439	87.8%
5 N-grams	must come with clean hands	219	87.6%
6 N-grams	equity must come with clean hands	201	80.4%

N-grams, which are sub-N-grams of other longer N-grams, are grouped together with the longer N-grams. For example, *of the*, *in the case* and *the case of* are sub-N-grams of *in the case of*. The grouping reveals that 3 N-grams takes the highest percentage in the text, and the rest share equal values but in most cases, some are recurring in combined structures which we refer to as nesting ranges of n-gram values. This kind of nesting is distributed in the table 5.

Lexical Bales Structural Groups

The lexical bales in the list were investigated based on the structural groups found in nesting the tokens. Biber (1988; 2006)

Table 5: Shared lexical bundles of selected N-grams.

Shared lexical bundles	Shared 2 N-grams	Shared 3 N-gram	Shared 4 N-grams	Shared 5 N-grams	Shared 6 N-grams
in the case of	5,032	851	439	76	-
	2,809	497	351		
	2,573	345	311		
	1,251				
Total	11,665	1687	1101	76	0

<i>Shared lexical bundles</i>	<i>Shared 2 N-grams</i>	<i>Shared 3 N-gram</i>	<i>Shared 4 N-grams</i>	<i>Shared 5 N-grams</i>	<i>Shared 6 N-grams</i>
come with clean hand	---	226 215	249 227 225 186 183 181 140	219 207 181 180 173 136	201 173 171 166 134
Total	0	481	1,391	1096	845
who comes to equity must come	---	202	183 181 140 186	136 97	134
Total	0	202	690	285	134
the presumption of resulting trust	597	272	165 147	65	48
Total	597	272	312	65	48
that all responses to the questions	788	254	102	87	87 87
Total	788	254	102	87	174

and Norris (2014) indicate that the N-gram categories indicate the narrow positioning of phrases, long phrases and clauses in Law writing. These clauses and phrases do not reveal any pattern but are occurring because the students have acquired these bundles over time to the other. 2 *N-gram* records 13150, 3 *N-grams* 2896, 4 *N-grams* 3598, 5 *N-grams* 1609 and 6 *N-grams* 1201. The 2 *N-gram* that carries trivial pairs and chunks nested itself in many ways that caused the occurrence of the high number of chunks, 3, 4, and 5 *N-grams* reveal reasonable nest while 6 *N-grams* is nesting in a very minimal pair, this shows a decline in the use of the bundles. The actual bundle or formulaic patterns for Law students in this ESL corpora show that *N-grams* 3, 4, 5 are important shared bundles and formulaic patterns among the discipline.

CONCLUSION

The analysis of huge corpora reveals aspects of language such as lexical bundles and other formulaic linguistic patterns. A short single-authored sample of the language did not reveal the full spectrum of formulaic language, which is characterised by ESL. In a similar vein, no single text will contain every repetition of language found in the larger corpus. It is only the huge patterns exposed in linguistic features like lexical bundles, verb forms and other formulaic linguistic patterns. A little sample written by a single author can never capture the whole range of formulaic language that is unique to that particular discipline language because no

single text will include all instances of words that are repeated throughout the larger corpus. This is the contribution this literature has added to previous literature. In furtherance, verb forms identified in the multi-patterns reveal some active forms that the writers used to achieve lucidity and clarity in the patterns. The repeated forms that yield to active forms of the verbs are common phenomena than the less passivised voice for law language.

As part of the educational process, teachers and students deal with single texts in a given course. These people acquire the patterns in frequent interaction in lecture halls, textbooks, and journal articles and, other subject-related books in chunks. In fact, ESL in academic writing fragments is sometimes mistaken for formulaic patterns. For instance, the students at times include part of online paragraphs, and part of structures from the pamphlets the teachers have given to them and call these formulaic patterns. In fact, these are only chunks they do not even understand. when we consider the positioning of lexical bales in the framework of teaching goal-directed Law English in ESL contexts, then the teachers and the students do not really know what they are looking for or what they intend to look for.

To discover how to use Law lexical bales or bundles at the text level for Law students, the analysis covered texts that are assignments that were submitted online from a variety of wide courses. In this work, we ran 792,237 words in the Sketch Engine as check on the common lexical features of the

bundles in Law texts. This study did not consider single words rather we considered the general features of the bundles found in most Law corpora across law courses.

Acknowledgements

The authors would like to express their profound appreciation and gratitude to all the participants of the study.

REFERENCES

- Adika, G S. K. (2019). Credibility and accountability in academic discourse: Increasing Awareness of Ghanaian Graduate Students. In: Philosophical Foundations of the African Humanities through Postcolonial Perspectives. 197-219. DOI: https://doi.org/10.1163/9789004392946_013
- Afful, J. B. A. (2007). Academic literacy and Communicative Skills in the Ghanaian university: A proposal. *Nebula*, 4(3), 141-159.
- Afzal, N. (2019). A Study on Vocabulary-Learning Problems Encountered by BA English Majors at the University Level of Education. *Arab World English Journal*. 10.(3), 81-98.
- Alfers, H. & Dison, A. (2000). Tracking students through the system: An investigation into EAP students' progress at Rhodes University. *Proceedings of the SAALA 2000 conference vol 1*: 1-10.
- Akmajian, A., Farmer, A. K., Bickmore, L., Demers, R. A., & Harnish, R. M. (2017). *Linguistics: An introduction to language and communication*. MIT press.
- Breeze, R. (2013). Lexical bundles across four legal genres. *International Journal of Corpus Linguistics*, 18 (2), 229-253.
- Biber, D., & Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26(3), 263-286.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, England: Pearson Education Limited.
- Breeze, R., 2013. Lexical bundles across four legal genres. *International Journal of Corpus Linguistics*, 18 (2), 229-253.
- Byrd, P. and Coxhead, A. V. (2010). On the otherhand: Lexical bundles in academic writing and in the teaching of EAP. *University of Sydney Papers in TESOL*, 5, 31-64.
- Choi, Y., & Zhang, D. (2021). The relative role of vocabulary and grammatical knowledge in L2 reading comprehension: A systematic review of literature. *International Review of Applied Linguistics in Language Teaching*, 59(1), 1-30.
- Cravens, S. M. (2024). Recording Judges: Filling Gaps to Improve Judicial Discipline. *J. App. Prac. & Process*, 24, 1.
- Castellano-Risco, I. (2024). Profiling lexical coverage in EMI academic seminars: a corpus-based study. *Ibérica*, (47), 227-250.
- Coxhead, A. V. and Byrd, P. (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing*, 18 (3). 129-147.
- Chalikandy, M. A. (2013). A comprehensive Method for Teaching English for Specific Purpose. *Arab World English Journal*. . International peer-reviewed journal ISSN: 2229 -9327, 310-320.
- Conrad, S. (2008). Writing Myth 6: Corpus-based research is too complicated to be useful for writing teachers. In J. Reid (Ed.), *Myths about teaching writing* (pp.115-139). Ann Arbor: University of Michigan Press.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23, 397-423.
- Coxhead, A. (2020). Academic vocabulary. In S. Webb (Ed.), *Handbook of vocabulary studies* 97-110. Routledge.
- Coxhead, A. (2008). Phraseology and English for academic purposes: Challenges and opportunities. In F. Meunier & S. Granger (Eds.), *Phraseology in language learning and teaching* 149-161.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-238.
- Duszak, A. (1997) 'Cross-cultural academic communication: A discourse community view.' In:
- Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics*, 38(2), 165-193.
- Duszak, A. (ed.) *Culture and Styles of Academic Discourse*. Berlin and New York: Mouton de Gruyter. 11-39.
- Eliazarian, H. A. (2024). Moving from a rule-based approach to a functional understanding in Plain Legal Language. *The International Journal of Law, Language & Discourse*, 12(1).
- Ellis, N. C. (2012). Formulaic language and second language acquisition: ZZIPP and the phrasal teddy bear. *Annual Review of Applied Linguistics* 32, 17-44. Cambridge University Press, doi: 10.1017/S0267190512000025.
- Evans, S. & Green, C. (2007) Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*. 6 (1), 1-17.
- Frese, M., & Sabini, J. (Eds.). (2021). *Goal directed behavior: The concept of action in psychology*.
- Gardner, H. (2021). *Disciplined mind: What all students should understand*. Simon & Schuster.
- Garner, B. A. (2023). *Legal writing in plain English: A text with exercises*. University of Chicago Press.
- Granger, S. & Meunier, F. (2008). Phraseology in language learning and teaching: Where to from here? In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching* (pp.247-252). Amsterdam: John Benjamins.
- Granger, S. (2018). Tracking the third code. *The Corpus Linguistic Discourse: In Honour of Wolfgang Teubert, Studies in Corpus Linguistics*, 185-204.
- Greenawalt, K. (1995). *Law and objectivity*. Oxford University Press.
- Gilquin, G., Granger, S., & Paquot, M. (2007). Learner corpora: The missing link in EAP pedagogy. *English for Academic Purposes*, 6, 319-335.
- Golparvar, S. E., & Barabadi, E. (2020). Key phrase frames in the discussion section of research articles of higher education. *Lingua*, 236, 102804.
- Hadley, G. (2014). *English for academic purposes in neoliberal universities: A critical grounded theory* (Vol. 22). Springer.

- Hall, C. J., Smith, P. H., & Wicaksono, R. (2017). *Mapping applied linguistics: A guide for students and practitioners*. Routledge.
- Heatley, A., Nation, I. S. P., & Coxhead, A. (2002) Range: A program for the analysis of vocabulary in texts. <http://www.vuw.ac.nz/lals/staff/paulnation/nation.aspx>
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning centered approach*. Cambridge University Press.
- Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing*, 6(2), 183–205.
- Hyland, K. and Tse, P. (2012). Academic Lexis and Disciplinary practices: Corpus evidence for specificity. *International Journal of English Studies (IJES)*. 9(2). 111-129.
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4-21.
- Hyland, K. (2012). Bundles in academic discourse. *Annual review of Applied Linguistics*. (32)
- Jordan, B. (1997). *English for Academic Purposes*. Cambridge: CUP.
- James, C. (2013). *Errors in language learning and use: Exploring error analysis*. Routledge.
- Jones- Mensah, I., Tabiri, M.O., Fenyi, D.A., Kongo, A.E. & Amexo, D. (2022). Vocabulary Knowledge of Collocation in Business Texts: A Case of ESL Tertiary Students. Available online at globets.org/journal International Journal of Education, Technology and Science. 2(1) 001–023.
- Kalen, S. (2022). Refining Statutory Construction: Contextualism & Deference. *UNHL Rev.*, 21, 261.
- Kim, J. (2013). Lexical bundles in Korean college students' English essays: A corpus-based comparative study. *English Language and Literature Teaching*, 19(3), 157-179.
- Law, H. Q. (2022). *Corpus-based analysis of the usage pattern of homonymys among malaysian ESL learners* (Doctoral dissertation, UTAR).
- Lee, J. H., & Macaro, E. (2013). Investigating age in the use of L1 or English-only instruction: Vocabulary acquisition by Korean EFL learners. *The Modern Language Journal*, 97(4), 887-901.
- Leech, G., & Rayson, P. (2014). *Word frequencies in written and spoken English: Based on the British National Corpus*. Routledge.
- Liu, D. (2013). *Describing and explaining grammar and vocabulary in ELT: Key theories and effective practices*. Routledge.
- Lockwood, J. (2012). Are we getting the right people for the job? A study of English language recruitment assessment practices in the Business Processing Outsourcing (BPO) sector: India and the Philippines. *Journal of Business Communication*, 49(2), 107–127.
- Martínez, A. C. L. (2024). The role of syntactic and lexical complexity in undergraduate writing quality. *ibérica*, (47), 251-274.
- Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics*, 32, 111–129. <https://doi.org/10.1017/S0267190512000128>
- Nartey M. and Dorgbetor N. (2014). Towards a Needs Analysis of why Students in a Ghanaian University Fail the Academic Writing Course.
- Norris, M. (2014). *A theory of nominal concord* (Doctoral dissertation, UC Santa Cruz).
- Paquot, G. and Granger, M. (2007). Spoken features in learner academic writing: identification, explanation and solution. *Corpus linguistics conference*. In: Davies M.; Rayson P.; Hunston S.; Danielsson P., *Proceedings of the Fourth Corpus Linguistics Conference*, University of Birmingham: <http://hdl.handle.net/2078.1/75861>.
- Rees, G. P. (2018). A phraseological multi-discipline approach to vocabulary selection for English for academic purposes.
- Rose, H., McKinley, J., Xu, X., & Zhou, S. (2020). Investigating policy and implementation of English medium instruction in higher education institutions in China. *London: British Council*. Retrieved October 14, 2020.
- Rodrigo, A. Á. V. (2023). A lexical bundle analysis of art-related crowdfunding projects. *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)*, (46), 321-349.
- Shintani, N. (2012). Input-based tasks and the acquisition of vocabulary and grammar: A process-product study. *Language Teaching Research*, 16(2), 253-279.
- Smith, G. (2021). *Towards a Politically Engaged Anthropology*. London: Routledge. <https://doi.org/10.4324/9781003085010>
- Springer, J. W. (1981). An ethnohistoric study of the smoking complex in eastern North America. *Ethnohistory*, 217-235.
- Tripto, J., Assaraf, O. B. Z., & Amit, M. (2018). Recurring patterns in the development of high school biology students' system thinking over time. *Instructional Science*, 46, 639-680.
- Véliz, C. (2023). *The new world of the gothic fox: culture and economy in English and Spanish America*. Univ of California Press.
- Vuković, M., & Čarapić, D. (2024). Vocabulary profile, lexical density and speech rate of science podcasts: How adequate are they for EAP and ESP listening?. *Ibérica: Revista de la Asociación Europea de Lenguas para Fines Específicos (AELFE)*, (47), 201-226.
- Yakut, I., Yuvayapan, F., & Bada, E. (2021). Lexical bundles in L1 and L2 English doctoral dissertations. *Journal of Teaching English for Specific and Academic Purposes*, 475-493.
- Zwiers, J. (2014). *Building academic language: Meeting common core standards across disciplines, grades 5-12*.