PASAA Journal
Volume 67, July – December 2023, 66–100

Investigating Proficiency of Academic English in Student Writing: A Comparative Case Study on Vocabulary Utilization in Student Research Article Writing vis-à-vis National and International Research

Donlawat Meebangsai, Pawarit Pongtin, Panjaphon Kitipoontanakorn, and Piyapong Laosrirattanachai*

Faculty of Hospitality Industry, Kasetsart University, Kamphaeng Saen Campus, Nakhonpathom, Thailand

**Corresponding author: piyapong.l@ku.th*

Article information	
Abstract	The researchers investigated the usage of appropriate academic English vocabulary in research papers authored by English major students, publications in Thai national journals, and international journals. In total, 708 papers published between 2016 and 2018 were analyzed, consisting of 300 international research papers, 300 national research papers, and 108 research papers authored by English major students. The study focused on four aspects of lexical evaluation to analyze the vocabulary in these papers: lexical profiling, lexical level, lexical variation, and lexical density. The analysis yielded the following results: 1) Lexical profiling revealed that the usage of academic words in students' research papers was 8.11%, which was lower than that in national research papers (10.96%) and international research papers (12.01%). 2) In terms of lexical level, students' research papers had a medium-frequency word usage of 3.85%, lower than that of international research papers (4.17%) but higher than that of national research papers (3.32%). 3) Concerning lexical variation, the research papers authored by

E-ISSN: 2287-0024

	English major students had a rate of 24.26%, which was lower than that of national research papers (25.01%) but higher than that of international research papers (21.69%). 4) Finally, in terms of lexical density, the ratio of function words to content words was similar across students' research papers (57.59%), publications in national journals (57.19%), and publications in international journals (57.52%). Students who are required to write research papers and novice researchers are recommended to adhere to the standardized ratios set by publications in international journals when aiming to publish their research in such journals.
Keywords	academic writing, lexical profiling, lexical level, lexical variation, lexical density
APA citation:	Meebangsai, D., Pongtin, P., Kitipoontanakorn, P., & Laosrirattanachai, P. (2023). Investigating proficiency of academic English in student writing: A comparative case study on vocabulary utilization in student research article writing vis-à-vis national and international research. *PASAA, 67*, 66–100.

1. Introduction

In the field of academic research, it is crucial for researchers to publish their completed work to facilitate engagement with the broader public. These publications can take various forms, including research conferences, national research journals, and the highly esteemed category of international research publications. Evaluating the quality of research is a multifaceted endeavor (Hug et al., 2013). Scholars encounter numerous challenges when composing research papers at the international level (Harzing, 2013). Additionally, the process of selecting and recognizing papers for publication in international academic journals requires a higher level of rigour compared to that for academic conferences and national journals. One pivotal criterion for attaining recognition as a high-quality study is the proficient use of standard English language, guaranteeing widespread acceptance within the scholarly community. To have a manuscript published, particularly in international research journals, it is crucial to adhere to several stringent guidelines, one of which pertains to the utilization of academically appropriate English language standards. Prior to submission, it is imperative to diligently edit manuscripts to uphold language excellence. The onus rests on the author to ensure that the manuscript has undergone thorough editing and proofreading and has adhered strictly to the requisite academic standards of English. Any manuscript that fails to meet this prerequisite will be returned to the author for revision before entering the peer review process. Several aspects of English language usage need to be carefully considered to determine whether they meet the criteria of “academically appropriate English language standards.” One of these aspects is word choice, which plays a major role in establishing an academic tone and adhering to the conventions of scholarly writing. By using appropriate vocabulary, writers can elevate the academic quality of their work.

Numerous studies have been conducted to explore English vocabulary across various domains. These studies have encompassed several aspects, including: 1) lexical profiling, which involves the classification of terms based on the General Service List (West, 1953), the Academic Word List (Coxhead, 2000),

and the Outside Word List (Coxhead, 2000; Coxhead & Hirsch, 2007); 2) lexical level, which measures the linguistic proficiency of words, encompassing low-frequency words, medium-frequency words, and high-frequency words (Schmitt, 2000); 3) lexical variation, which focuses on the analysis of the Type-Token Ratio (TTR) (Geeraerts, 1993); and 4) lexical density, which involves comparing the usage of function words and content words (Ure, 1971). These comprehensive studies have made important contributions to the general understanding of the complexities and nuances of English vocabulary in diverse contexts. Several studies have utilized these four aspects to analyze vocabulary, such as in the analysis of written assignments (Llach, 2014), a standardized English proficiency test (Webb & Paribakht, 2015), the International English Language Testing System (IELTS) (Phung & Ha, 2022), research article abstracts (Nguy & Ha, 2022), and the development of various specialized word lists (Coxhead, 2000; Laosrirattanachai & Laosrirattanachai, 2021). However, to the best of our knowledge, there has not been an analysis of academic vocabulary use in research writing applying these four aspects. Furthermore, no comparisons have been made between the vocabulary use in research articles published in various quality journals and those written by undergraduate students. The results of this research should underscore the importance of being acquainted with academic vocabulary for use in research writing, as well as understanding the levels of academic vocabulary proficiency that should be attained based on the analysis of international research as a standard. The studies mentioned above are intriguing, specifically those that examined the utilization of vocabulary in research papers across different groups, including research papers authored by English major students, publications in Thai national journals, and publications in international journals. The objective of the present study was to conduct a comprehensive analysis to determine the established criteria governing the usage of English vocabulary in research paper writing. The primary emphasis was on scrutinizing the application of English vocabulary in research papers.

The present study applied four criteria, namely lexical profiling, lexical level, lexical variation, and lexical density, to analyze the vocabulary used in research papers published across various levels. Specifically, the analysis focuses on international research papers published in SCOPUS-indexed journals that are renowned for their exemplary quality. The primary objective was to establish benchmarks for the usage of English vocabulary in published research papers. Once the criteria were established, a comparison was conducted between national research and the findings of international research to assess the degree of adherence to international standards of English language usage in research. Furthermore, the research papers authored by undergraduate students enrolled in the English program were examined and juxtaposed with both national and international research to discern any divergences in the standards of English vocabulary applied in research writing among undergraduate students.

2. Literature Review

In the realm of English communication skills, vocabulary stands out as a foundational component (Laufer & Nation, 1995; Lestari & Hardiyanti, 2020). The acquisition of vocabulary holds paramount importance for second language learners (Alqahtani, 2015) because a robust vocabulary serves as the linchpin for successful communication (Nation, 2022). Even if English as a Foreign Language (EFL) learners lack proficiency in grammar, they can still comprehend English; their pronunciation may not be flawless, yet communication remains possible. However, the absence of vocabulary knowledge poses substantial challenges in both communication and writing (Schmitt, 1997; Wilkins, 1972). It has been widely recognized that vocabulary knowledge plays a pivotal role in students' academic achievements (Nagy & Townsend, 2012). This is because EFL learners with limited vocabulary encounter barriers to success in academic discourse (Mozaffari & Moini, 2014). It is clear that vocabulary knowledge is indispensable for academic writing. To effectively enhance academic vocabulary, it is necessary to evaluate the existing vocabulary proficiency of language users. The present study aimed to explore vocabulary knowledge in academic writing based on the following aspects.

2.1 Lexical Profiling

In 1995, Laufer and Nation introduced the concept of lexical profiling, which provides a methodology for classifying words into specific categories using word lists. The primary objective of lexical profiling is to assign words to a singular profile or reference word list, effectively eliminating any redundant or unrelated words from the compiled lists. During its inception, the reference word lists comprised the General Service List (GSL), which consisted of the initial 1,000 and a subsequent 1,000 high-frequency words (West, 1953), along with the University Word List (UWL) (Xue & Nation, 1984). Subsequently, the UWL was replaced by the Academic Word List (AWL) (Coxhead, 2000). Nation (2022) employed a systematic classification to facilitate the acquisition of English vocabulary. This classification delineates four discernible groups: high-frequency words, academic words, technical words, and low-frequency words. Each group has its distinct attributes and scholarly importance. Certain researchers amalgamate technical words and low-frequency words and refer to them collectively as “outside words” (Coxhead & Hirsh, 2007). For a more comprehensive depiction of these categories and a thorough elucidation of their distinctive attributes, see the subsequent explanations.

The General Service List (GSL) encompasses English words that are commonly applied in everyday discourse. It stands as the preeminent compilation of high-frequency vocabulary in the English language. Originally formulated by West in 1953, the GSL consists of a comprehensive set of 2,000 frequently occurring words. As posited by Nation and Waring (1997), the GSL is estimated to account for approximately 80% of the lexical composition encountered in a given text.

The Academic Word List (AWL) is a compilation of English words commonly utilized in academic contexts. Early contributors to the development of an academic word list include Campion and Elley (1971), Praninskas (1972), Lynn

(1973), and Ghadessy (1979). Subsequently, Xue and Nation (1984) refined their previous word list and transformed it into the University Word List (UWL), which underwent further enhancements and modifications. A significant advancement in academic word list construction took place in 2000, when Coxhead introduced the AWL, consisting of 570 new academic words. Coxhead (2000) suggested that the AWL should account for approximately 10% of the vocabulary found in a given text.

The Outside Word List (OWL) is a distinct component of vocabulary that falls outside the GSL and AWL. Coxhead (2000) and Coxhead and Hirsh (2007) introduced the concept of the OWL, which comprises English-specific terms, proper nouns, technical jargon, and words borrowed from other languages. Typically, the OWL is expected to account for approximately 10% of the total vocabulary in a given text.

The lexical profiling approach has been applied to analyze the vocabulary utilized in a wide range of research studies focusing on vocabulary. These studies include investigations, such as analyzing the lexical profile of written assignments completed by young CLIL learners (Llach, 2014), exploring the relationship between the lexical profile of test items and performance on a standardized English proficiency test (Webb & Paribakht, 2015), determining the proportion of different lexical profiles necessary for comprehending research article abstracts (Nguy & Ha, 2022), and developing various specialized word lists using the lexical profiling as a key criterion (Arunvong Na Ayuthaya et al., 2022; Laosrirattanachai & Ruangjaroon, 2021; Namwong et al., 2022; Rungrueang et al., 2022).

In the present study, the lexical profiling approach was applied, utilizing three reference word lists: the GSL (West, 1953), AWL (Coxhead, 2000), and OWL. These word lists were used to classify the vocabulary present in research papers authored by English major students, publications in Thai national journals, and international journals.

2.2 Lexical Level

The lexical level utilized in this study was based on Paul Nation's BNC-COCA (British National Corpus/Corpus of Contemporary American) Word family lists, which are renowned resources for analyzing vocabulary. These lists have been thoughtfully organized into 1,000-word frequency bands, allowing for a systematic examination of lexical patterns. The Vocabulary Profiler, accessible at https://www.lextutor.ca/vp/comp/, effectively uses this version of the lexical profiling tool, encompassing a comprehensive spectrum from the most frequently occurring 1,000 words to the uppermost echelon of 25,000 words in the English language. Notably, the BNC-COCA word list (Nation, 2016) serves as an indispensable compendium of English vocabulary derived from contemporary corpora. Its authoritative status is evident through its consistent utilization as a reference for assessing vocabulary size and level in various examinations (Coxhead et al., 2014; McLean et al., 2015; McLean & Kramer, 2015; Nation & Beglar, 2007).

Lexical levels are connected to the ranking of word frequencies, which include high, medium, and low-frequency words. According to Schmitt and Schmitt (2014), it is more advantageous to use a three-tiered classification system for words, rather than simply categorizing them as high-frequency or low-frequency. This system entails identifying approximately 3,000 high-frequency words, 6,000 medium-frequency words, and categorizing the remaining words as low-frequency. In this study, the primary objective of this selection was to emphasize the importance of systematically acquiring medium-frequency words. To achieve this, the researchers divided the lexical range from 1,000 to 25,000 (K1-K25) into three sections. The first section encompassed high-frequency words ranging from 1,000 to 3,000 (K1–K3), the second section consisted of medium-frequency words spanning from 4,000 to 9,000 (K4–K9), and the third section included low-frequency words from 10,000 to 25,000 (K10–K25). Lexical levels were also linked with lexical coverage. In fact, several scholars have posited that achieving a vocabulary coverage of 95% is deemed acceptable, while striving for a minimum of

98% is considered the ultimate objective for learners (Laufer & Ravenhorst-Kalovski, 2010; Schmitt et al., 2011). Given the high standards and linguistic complexity demanded in writing academic research papers, there is a strong likelihood of utilizing lexical levels derived from higher base lists (K10-K25) to achieve coverage levels of 95% and 98% in writing.

Öztürk (2022) highlighted both the strengths and weaknesses of the BNC/COCA word frequency list. The organization of the list based on frequency is considered a strength as it enables a systematic and unbiased selection of target terms for vocabulary range exams. However, relying solely on frequency as the criterion for selecting target words can be seen as a weakness, given that lexicons are inherently diverse and word complexity is influenced by various factors.

Various studies in the field of vocabulary exploration utilize the concept of lexical level to analyze vocabulary usage. For example, Laosrirattanachai and Laosrirattanachai (2023) examined the lexical level applied in conducting press conferences, while Phung and Ha (2022) investigated the lexical level required to achieve 95% coverage in the listening sub-test of the widely recognized International English Language Testing System (IELTS). In line with these investigations, the present study aimed to explore the lexical level utilized in research papers authored by English major students, as well as publications in Thai national journals and international journals.

2.3 Lexical Variation

Lexical variation, also known as lexical diversity, encompasses the extent to which a text showcases a rich and varied vocabulary. A high degree of lexical variation signifies a wide range of distinct lexical choices used by the speaker or writer, reducing the prevalence of repetitive word usage. To achieve an elevated level of lexical variation, a text requires the incorporation of numerous distinct words, thereby augmenting the richness and intricacy of the discourse (Johansson, 2008).

The measure of lexical variation is the type-token ratio (TTR), which calculates the ratio of unique words (types) to the total number of words (tokens) in a given text (Lieven, 1978; Bates et al., 1988). The TTR serves as a widely accepted metric for quantifying the lexical richness and diversity exhibited within a text (Davis & Brewer, 1997; Wolfe-Quintero et al., 1998). Tokens represent the total count of individual words within a given text, whereas types are defined as the total count of distinct or unique words in the text. The TTR quantifies the ratio of unique words to the total number of words in a given text, thereby providing valuable insights into the extent of lexical diversity present within the discourse. A higher TTR signifies a broader spectrum of linguistic forms and potential transformations in the writer's register or style (Baker, 2006). Furthermore, as the length of the text increases, the vocabulary used tends to manifest greater variability. Consequently, to attain a heightened level of lexical diversity, speakers and writers must demonstrate adeptness in utilizing an extensive repertoire of terms while minimizing redundancy. The lexical variation can be determined by applying the following formula to calculate the TTR (Laufer & Nation, 1995):

$$\text{Type-Token Ratio} = (\text{Number of Types} / \text{Number of Tokens}) \times 100$$

In the context of the present study, the TTR was utilized as a fundamental tool to assess and measure the lexical variation present in research papers written by English major students, as well as publications in Thai national journals and international journals.

2.4 Lexical Density

Lexical density, a concept employed in text analysis, pertains to the ratio of content words, which consist of nouns, verbs, adjectives, and adverbs, to function words within a given text (Thornbury & Slade, 2006). Empirical evidence suggests a pronounced disparity in the levels of lexical variation and lexical density between written and spoken language contexts (Ure, 1971; Halliday, 1985). Notably, these

studies demonstrate that both measures exhibit considerably greater magnitudes within written discourse compared to oral communication. The study by Ure (1971) suggests that texts produced by proficient speakers typically exhibit a lexical density below 40%, whereas written texts generally demonstrate a lexical density of approximately 40% or higher. It is crucial to emphasize that these measurements are considerably influenced by the language proficiency of the speaker or writer.

Kondal (2015) conducted a study to investigate the influence of lexical density on communication ability. The findings indicated that written proficiency is reflected in the lexical density and diversity of the text, both of which have a major impact on language performance. Additionally, the study highlighted a positive correlation between higher lexical density and increased proficiency levels among learners. The variations in variety and lexical density observed in their writing serve as distinguishing factors for these proficiency levels. It is evident that lexical density contributes to the enhancement of students' writing in academic contexts. An accomplished academic writer carefully selects words that are pertinent to the topic at hand and strives to refine ideas by applying more precise terminology. In contrast, writers with lower proficiency levels often adopt a more elaborate writing style while using fewer words, resulting in a reduction in lexical density. Determining lexical density can be accomplished by utilizing the formula provided by Laufer and Nation (1995):

Lexical Density = (Number of content words / Number of Tokens) x 100

A text is categorized as "dense" when it contains a high proportion of content words relative to the total number of words, as content words play a crucial role in conveying information. However, notably, lexical density may not solely reflect the presence of lexicon, as it is also influenced by the syntactic and cohesive aspects of the composition.

In the present study, the researchers applied lexical density to analyze the proportion of function words and content words, aiming to determine whether the research papers authored by English major students exhibited higher or lower density compared to publications in Thai national journals and international journals.

The present study used lexical profiling, lexical level, lexical variation, and lexical density to analyze the vocabulary utilized in research paper writing. The study aimed to address the following research questions:

1. What percentage of vocabulary coverage of the lexical profiles, including the GSL, AWL, and OWL, is contained in the research papers published in international academic journals, Thai national academic journals, and those authored by English major students?
2. To what extent are the high, mid, and low-frequency word levels distributed in the research papers published in international academic journals, Thai national academic journals, and those authored by English major students, specifically concerning the K1–K25 word levels?
3. What ratio of lexical variation is observed in research papers published in international academic journals, Thai national academic journals, and those authored by English major students?
4. What is the lexical density ratio in research papers published in international academic journals, Thai national academic journals, and those authored by English major students?

3. Methodology

3.1 Data Collection

The present study aimed to investigate the use of vocabulary in research article writing associated with the study of English. Therefore, the collected research papers published at both national and international levels must exhibit strong relevance to various disciplines within the field of linguistics. These disciplines encompassed areas such as English for Specific Purposes, English

Language Teaching, and other domains closely associated with the study of English. To align with the analysis, undergraduate research articles must have been authored by students in the English language major. In this study, research papers were collected from three main sources: research papers authored by English major students, publications in Thai national journals indexed in TCI (Tier 1), and international journals indexed in SCOPUS. The research papers authored by students majoring in English originated from Thai students in their fourth academic year who used Thai as their primary language. There were no specific criteria for selecting research papers authored by students for this study, as all research papers were utilized for the comprehensive analysis. The research papers authored by English major students spanned the academic years 2016 to 2018. Similarly, 100 papers per year were chosen from Thai national journals indexed in TCI (Tier 1) and international journals indexed in SCOPUS, covering the years 2016 to 2018. Notably, the selected research papers published in national journals were authored by Thai researchers who were Thai citizens and used Thai as their primary language. However, there was no specification of the nationality and primary language of the authors of research papers published in international journals, as it was believed that these papers had undergone review and scrutiny for appropriateness in academic discourse as a part of the publication process used by reputable international journals. The total number of research papers collected was 708, consisting of 300 international research papers with a total size of 289,542 tokens, 300 national research papers with a total size of 1,090,357 tokens, and 108 papers from English major students with a total size of 1,583,014 tokens. Overall, there were 2,956,326 tokens. Abstracts were excluded from the corpora compilation because the abstract contains a summary of the entire research paper, which could result in duplicate data. Additionally, the reference and appendix (if any) sections mainly comprised specific names and terminology, so they were also excluded from the compilation process.

3.2 Research Instruments

The present study used two primary research instruments for data collection and analysis:

1. The AntWordProfiler (Anthony, 2022) is a software tool designed for classifying words in a corpus into distinct profiles. These words are categorized into four groups. The first and second groups comprise the 2,000 most frequently used words according to the GSL (West, 1953). The third group consists of specialized terminology found in the AWL (Coxhead, 2000), whilst the fourth group encompasses words not included in these lists, such as technical terms or infrequently used terms. Furthermore, the program provides information on the type and token of the words (Lieven, 1978). Users have the flexibility to include additional word lists as reference lists, depending on their specific analysis goals, such as incorporating function word lists for studying function words and content words. In the present study, the researchers used the AntWordProfiler to investigate lexical profiling, lexical density, and lexical variation.

2. The VocabProfile (Cobb, 2022) is an online program developed by Cobb and can be accessed at https://www.lextutor.ca/vp/comp/. This software has been specifically designed to classify words into 26 essential lists. The first 25 lists are generated based on frequency ratings derived from the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA) combined. Each of these base lists comprises 1,000 words. The 26th base list, also referred to as the off-list, encompasses various words that may appear, including less common words with a frequency below 25,000, new vocabulary, transliterated words, and misspelled words. In this study, the VocabProfile was utilized to analyze the lexical level.

3.3 Data Analysis

The data collected were examined using the following methods:

1. Lexical profiling: The GSL, AWL, and OWL were utilized as reference lists for the analysis. These lists are already integrated into the AntWordProfiler software. Through the examination of a total of 708 research papers, the program determined the proportion of GSL, AWL, and OWL words in each paper.

2. Lexical level: The researchers utilized the VocabProfile program to assess the vocabulary level. Within the lexical level, there were categories ranging from K1 to K25, as well as the off-list category. After analyzing the 708 research papers on the website, the program displayed the number of words falling within the K1 to K25 range and the off-list category for each research paper. The researchers further divided the levels into three main groups: high-frequency words (K1–K3), medium-frequency words (K4–K9), and low-frequency words (K10–K25), plus the off-list.

3. Lexical variation: The researchers utilized the AntWordProfiler program to analyze all the research papers and to examine the number of types and tokens. After analyzing the 708 research papers, the program provided the count of types and tokens for each research paper. Then, the researchers applied the TTR formula to calculate the lexical variation.

4. Lexical density: The researchers utilized the AntWordProfiler program to analyze all the research papers and to distinguish between function words and content words. The analysis of the 708 research papers using the program provided the counts of function words and content words for each research paper. Then, the researchers applied the equation to calculate the lexical density.

After compiling all the statistics, a comparative analysis of the results was conducted among the research papers authored by English major students, publications in Thai national journals, and publications in international journals.

4. Findings

The overall analysis of the 708 research papers authored by English major students, publications in Thai national journals, and publications in international journals had a combined size of 2,956,326 running words. The explanations for each research question (RQ) are presented below.

RQ1. What percentage of vocabulary coverage do the lexical profiles, including the GSL, AWL, and OWL, offer in the research papers published in international academic journals, Thai national academic journals, and those authored by English major students?

To address research question 1, the AntWordProfiler program was utilized to perform the lexical profiling. Table 1 provides an illustration of the proportions of vocabulary categorized into the three reference word lists: the GSL, AWL, and OWL.

Table 1

Lexical Profiling of Vocabulary Used in Research Papers from Three Groups

Lexical Profile	%	Students' papers	National papers	International papers
General Service List	Max	88.39	85.81	87.02
	Min	56.14	60.88	64.84
	Average	**79.80**	**76.41**	**75.21**
Academic Word List	Max	14.40	18.01	20.31
	Min	1.91	5.08	4.70
	Average	**8.11**	**10.96**	**12.01**
Outside Word List	Max	39.03	28.53	25.24
	Min	5.65	5.58	4.00
	Average	**12.09**	**12.63**	**12.78**

Based on the analysis results presented in Table 1, the GSL had average coverages of 79.80%, 76.41%, and 75.21% in the research papers authored by

English major students, publications in Thai national journals, and publications in international journals, respectively. Considering the AWL, the average coverage levels were 8.11%, 10.96%, and 12.01% for research papers authored by English major students, publications in Thai national journals, and publications in international journals, respectively. These findings indicate the prevalent usage of academic terminology in international journals, which align their publications with the scholarly discourse. Furthermore, a comparative analysis of the AWL coverage among the international journals, Thai national journals, and research papers authored by English major students highlights a deficiency in the utilization of academic terminology by both Thai researchers and students, particularly the latter group, in their scholarly works. Some examples are provided below of the lexical profiles used in research papers. Words from the GSL are presented in italics, while words from the AWL are highlighted in bold, and words from the OWL are underlined.

Example 1: Excerpt from a research paper authored by a student

> "Having knowledge of language in sports games is **crucial** for keeping up with the modern world. Whamsiri (2001) discussed the importance of language in sports journalism, **emphasizing** that the language and vocabulary used in sports news writing should evoke a sense of excitement, fun, and liveliness in readers. By using **appropriate** language, readers can **accurately** understand the news content and easily engage their imagination for enjoyable reading."

Example 2: Excerpt from a publication in a Thai national journal

> "The difference in **publishing procedures** in terms of the reviewing **process** is a possible **factor** leading to differences in the information **structure** in standard and predatory **journal** articles. Therefore, the move **analysis** in this study will help **researchers** to **differentiate** between these two **journals** and avoid being victims of predatory journals, but rather serve to raise **awareness** of these **researchers.**"

Example 3: Excerpt from a publication in an international journal

> "The current study **sought** to examine the influence of Korean college students' **initial motivation** on their English as a Foreign Language (EFL) writing development, as well as the relationships among their learner characteristics, including EFL writing performance, self-efficacy, and interest. **Motivation** has been shown to hold importance in EFL writing development, as EFL learners **require psychological stability** and **sufficient energy** to overcome anxiety and difficulties stemming from their unfamiliarity with the new language."

RQ2. To what extent are the high, mid, and low-frequency word levels distributed in the research papers published in international academic journals, Thai national academic journals, and those authored by English major students, specifically concerning the K–1 to K–25 word levels?

To address the second research question, the researchers utilized the VocabProfile, an advanced tool specifically designed for analyzing vocabulary proficiency. In total, 708 research papers were selected from a diverse range of sources, consisting of those authored by English major students, publications in Thai national journals, and publications in international journals. These papers were subjected to analysis on the VocabProfile website, which provided comprehensive insights into the vocabulary levels observed within the 1K–25K range, including the incorporation of off-list words. To ensure effective categorization of the findings, the vocabulary levels were classified into three distinct groups: high-frequency words (K1–K3), medium-frequency words (K4–K9), and low-frequency words (K10–K25), as well as the off-list category. The resulting outcomes are outlined in Table 2.

Table 2

Lexical Level of Vocabulary Used in Research Papers from Three Groups

Lexical Level	%	Students' papers	National papers	International papers
High-frequency word (K1–K3)	Max	98.50	98.50	98.30
	Min	80.00	79.80	87.40
	Average	**94.54**	**94.67**	**94.15**
Mid-frequency word (K4–K9)	Max	11.60	9.30	11.10
	Min	1.30	0.90	1.20
	Average	**3.85**	**3.32**	**4.17**
Low-frequency word (K10–K25) + Off-list	Max	11.89	13.65	5.90
	Min	0.04	0.04	0.03
	Average	**1.29**	**1.59**	**1.52**
Mid-, Low-frequency words and Off-list combination	**Average**	**5.14**	**4.91**	**5.69**

According to Table 2, the research papers in the three groups had remarkably similar utilization levels of high-frequency words in their papers, with percentages of 94.54%, 94.67%, and 94.15% for English major students, Thai national journals, and international journals, respectively. Notably, in terms of mid-frequency word usage, research papers published in international journals had a higher proportion (4.17%) compared to the other two groups (3.85% and 3.32%, respectively). Combining the proportions of mid-frequency words, low-frequency words, and off-list words indicated that research papers in international journals consistently utilized a higher number of non-high-frequency words. This finding underscores the standard of research paper writing by international scholars, who incorporate 4.17% to 5.69% of non-high-frequency words. Several examples of word extractions are provided, where high-frequency words are italicized, mid-frequency words are bolded, and low-frequency words and off-list words are underlined.

Example 4: Excerpt from a research paper authored by a student

> "Technology also provides opportunities for people to enjoy online information and **recreational** activities. Among these activities, games have gained popularity as they offer emotional benefits and facilitate the development of new skills, such as hand-eye coordination, intellectual abilities, **tact**, and determination. Additionally, games can serve as a means to practice foreign language skills."

Example 5: Excerpt from a publication in a Thai national journal

> "Afterwards the researcher teacher **transcribed** the responses from videotape into written form, and then translated the **transcriptions** into English. All questions were designed to capture how participants responded to the experiment to help answer the research questions. Twelve structured interview questions related to integrative motivation were divided into three parts following the integrative motivation constructs: interest, effort, and attitude (alphanumeric codes below refer to these questions) have been summarized here…"

Example 6: Excerpt from a publication in an international journal

> "The present findings showed that participants' motivational orientation significantly predicted their EFL writing performance at the beginning of a semester. This result was consistent with previous studies regarding EFL acquisition and general learning. Students' EFL **proficiency** develops as their self-**efficacy** is encouraged or vice versa."

RQ3. What ratio of lexical variation is observed in research papers published in international academic journals, Thai national academic journals, and those authored by English major students?

To explore the third research question, the AntWordProfiler program was used to examine the types and tokens in each research paper within all three

groups. Subsequently, the TTR value of each research paper was calculated, from which the average value was determined for comparisons among the three groups. The analysis results are summarized in Table 3.

Table 3

Lexical Variation of Research Papers from Three Groups

Lexical Variation	%	Students' papers	National papers	International papers
Type-Token Ratio (TTR)	Max	35.57	43.60	32.94
	Min	3.86	11.12	9.52
	Average	**24.26**	**25.01**	**21.69**

Based on the findings presented in Table 3, the lexical variation of research papers varied among the three groups, with national papers having the highest TTR (25.01%), which contrasted with international journals having the lowest TTR (21.69%). In many linguistic contexts, the ability to apply a wide range of vocabulary while minimizing redundancy may indicate language proficiency. However, when writing research papers intended for publication within the academic discourse, the opposite holds true. Drawing from our prior experiences in publishing research papers, particularly in reputable international journals indexed in SCOPUS and ERIC, we have observed that peer reviewers and language editors strongly advocate for consistency in word choice. They emphasize the importance of using consistent terminology throughout the paper to enhance reader comprehension. These insights align with the results presented in Table 3, which reveal that publications in international journals had less variety of terms in conveying their messages. Below, examples highlighting the use of the near-synonym words "ensure" (bolded) and "guarantee" (underlined and italicized) are provided.

Example 7: Excerpt from a research paper authored by a student

"The researcher chose to study the discourse in this film because it explores the participation of 19th-century or Victorian era political

movements that campaigned for electoral reform laws to **ensure** equal rights and freedoms for all individuals. …. In the past, particularly until the 1950s, it was customary for women to wear hats when they ventured outside as it was deemed inappropriate to expose or reveal their hair in public. Such rules were selected by the old society, guaranteeing that a woman adhere to societal norms and expectations."

Example 8: Excerpt from a publication in a Thai national journal

"Lecturing can guarantee teaching, but it might not necessarily confirm students' learning. In English language teaching, a number of teaching methods including the grammar translation method, the audio-lingual method, the silent way, and the communicative approach, were developed throughout the last century to **ensure** students' learning and mastering of the language at a level where they can communicate effectively. Actually, every type of teaching method is different from another, but there should be one thing in common: that is, any classroom instruction should guarantee students' learning, not the teacher's teaching."

Example 9: Excerpt from a publication in an international journal

"Another important criterion in the selection of those nouns was to **ensure** that they could activate a range of EAP collocations rather than a single target collocation. They were chosen among the fifty most frequent nouns in AVL, so it can be assumed that all the participants taking part in the experiment would be familiar with them. Another important criterion in the selection of those nouns was to **ensure** that they could activate a range of EAP collocations rather than a single target collocation. The nouns were thus presented within contexts pertaining to gapped academic English concordances from COCA_ac. These were piloted with two experienced academic writers, and a few adjustments were made to **ensure** the test items elicited the data anticipated."

RQ4. What is the lexical density ratio in research papers published in international academic journals, Thai national academic journals, and those authored by English major students?

Content words and function words used in each research paper were examined using the AntWordProfiler program. Once the count of content words and function words for each research paper was obtained, the Dense Ratio value was calculated. In the final step, the average Dense Ratio values among groups were compared. The analysis results are presented in Table 4.

Table 4

Lexical Density of Research Papers from Three Groups

Lexical Density	%	Students' papers	National papers	International papers
Dense Ratio	Max	71.68	69.77	67.14
	Min	48.69	46.44	48.64
	Average	**57.59**	**57.19**	**57.52**

Table 4 indicates that the research papers authored by English major students and those published in Thai national journals and international journals had similar ratios of content words to function words, with proportions of 57.59%, 57.19%, and 57.52%, respectively. Notably, even the lowest lexical density among the three groups (48.69%, 46.44%, and 48.64%) surpassed the findings in Ure (1971), which suggests that written texts generally have a lexical density of approximately 40% or higher.

5. Discussion, Conclusion, and Recommendations

Based on the lexical profiling results, the analysis revealed varying proportions of vocabulary from the GSL across publications in international journals, Thai journals, and students' research papers. Notably, students' research papers had the highest average coverage of the GSL vocabulary (79.8%), followed by Thai journals (76.41%) and international journals (75.21%).The average

proportion of the GSL vocabulary typically constitutes around 80% of the text, as stated by Nation and Waring (1997), since this list consists primarily of common vocabulary. Across these three groups, the average usage of the GSL vocabulary falls below the threshold, indicating that all three groups utilized a lower proportion of the GSL vocabulary than expected and tended to incorporate vocabulary from other lexical profiles. The findings regarding the AWL indicate that as anticipated, the research papers authored by students had the lowest level of average academic vocabulary profile, with a coverage of only 8.11%. Conversely, publications in Thai journals had an average coverage of 10.96%, while publications in international journals produced the highest average AWL coverage at 12.01%. As noted by Coxhead (2000), it is considered optimal for the AWL vocabulary to comprise approximately 10% or more of the academic texts. Thus, the substantial coverage of the AWL in international journals highlights the importance of using academic vocabulary extensively when crafting research papers. It is worth noting that among the three levels of research, only the students' research papers fell below the recommended threshold for AWL vocabulary. Consequently, the researchers propose the inclusion of supplementary courses in academic English writing for students enrolled in research methodology courses as part of their undergraduate programs. An example of adjusting and enhancing the instruction of research paper writing skills for undergraduate students in all majors undergoing research methodology courses, such as English and non-English majors, is to incorporate specialized content related to teaching research paper writing in the final 2–3 weeks of the course. Additionally, the use of vocabulary lists associated with academic language, such as the Academic Word List (Coxhead, 2000), can be implemented as supplementary material for students to study outside the classroom. This approach should aid students in developing their academic vocabulary skills, contributing to more effective use of academic language in research paper writing. In addition, the relevant government ministry overseeing education in the country should promote the inclusion of research paper writing as a key topic in research methodology courses to ensure that research papers are of high quality and are accepted for publication in reputable

academic journals. Focusing solely on research knowledge may equip learners with research skills and rational thinking. However, lacking high-quality academic language proficiency, researchers may struggle to disseminate the knowledge they generate to the global community.

The findings from the lexical profiling highlight the importance of the utilization of academic vocabulary in research writing. As emphasized by Brun-Mercer and Zimmerman (2016), authors should incorporate a greater proportion of academic terminology in their research articles compared to general vocabulary. This not only enhances the credibility of the research but also ensures the appropriateness of the scholarly discourse. Furthermore, Phillips (1986) underscores the crucial role of the AWL in developing the writer's proficiency in producing scholarly articles. Thus, adhering to the AWL becomes an essential aspect of enhancing the writer's skillset in scholarly writing.

Regarding the findings of the lexical level analysis, as noted by Schmitt and Schmitt (2014), the researchers categorized English words into three main frequency categories: high-, mid-, and low-frequency words. In the present study, the focus was specifically on the medium-frequency words (K4–K9) due to their high relevance to research writing. These words offer a balance between being sufficiently complex to demonstrate vocabulary proficiency and avoiding the simplicity of high-frequency words, as well as the potential difficulty posed by low-frequency words, which may involve proper names or highly specialized terminology. Furthermore, the average usage of medium-frequency words in publications in international journals will serve as a benchmark for comparing students' research papers and publications in Thai journals, as this level of research is widely regarded as a reliable and reputable standard. The investigation in the present study of the utilization of high-frequency words (K1–K3) in publications from international journals, Thai journals, and students' research papers revealed that students' research papers had the highest average coverage of high-frequency words (94.54%), followed by publications in Thai journals

(94.67%), while publications in international journals had the lowest average coverage (94.15%). Importantly, research papers across all three groups used comparable numbers of common terms, underscoring the continued importance of high-frequency words in effectively communicating the main ideas in research paper content. Shifting our focus to the medium-frequency words (K4–K9), the average coverage of these words was moderately comprehensive in research papers. Publications in Thai journals had the lowest average coverage (3.32%), followed by students' research papers (3.85%) and publications in international journals with the highest average (4.17%). This level of terminology garnered specific attention from the researchers. It was observed that students' research papers and publications in Thai journals used mid-frequency vocabulary less frequently compared to publications in international journals. The following observation is noteworthy: comparing the average coverage of mid-frequency words between research papers authored by students and those published in national journals, on average, the students use mid-frequency words more frequently in a higher proportion than researchers who publish national-level research papers. This may lead to the understanding that students use higher-level vocabulary compared to Thai researchers. However, considering the average coverage of low-frequency words, it is found that conversely, Thai researchers use low-frequency words in a higher proportion than students. According to Schmitt (2000), low-frequency words often consist of technical words. This may imply that the reason Thai researchers use mid-frequency words less frequently than students is because they focus more on using technical words in research writing. However, the average use of vocabulary levels in research papers published in international journals indicates that it is advisable to focus on using mid-frequency words to reflect the quality of academic language use. Furthermore, it is recommended to reduce the use of low-frequency words, which may cause readers from outside the field to experience difficulty in understanding the research.

Consequently, it is recommended that students enrolled in undergraduate programs prioritize the inclusion of intermediate frequency terminology in their research writing endeavors. In was observed that the average coverage of low-

frequency words (K10–K25 and off-list) was the least comprehensive in research papers across the three groups. Students' research papers had the lowest average coverage (1.29%), followed by publications in international journals with a moderate average (1.52%), and publications in Thai journals had the highest average (1.59%). However, it is important to note that the utilization of low-frequency words cannot serve as a sole indicator of an author's proficiency in academic vocabulary, as this word group primarily consists of proper names and highly specialized terms specific to particular fields. In light of the lexical level, it is evident that the utilization of medium-frequency terminology in research writing holds great importance. Medium-frequency words, being situated at a vocabulary level that strikes a balance between simplicity and complexity, are deemed essential for developing a strong command of the English language, as supported by Schmitt (2000). Therefore, it is advisable to engage with more challenging vocabulary at the intermediate level and to apply it in research endeavors to enhance the quality of scholarly work. By incorporating more advanced terminology, especially medium-frequency words, researchers can further elevate the academic rigour and sophistication of their studies.

In terms of the lexical variation, a higher value typically indicates a larger vocabulary size for a language user, as measured by the Type-Token Ratio (TTR) calculation. However, when it comes to writing research papers, the focus lies more on the consistency of word usage rather than the breadth of vocabulary. Therefore, a lower lexical variation value is considered more desirable in scholarly discourse. Publications in international journals had the highest level of consistency regarding the lexical variation results. This could be attributed to the use of challenging terminology and a consistent application of original terminology, resulting in a consistency percentage of 21.69%. Such consistency surpasses that of publications in Thai journals and students' research papers. In addition, students' research papers had a higher degree of word usage consistency than research papers in Thai journals, suggesting that the syllabus for students' research papers has effectively fostered word consistency. Notably, publications

in Thai journals had the lowest degree of word repetition at 25.01%. However, both Thai journal publications and students' research papers still lack the level of consistency found in international journal research. It is crucial to emphasize the use of the same words, particularly at a more challenging level, without altering the vocabulary throughout the research process.

Based on the findings from the lexical density analysis, there was minimal variation among the three groups, with proportions of 57.59%, 57.19%, and 57.52% for research papers authored by English major students, publications in Thai national journals, and publications in international journals, respectively. According to Ure (1971), language users typically have a lexical density below 40% in spoken contexts, whereas written texts generally have a lexical density of approximately 40% or higher. The present results indicated that research papers authored by English major students, as well as publications in both national and international journals, had a lexical density of around 57%, surpassing the findings by Ure. Therefore, among the four aspects of lexical proficiency evaluation in the present study, it is only in terms of lexical density that papers authored by English major students and publications in national journals meet the guidelines set by publications in international journals.

The established proportions of lexical profile, lexical level, lexical variation, and lexical density utilized in publications in international journals can serve as valuable guidelines for authors when composing research papers intended for publication in academic journals. To facilitate learners and novice researchers, one recommended approach is to calculate the values of lexical profile, lexical level, and lexical variation for their self-written research papers. Then, these values can be compared with those obtained from publications in international journals, as identified in the present study. By examining their word choices in relation to those found in international research papers, students can make appropriate adjustments to enhance the academic tone and sophistication of their own writing. For practicality, if students or novice researchers analyze the lexical profile, lexical

level, and lexical variation using the tools specified in the present study, they may find that the rate of vocabulary usage in various aspects is lower than the standard set by research papers published in international journals. They might consider revising vocabulary usage by substituting words with similar meanings in the GSL with words from the AWL, such as using 'appropriate' instead of 'right' or 'suitable,' and so forth. This iterative process empowers students and aspiring researchers to expand their lexical repertoire, particularly within the academic domain, while concurrently honing their ability to incorporate challenging and scholarly vocabulary into their writing.

6. Limitations of the Present Study and Recommendations for Future Studies

The first limitation of this study was its narrow focus on assessing vocabulary usage in research papers. Whilst this aspect provides valuable insights, it is important to acknowledge that there are other linguistic features that contribute to the overall quality of academic writing. Therefore, future studies could consider examining additional aspects, such as the usage of academic lexical bundles, which are recurrent word combinations that play a crucial role in conveying disciplinary knowledge and fostering effective communication in scholarly writing.

The second limitation of the current study was the potential presence of direct quotations and examples of language presented in the research findings section of the collected research article, which may have result in slight inaccuracies in analyzing the four aspects of vocabulary use.

Another limitation was that the research papers examined in this study were sourced from journals pertaining to Linguistics, English for Specific Purposes, English Language Teaching, and related areas within English Studies. Notably, the proportion of vocabulary usage, particularly in terms of lexical profile and lexical level, may vary across different academic disciplines (Chung & Nation, 2003; Hyland & Tse, 2007). Therefore, future research investigating the lexical profile,

lexical level, lexical variation, and lexical density in other fields could provide valuable insights into this matter. Expanding the scope of inquiry beyond the current domains should contribute to a more comprehensive understanding of the role of vocabulary in academic writing.

7. About the Authors

Donlawat Meebangsai, Pawarit Pongtin, and Panjaphon Kitipoontanakorn are graduates of the English for Service Industry Program at the Faculty of Hospitality Industry, Kasetsart University, Nakhon Pathom, Thailand.

Piyapong Laosrirattanachai is an Assistant Professor at the Faculty of Hospitality Industry, Kasetsart University, Nakhon Pathom, Thailand. He obtained his Ph.D. in Linguistics from Srinakharinwirot University, Thailand. His research focuses on various areas including vocabulary teaching and learning, development of word lists and lexical bundles, corpus linguistics, and genre and discourse analysis.

8. References

Anthony, L. (2022). *AntWordProfiler* (Windows 64-bit version 1.5.1) [Computer software]. Laurence Anthony's Website. Waseda University. https://www.laurenceanthony.net/software.html

Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education, III*(3), 21–34.

Arunvong Na Ayutthaya, J., Kunthonjinda, K., Somwang, K., & Laosrirattanachai, P. (2022). Making beverage service word list for English for specific purposes classroom. *rEFLections, 29*(2), 325–343. https://doi.org/10.61508/refl.v29i2.259524

Baker, P. (2006). *Using corpora in discourse analysis*. Continuum.

Bates, E., Bretherton, I., & Snyder, L. (1988). *From first words to grammar: Individual differences and dissociable mechanisms*. Cambridge University Press.

Brun-Mercer, N., & Zimmerman, C. B. (2015). Fostering academic vocabulary use in writing. *The CATESOL Journal, 27*(1), 131–148.

Campion, M. E., & Elley, W. B. (1971). *An academic vocabulary list*. New Zealand Council for Educational Research.

Chung, T. M., & Nation, P. (2003). Technical vocabulary in specialized texts. *Reading in a Foreign Language, 15*(2), 103–116.

Cobb, T. (2022). *Vocabprofile*. [Computer program]. http://www.lextutor.ca/vp/

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238. https://doi.org/10.2307/3587951

Coxhead, A., & Hirsh, D. (2007). A pilot science-specific word list. *Revue Francaise de Linguistique Appliquee, XII*(2), 65–78.

Coxhead, A., Nation, I. S. P., & Sim, D. (2014). Creating and trialling six versions of the Vocabulary Size Test. *TESOLANZ Journal, 22*, 13–27.

Davis, B. H., & Brewer, J. P. (1997). *Electronic discourse: Linguistic individuals in virtual space*. State University of New York Press.

Geeraerts, D. (1993). Generalized onomasiological salience. *Belgian Journal of Linguistics, 8*(1), 43–56. https://doi.org/10.1075/bjl.8.04gee

Ghadessy, M. (1979). Frequency counts, words lists, and materials preparation: A new approach. *English Language Teaching Forum, 17*(1), 24–27.

Halliday, M. A. K. (1985). *Spoken and written language*. Deakin University Press.

Harzing, A.-W. (2013). A preliminary test of Google Scholar as a source for citation data: A longitudinal study of Nobel Prize winners. *Scientometrics, 94*(3), 1057–1075. https://doi.org/10.1007/s11192-012-0777-7

Hug, S. E., Ochsner M., & Daniel H. (2013). Criteria for assessing research quality in the humanities: A Delphi study among scholars of English literature, German literature and art history. *Research Evaluation, 22*(5), 369–383.

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

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Working Papers, 53*, 61–79.

Kondal, B. (2015). Effects of lexical density and lexical variety in language performance and proficiency. *International Journal of IT, Engineering and Applied Sciences Research, 4*(10), 25–29.

Laosrirattanachai, P., & Laosrirattanachai, P. (2021). Applying lexical profiling to construct technical word lists for Thai tourist guides. *PASAA, 62*, 61–91.

Laosrirattanachai, P., & Laosrirattanachai, P. (2023). Analysis of vocabulary use and move structures of the World Health Organization Emergencies press conferences on Coronavirus Disease: A corpus-based investigation. *LEARN Journal: Language Education and Acquisition Research Network, 16*(1), 121–146.

Laosrirattanachai, P., & Ruangjaroon, S. (2021). Corpus-based creation of tourism, hotel, and airline business word lists. *LEARN Journal: Language Education and Acquisition Research Network, 14*(1), 50–86.

Laufer, B., & Nation, P. (1995) Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*(3), 307–322.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22*(1), 15–30.

Lestari, I. W., & Hardiyanti, N. (2020). Vocabulary learning autonomy through incorporation of English songs: Indonesian EFL students' perspectives. *3L:*

The Southeast Asian Journal of English Language Studies, 26(2), 91–104. http://doi.org/10.17576/3L-2020-2602-07

Lieven, E. V. M. (1978). Conversations between mothers and young children: Individual differences and their possible implication for the study of child language learning. In N. Waterson & C. Snow (Eds.), *The Development of Communication: Social and Pragmatic Factors in Language Acquisition* (pp. 173–187). Wiley.

Llach, M. P. A. (2014). Exploring the lexical profile of young CLIL learners: Towards an improvement in lexical use. *Journal of Immersion and Content-Based Language Education, 2*(1), 53–73. https://doi.org/10.1075/jicb.2.1.03agu

Lynn, R. (1973). Preparing word lists: A suggested method. *RELC Journal, 4*(1), 25–32.

McLean, S., & Kramer, B. (2015). The creation of a new vocabulary levels test. *Shiken, 19*(2), 1–11. https://www.lextutor.ca/tests/levels/recognition/nvlt/paper.pdf

McLean, S., Kramer, B., & Beglar, D. (2015). The creation and validation of a listening vocabulary levels test. *Language Teaching Research, 19*(6), 741–760. https://doi.org/10.1177/1362168814567889

Mozaffari, A., & R. Moini. (2014). Academic words in education research articles: A corpus study. *Procedia - Social and Behavioral Sciences, 98*, 1290–1296. https://doi.org/10.1016/j.sbspro.2014.03.545

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*(1), 91–108. https://doi.org/10.1002/RRQ.011

Namwong, B., Vorakhan, P., Suntorn, W., & Laosrirattanachai, P. (2022). Building a semi-technical political news word list for political science students. *Thoughts, 2*, 1–21.

Nation, I. S. P. (2016). *Making and using word lists for language learning and testing.* John Benjamins Publishing Company.

Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781009093873

Nation, P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9–13.

Nation, I. S. P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6–19). Cambridge University Press.

Nguy, N. L. Q., & Ha, H. T. (2022). Lexical profile of academic written English revisited: What does it take to understand scholarly abstracts? *SAGE Open, 12*(3), 1–12. https://doi.org/10.1177/21582440221126342

Öztürk, M. (2022). Parts of speech distribution in the BNC-COCA word lists. *Language, Education & Technology, 2*(2), 128–140.

Phillips, J. P. N. (1986). Essay-writing phobia in undergraduates. *Behaviour Research and Therapy, 24*(5), 603–604. https://doi.org/10.1016/0005-7967(86)90042-2

Phung, D. H., & Ha, H. T. (2022). Vocabulary demands of IELTS listening test: An in-depth analysis. *SAGE Open, 12*(1), 1–13. https://doi.org/10.1177/21582440221079934

Praninskas, J. (1972). *American university word list*. Longman.

Rungrueang, T., Boonprasert, P., Poempongsajaroen, S., & Laosrirattanachai, P. (2022). Corpus-based approach to generate a word list for food service. *THAITESOL Journal, 35*(1), 57–76.

Schmitt, N. (1997). Vocabulary learning strategies. In D. N. Schmitt, & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 199–227). Cambridge University Press.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University Press.

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*(1), 26–43. https://doi.org/10.1111/j.1540-4781.2011.01146.x

Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching, 47*(4), 484–503. https://doi.org/10.1017/S0261444812000018

Thornbury, S. & Slade, D. (2006). *Conversation: From description to pedagogy.* Cambridge University Press.

Ure, J. (1971). Lexical density and register differentiation. *Contemporary Educational Psychology*, *5*, 96–104.

Webb, S., & Paribakht, T. S. (2015). What is the relationship between the lexical profile of test items and performance on a standardized English proficiency test? *English for Specific Purposes, 38,* 34–43. https://doi.org/10.1016/j.esp.2014.11.001

West, M. (1953). *A general service list of English words*. Longman.

Wilkins, D. (1972). *Linguistics in language teaching*. Edward Arnold.

Wolf-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity.* University of Hawai'i, Second Language Teaching & Curriculum Center.

Xue, G., & Nation, P. (1984). A university word list. *Language Learning and Communication, 3*(2), 215–229.