A CORPUS-BASED VOCABULARY ANALYSIS OF FIRST-YEAR UNDERGRADUATE TEXTBOOKS OF AN INTERNATIONAL PROGRAM IN ECONOMICS

BY

MR. PONGSATHON WASIKARAT

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN ENGLISH LANGUAGE TEACHING (INTERNATIONAL PROGRAM)

LANGUAGE INSTITUTE

THAMMASAT UNIVERSITY

ACADEMIC YEAR 2023

A CORPUS-BASED VOCABULARY ANALYSIS OF FIRST-YEAR UNDERGRADUATE TEXTBOOKS OF AN INTERNATIONAL PROGRAM IN ECONOMICS

BY

MR. PONGSATHON WASIKARAT

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN ENGLISH LANGUAGE TEACHING (INTERNATIONAL PROGRAM)

LANGUAGE INSTITUTE

THAMMASAT UNIVERSITY

ACADEMIC YEAR 2023

THAMMASAT UNIVERSITY
LANGUAGE INSTITUTE

DISSERTATION

BY

MR. PONGSATHON WASIKARAT

ENTITLED

A CORPUS-BASED VOCABULARY ANALYSIS OF FIRST-YEAR UNDERGRADUATE TEXTBOOKS OF AN INTERNATIONAL PROGRAM IN ECONOMICS

was approved as partial fulfillment of the requirements for
the degree of Doctor of Philosophy in English Language Teaching
on April 22nd, 2024

Chair _____
(Assistant Professor Apasara Chinwonno, Ph.D.)

Member and Advisor _____
(Assistant Professor Kittitouch Soontornwipast, Ed.D.)

Member _____
(Associate Professor Supong Tangkiengsirisin, Ph.D.)

Member _____
(Ajarn Panna Chaturongakul, Ph.D.)

Member _____
(Ajarn Chanika Gampper, Ph.D.)

Director _____
(Associate Professor Supakorn Phoocharoensil, Ph.D.)

Dissertation Title	A CORPUS-BASED VOCABULARY ANALYSIS OF FIRST-YEAR UNDERGRADUATE TEXTBOOKS OF AN INTERNATIONAL PROGRAM IN ECONOMICS
Author	Mr. Pongsathon Wasikarat
Degree	Doctor of Philosophy
Major	English Language Teaching
Field/Faculty/University	Language Institute Thammasat University
Dissertation Advisor	Assistant Professor Kittitouch Soontornwipast, Ed.D.
Academic Year	2023

ABSTRACT

The purpose of this current study was to (1) investigate the text coverage that provided by the BNC/COCA Word Family Lists (Nation, 2017) and the Academic Word List (Coxhead, 2000) in the first-year Bachelor of Economics International Program (BEIP) textbooks, (2) investigate the vocabulary size required to read the BEIP textbooks, and (3) identify the AWL academic words that occur most frequently in the BEIP textbooks. A corpus of 1,343,493 words from the economics course textbooks was compiled into the ECON corpus and then analyzed using the AntWordProfiler software program (Anthony, 2020). It was found that the 4th 1,000 BNC/COCA word family lists reached 96.10% and the 9th 1000 BNC/COCA word family lists reached 98.11% in the ECON corpus. The results in this current study implied that a vocabulary size of around 3,500–4,000 word families is needed for reading the BEIP textbooks that requires 95% coverage for reasonable comprehension, while a vocabulary size of approximately 8,500–9,000 word families is needed for reading the BEIP textbooks that requires 98% coverage for optimal comprehension. Moreover, the AWL accounted for 9.62% of the text coverage, while the first and second 1,000-word levels of the GSL gave a cumulative coverage of 79.71%. Regarding the most frequently occurring words, AWL academic words must be selected using Coxhead's (2000)

criteria, with the word frequency of at least 38 times and 4 times in each subcorpus. In this current study, there were 187 frequently-occurring AWL academic words and 133 AWL academic family headwords. The results of the ECON corpus compilation suggest curriculum developers and materials designers include these words and emphasize their importance in an economics textbook for effective word selection. The findings also suggest implications for instructional practices, and recommendations for future research.

Keywords: Text coverage, vocabulary size, word frequency, economics textbooks, word families, BNC/COCA word family lists, AWL

ACKNOWLEDGEMENTS

This Ph.D. journey was a profoundly transformative life experience which resulted from long and constructive discussions. I would like to start by expressing my sincere gratitude and appreciation to my dissertation advisor, Assistant Professor Kittitouch Soontornwipast, Ed.D., for the intellectual supervision, extra-high level of time contribution, and caring he tirelessly demonstrated during the milestones of my dissertation as follows: the qualifying examination (QE), the dissertation proposal defense, and the final dissertation defense. He took time to review and demystify my drafts, gave me critical advice and encouragement with a perfect blend of insight and humor, and even rehearsed the timing for slide show presentations. He is exceptional and very supportive. His dedication goes far beyond what is required of a dissertation advisor. However, I would like to apologize to him for not being able to work very well under pressure while working on the dissertation. My classmates always say that being supervised by him is a great blessing for me. He is undoubtedly a treasure among professors. The completion of my dissertation would not have been possible without him.

Secondly, my family is my greatest blessing in life. Without my family's encouragement, I could never have achieved this great goal in my educational journey. Their belief in me has kept my spirits and motivation high throughout my studies. This journey would not have been possible without the unconditional love and eternal support of my loved ones, including my father, Mr. Somsak Wasikarat; my mother, Mrs. Pensri Wasikarat; and my aunt, Ms. Somying Konpha. As this study is dedicated to them, family is everything, always and forever.

Furthermore, I would like to express my sincere gratitude to all the respected members of the Ph.D. dissertation committee, as follows: Assistant Professor Apasara Chinwonno, Ph.D., Associate Professor Supong Tangkiengsirisin, Ph.D., Ajarn Panna Chaturongakul, Ph.D., and Ajarn Chanika Gampper, Ph.D., for generously providing knowledge, expertise, and academic guidance throughout the dissertation. Without them, I would not have been able to take on this epic journey. I would also like to thank all the professors at the Language Institute of Thammasat University and all

the visiting professors for the classes that keep on inspiring us with knowledge and wisdom. Also, I would like to extend my gratitude to Associate Professor Supakorn Phoocharoensil, Ph.D., for always supporting and encouraging me during my time in the corpus linguistics course and Ph.D. program. My special thanks for the corpus linguistics classes are extended to Assistant Professor Passapong Sripicharn, Ph.D., and Associate Professor Raksangob Wijitsopon, Ph.D. In addition, I would like to acknowledge the support I received from Ms. Kobkorn Panmuang, or P-Pang, and Ms. Jiraporn Petchthong, or P-Add, for their excellent academic services and student support in the Ph.D. program.

Friends make life memorable; the friendship of Ph.D. classmates (the sixth batch) is a good example of this. It has been able to drive my growth and academic success in life. I would like to express my heartfelt gratitude to Ajarn Chonrachai Ketworrachai, Ph.D., an English lecturer at Assumption University (ABAC), for his constant support in every exam preparation and presentation rehearsal. He did not hesitate to share lecture notes or anything else. His boundless generosity will not be forgotten by his classmates. I am also grateful to be friends with Ajarn Poonyapat Boonyarattanasoontorn, Ph.D., an English instructor at Srinakharinwirot University. I admire her exemplary behavior as a role model of hardworking instructors, as she worked hard day and night on her dissertation while also working full-time. I would also like to thank the little sister of our group, Ajarn Tivawan Phramphun, Ph.D., an English instructor at Srinakharinwirot University, for friendship and her moral support. P-Kwan, or Ajarn Keeratikan Fuangkarn, Ph.D., the head of the foreign language department at Prince Royal's College in Chiang Mai, is a reliable, unforgettable friend from northern Thailand. I appreciate how good she is with people and that so many people love her. I appreciate having Ajarn Nateethorn Narkprom, Ph.D., an English instructor at Phetchabun Rajabhat University, as my only classmate in the Corpus Linguistics course. The Dream Team, sometimes called the Nightmare Team, will never be forgotten. Ajarn Orabudh Mayanondha, Ph.D., nicknamed Kru Soi, is one of the members. She is brilliant and super productive, having earned the batch's first Ph.D. while holding down a full-time job. Ajarn Sophia Arwae, Ph.D., a full-time English teacher at Kamnoetvidya Science Academy (KVIS), is also part of the Dream Team. I am so grateful for her constant encouragement.

I would like to thank Nakhon Phanom University (NPU) for the fully-funded financial support and employee scholarship benefits during the study in fiscal years 2017–2019, as I was on study leave for three and a half years during my doctoral studies. My thanks go to Ajarn Thanjira Srikachang, a kind neighbor from the Department of Chinese, for her food support. I am also grateful to Ajarn Phatcharaphong Phubetpeerawat, Ph.D., my colleague since I started working at NPU, for his invaluable support when I was sick or had problems in Nakhon Phanom.

Aristotle said, “The roots of education are bitter, but the fruit is sweet.” No matter how many years and months I spent on my research, learning new things was always the sweetest in the journey. Thammasat University, more specifically the Language Institute of Thammasat University (LITU), has become my second home since the day I started my master’s degree here. LITU is not just a group of campus buildings. It is a place of light, civility, and learning with a family atmosphere, study spaces, a wide range of student support services and facilities, and learning resources. In addition, I did not decide to pursue a Ph.D. immediately after completing my master’s degree. Otherwise, I worked as an English instructor for 3 years before the study leave. It is tedious and challenging to earn a Ph.D. If it were easy, everyone would do it. Although I am a slow learner, I am grateful to myself that I continue to grow and appreciate every moment of what I have gone through since the first day of study. I am thankful for everything I have been through today and always.

Mr. Pongsathon Wasikarat

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ABBREVIATIONS

Symbols/Abbreviations	Terms
AEC	ASEAN Economic Community
ASEAN	Association of Southeast Asian Nations
AWL	Academic Word List
BEIP	Bachelor of Economics International Program
BNC	British National Corpus
CAL	Calculus for Social Science 1
CEFR	Common European Framework of Reference for Languages
CLIL	Content and Language Integrated Learning
COCA	Corpus of Contemporary American English
CU-TEP	Chulalongkorn University Test of English Proficiency
ECON	Economics
EFL	English as a Foreign Language
ELF	English as a Lingua Franca
ELT	English Language Teaching
EMI	English-Medium Instruction
GSL	General Service List
MIC	Principles of Microeconomics
MAC	Principles of Macroeconomics
STA	Statistics for Social Science 1

CHAPTER 1
INTRODUCTION

In this chapter, the background of the study is explained, starting from the importance of the English language, scoped down into the English Language and internationalization in Thailand, followed by the problem statements, the research objectives, the research questions, the scope of the study, the definition of terms, the significance of the study, the organization of the study, and the limitations.

1.1 Background of the Study

1.1.1 Importance of the English Language

The English language is considered one of the most widely spoken languages in the world. With the number of non-native English speakers currently outnumbering native English speakers, English as a Lingua Franca (ELF) has become more prominent as a means of international communication between speakers whose native language is different in various fields such as business, diplomacy, and education, to mention but a few. Global companies require their employees to have a certain level of English proficiency and skills in order to communicate effectively, especially when dealing with international clients and business partners. As the Association of Southeast Asian Nations (ASEAN) has established the economic integration of the ASEAN Economic Community (AEC) since 2015, English has become a common working language. This significantly affects the language policies of individual member countries. For example, Singapore is one of the ASEAN countries where English proficiency is actively promoted through a bilingual policy, namely an 'English + 1' policy, by adopting English as a medium of instruction along with native language learning to ensure the preservation of Singapore's ethnic identities (Kirkpatrick, 2010; Wee et al., 2010). English proficiency is also critical for Thailand's economic competitiveness and for most Thai students (Chatsawat, 2013; Jaiprasertsook, 2014). However, current teaching approaches aimed at preparing students for AEC are not effective enough; in other words, there is a need for practical curriculum reform in English teaching (Bunwirat, 2017; Sanonguthai, 2014). To

overcome the learning problems, Thai undergraduate students themselves needed activities such as self-study, speaking with foreigners, and visiting abroad (Jindathai, 2013). Therefore, it could be concluded that they were aware of the need for authentic English.

1.1.2 The English Language and Internationalization in Thailand

In the wake of educational reform, there is a new wave of interest in English-language degree programs. This is due to the requirements of the Thai education system, which demands competent English proficiency. Despite not having an official language status, the critical importance of teaching English as a required subject to all students in Thai primary and secondary schools from grades 1 through 12 has been highlighted by the English language. According to the policy of Thailand's Ministry of Education, many schools offer a bilingual English-Thai program, often referred to briefly as the 'English program' among Thais, which is better known. In other words, each subject in an international or English program—aside from Thai language and social studies—uses English as the primary medium of instruction.

Thailand's Ministry of Education has long sought to promote English language teaching through government policy. However, due to challenges and societal obstacles, the implementation of English language reform policies has produced seemingly unsatisfactory results (Kaur & Young, 2016). Since 2014, the Ministry of Education has officially adopted the Common European Framework of Reference for Languages (CEFR) to raise the educational standards for English. Thailand was rated in the low proficiency band in the EF English Proficiency Index for 2017, and in the following year, the rank was classified as lower (Mala, 2018). In addition, the CEFR-based language proficiency levels of Thai undergraduate students are at the Basic User (A1 and A2) level rather than the Independent User (B2) level as expected by the Ministry of Education (Waluyo, 2019). The CEFR is viewed as an assessment tool rather than a language skill development tool (Foley, 2019). For instance, it is recommended that Thai university students have an acceptable English proficiency level for admission and graduation. A large number of Thai students who completed the Chulalongkorn University Test of English Proficiency (CU-TEP) in science and non-science disciplines were below the standards needed for graduate-level study in Western universities (Prapphal, 2003). Similar to Padermprach's (2017) study, English

is perceived as a very important area of study, but it does not meet international standards.

Generally, higher education programs are authorized to offer courses based on either institutional accountability and public policy needs or student-related demands (Kaewphanngam & Tharmmapornphilas, 2013). The number of English-taught degree programs, also referred to as 'international programs' or literally 'English programs' in Thailand, has increased rapidly due to educational offerings and reform policies on internationalization. As more institutions develop international degree programs to promote internationalization, accredited Thai universities offer international and English undergraduate degree programs in major disciplines such as engineering, architecture, science, and business administration (see Appendix A for the international and English programs offered in Thai accredited higher education institutes). According to the QS World University Rankings 2023, Thai accredited colleges and universities in Thailand offer a great number of international and English programs across more diverse, specialized disciplines for 21st century education, including digital innovation, environmental engineering (Chiang Mai University, n.d.), arts and science in integrated innovation, nano-engineering, robotics and artificial intelligence engineering (Chulalongkorn University, n.d.), creative media technology, digital media engineering (Khon Kaen University, n.d.), to name just a few.

The internationalization of higher education institutions has reciprocally influenced national education policies. In non-English-speaking peripheral countries, the importance of English and EMI practices is emphasized in terms of personal endeavors and institutional internationalization (Hu et al., 2013). Higher education institutions claim to train graduates for international professions. Chalapati (2007) highlighted that international degree programs operated in association with Western universities support the Thai government's strategic goal of internationalization in order to bring the country's higher education system on a pace with other developed nations' standards. Significantly, university ranking systems are also driven by economic considerations through internationalization policies (Cavanagh, 2016).

Since there is a reciprocal relationship between English language learning and subject content learning, some teaching approaches on a continuum of

English language teaching (ELT) are necessarily interrelated, such as content-based language teaching (CBLT), content and language integrated learning (CLIL), English for academic purposes (EAP), English-medium instruction (EMI), and English for specific purposes (ESP), as proposed. There have been some complications, such as CLIL and EMI being used as umbrella terms on the continuum due to different contexts of EFL/ESL or regional variations (Bradford & Brown, 2016).

In Thailand, most international and English programs are EMI, which focus on content and are taught by content specialists. English-medium instruction, also known as EMI, is defined as "the use of the English language to teach academic subjects in countries or jurisdictions where the first language of the majority of the population is not English" (Dearden, 2014, p. 4). EMI, as used in this current study, is the method of teaching students in academic disciplines including science, math, literature, engineering, history, and geography using the English language, particularly in countries where English is not an official language. It is an increasingly educational trend and plays a critical role across the world in achieving a country's globalization and internationalization policies. The growing popularity of EMI programs in Thailand is also built upon the country's educational internationalization policies. In other words, internationalization results in improved standardization for an institute's international recognition and/or global university rankings. EMI universities increase international student enrollment, prepare domestic students for either overseas academic study or professional work, publish in English, and are more competitive in rankings. The goal is for students to become fully immersed in the language that will consequently be used for instruction and discourse.

According to the EMI programs in Thailand, many universities offer international and English undergraduate programs in economics, for example, (1) the Bachelor of Economics in Business Economics at Assumption University; (2) the Bachelor of Economics International Program at Chiang Mai University; (3) the Bachelor of Arts in Economics at Chulalongkorn University; (4) the Bachelor of Arts Program in Entrepreneurial Economics at Kasetsart University; (5) the Bachelor of Economics Program at Mae Fah Luang University; (6) the Bachelor of Business Administration Program in Business Economics (International Program) at Mahidol University; (7) the Bachelor of Arts in International Political Economy and

Development at Rangsit University; and (8) the Bachelor of Economics International Program at Thammasat University.

EMI courses in Thailand aim to enhance students' academic English skills, resulting in high proficiency in reading, writing, and interactional English, with significant inter-correlations among students' perceptions. (Tang, 2021). In addition, Chuang (2015) points out that EMI can significantly facilitate student learning. According to Hua's (2019) study, EMI students are encouraged to gain dual benefits as they simultaneously acquire English and content knowledge to better prepare for the global workforce. For them, mastery of content knowledge is the learning goal, while language is seen as a tool to master content knowledge (Hu & Wu, 2020). In addition, a previous research study showed that university students' general English proficiency increased significantly after years in an EMI program, for example, from level B2 to C1 (Vidal & Jarvis, 2020). However, they inevitably encounter a number of associated challenges in EMI classes. The usage of English remains limited even though it is required as the working language for all course-related activities (Jiang et al., 2019).

EMI implementation is difficult to manipulate because of a number of influencing factors. In non-English-speaking peripheral countries, several studies (Chuang, 2015; Hua, 2019; Jiang et al., 2019) have shown that EMI students' limited English proficiency is considered one of the learning challenges commonly encountered in EMI classrooms. English language proficiency of both instructors and students is considered a key indicator in EMI implementation in the Taiwanese context (Tseng et al., 2018). In addition, English language proficiency is considered crucial for students' success at an early stage of higher education. They also considered the integration of English language skills (listening, speaking, reading, and writing) essential (Dimova, 2020). Roche and Harrington (2013) emphasize the importance of vocabulary development in English language skills, with Aizawa & Rose (2020) highlighting the problematic nature of a lack of vocabulary in these skills. Regarding the implementation of EMI for all four English language skills, a perceived difficulty is a lack of vocabulary knowledge. As a result, academic vocabulary size and difficulties with language are significantly correlated with English proficiency. Less difficulty exists for students whose vocabulary is larger (Aizawa & Rose, 2020).

1.1.3 Vocabulary Knowledge and English Language Skills

There are countless vocabulary words in English because the language is so rich. This implies that learning every word in the English language is unlikely for beginners. Learners do not need to have a broad vocabulary as native speakers have in order to communicate effectively in the language, whether for reading, listening, writing, or speaking (Nation & Waring, 2002).

The number of words of one's vocabulary is viewed as a challenge that influences how well one succeeds in academic listening in a second language, according to previous studies (Aizawa & Rose, 2020; Chuenjit, 2011). For instance, the size of vocabulary knowledge had a statistically significant impact on Thai hospital pharmacists' English listening comprehension. In a recent study, Wang (2015) investigated 120 Chinese non-English major undergraduate students at a medical university to see how their vocabulary knowledge and L2 listening comprehension compared. It was found that test results for L2 listening comprehension are significantly influenced by both vocabulary breadth and depth. Regarding the study's pedagogical implications, teachers should help students understand the importance of vocabulary learning for the improvement of their listening skills. In order to broaden their vocabulary knowledge, advanced English proficiency students should have access to more and different reading. After taking the IELTS listening test, 290 Chinese undergraduate students were studied by Li and Zhang (2019) to determine their vocabulary level. Using an auditory vocabulary knowledge assessment, the connection between L2 listening comprehension and the three components of L2 vocabulary knowledge: vocabulary breadth, depth, and fluency, was investigated. Each of the three vocabulary knowledge dimensions significantly predicts L2 listening comprehension. In addition, the amount of L2 auditory vocabulary knowledge is an effective predictor. Students should be aware of both the breadth and depth of their vocabulary knowledge due to the relationship that exists between them. In order to determine the relationship between vocabulary size and listening, 59 Japanese EFL students were asked to listen to four brief sections in a study by Bonk (2000). A free written recall test and a dictation test were among the research instruments. Listening comprehension was found to be improved for learners with 95% text coverage. This shows that students require a strong vocabulary to comprehend listening materials. A large vocabulary is necessary for

watching movies or television shows (Webb & Rodgers, 2009). In addition, Bonk (2000) proposed that listening comprehension with effective coping strategies might be achieved with coverage far below 95%. According to Nation (2006), if the 98% coverage need for comprehension is taken into consideration, written texts require an 8,000–9,000 word-family vocabulary, but auditory texts only require a 6,000–7,000.

A person's ability in writing, reading, listening, and speaking also depend on vocabulary knowledge. It has an impact on the writing's quality and requires learning the definitions of words as well as effective word usage. It is regarded as an essential tool for L2 learners that may help them write better. English uses both spoken and written words to express thoughts, sensations, and feelings. When L2 learners are given a writing task or assignment, having a robust vocabulary is essential so that they can choose an appropriate word from their vocabulary repertoire for the assignment's specific genre. Diversity of vocabulary has often been proven to be a significant predictor of L2 writing performance and growth. Furthermore, the performance of productive skills is predicted by language knowledge (Kilic, 2019). Writing abilities and vocabulary knowledge were shown to be positively and significantly correlated (Firdaus, 2019). Consequently, it has an impact on the writing's quality and requires for understanding word usage and meaning. Vocabulary deficits appear to have a greater negative impact on writing quality than other elements like attitude, motivation, and test readiness (Leki & Carson, 1994).

Learners gain receptive vocabulary knowledge by reading and listening to texts, and they can comprehend language without speaking or writing as a result of it. Typically, teachers require students to basically spell and pronounce words while teaching the meaning of words by means of sentences. This method assists students in comprehending what they are reading on a deeper level. Vocabulary knowledge is a significant indicator of L2 reading comprehension and performance (Shen, 2008; Zhang, 2012). At the university level, the severe vocabulary knowledge deficiencies of Thai undergraduate students may hinder their reading mastery (Ward, 2009). Because there is only a little amount of explicit vocabulary attention throughout the duration of a week of lessons, many teachers continue to neglect the importance of vocabulary (Folse, 2010).

1.2 Problem Statements

While much attention has been paid in the past to language teaching topics such as grammatical competence, contrastive analysis, reading, or writing, the teaching and learning of vocabulary has not received much attention compared to the aforementioned topics (Richards, 1976). How much vocabulary does a student of a second language (L2) need to know, according to Nation and Waring (1997), is a straightforward question that is frequently addressed. The term "vocabulary size" or "vocabulary breadth" can be used to describe it. Without a sufficient vocabulary size, students tend to have learning difficulties. Although a student's L2 learning depends on having a strong vocabulary, not all English words are equally useful. Word frequency can show how frequently a word appears in a particular text or text genre. In the existing literature on corpus-based vocabulary analysis of undergraduate textbooks, it has been found so far that there are many studies conducted independently of an EMI context, e.g., textbooks of English for Engineering (Chanchanglek & Sriussadaporn, 2011), textbooks for English majors (Hajiyeva, 2015), and materials for Business English majors (Liangpanit, 2010). There is not much research in an EMI context, e.g., among computer science undergraduate students (Bi, 2020). Studies on the vocabulary analysis of first-year economics undergraduate textbooks are also scarce.

One possibility for why English language proficiency plays such an important role in academic success at university could be the interconnection between English language proficiency and vocabulary. To be successful with reading comprehension, without getting frustrated and discouraged, an individual needs knowledge of the suggested vocabulary thresholds of a text (Hu and Nation, 2000; Nation, 2001).

As Thailand's higher education becomes more globalized, more institutions are providing content-based bilingual or English-language courses for undergraduate students majoring in a variety of fields. To develop their academic knowledge and abilities, students enrolled in these courses are required to read textbooks written in English.

In general, first-year undergraduate students tend to have more challenges in that they are required to take specialized and academic courses that they are not

familiar with. The first year of study is a new beginning for life at university. Some students either drop out or quit at this early stage of the study. Time and attention are considerably needed. Most of the courses are new, and prerequisite courses must be completed before they enroll in the second-year courses. Having sensible goals for vocabulary learning is crucial for overall English proficiency development and vocabulary development itself. According to the correlation between vocabulary and English language abilities, students increase their vocabulary by understanding what they read and then using those words in their writing and speaking. The BEIP first-year undergraduate textbooks might be intimidating for students to read if they do not have a strong foundation in reading.

Consequently, it is critical to English language specialists, subject content teachers, materials developers, and even the learners themselves to know the text coverage, which may indicate if students can understand written input or the textbooks, since it might be challenging to choose texts for students at their level, which allows instructors to subsequently decide on what kind of vocabulary strategies are needed to employ in class. It serves as an authentic and practical guideline to improve the students' vocabulary knowledge and their overall language proficiency (Mehring, 2006). That is, how teachers or course designers decide which and how many words should be incorporated. A review of the literature shows that there is a need to carry out a study that analyzes the first-year undergraduate textbooks on corpus-based vocabulary analysis.

1.3 Research Objectives

The three main objectives of this study are as follows:

(1) to investigate the text coverage that the BNC/COCA Word Family Lists (Nation, 2017) and the Academic Word List (Coxhead, 2000) provided in the first-year Bachelor of Economics International Program (BEIP) textbooks,

(2) to investigate the vocabulary size required to read the BEIP textbooks, and

(3) to identify the AWL academic words that appear most frequently in the BEIP textbooks.

The goal of this current study is to investigate a corpus-based vocabulary analysis of the BEIP textbooks for the first year. The suggested textbooks are meant to provide as an example of the written materials in English that first-year undergraduate students should be familiar with or aware of in an environment where English is the medium of instruction (EMI). In order to meet their needs for vocabulary learning, the anticipated outcomes also target the most popular academic words included in the BEIP's first-year undergraduate textbooks. By analyzing the text coverage of the BNC/COCA Word Family Lists and the Academic Word List, as well as by selecting words which are most frequently appearing from the BEIP textbooks, the researcher intends to give an overview of the vocabulary required to comprehend the textbooks selected in the BEIP.

1.4 Research Questions

The following research questions were established to focus the investigation based on the study's objectives, which include

(1) What text coverage do the BNC/COCA word family lists and the Academic Word List provide in the first-year undergraduate economics textbooks of the BEIP?

(2) What vocabulary size is needed to read the first-year undergraduate economics textbooks of the BEIP?

(3) What are the most frequently occurring AWL academic words in the first-year undergraduate economics textbooks of the BEIP?

1.5 Scope of the Study

The research study is carried out to investigate the text coverage, estimate the vocabulary size required to read, and identify the AWL academic words that appear most frequently in the BEIP first-year textbooks at a Thai autonomous university. The selected textbooks in this current study are the textbooks used in the economics core courses listed in the first year of the BEIP study curriculum. The BEIP introduces the following economics core courses in the first year of study: Principles of Microeconomics, Principles of Macroeconomics, Calculus for Social Science 1, and Statistics for Social Science 1. In addition, the textbooks exclusively focus on those that

are officially required to be read as part of the curriculum. In the corpus analysis, it focuses on a list of single-word units, which are word families in the discipline of economics. Furthermore, a word family is a set of words that can help increase the number of words an individual can master in vocabulary learning.

1.6 Definition of Terms

(1) Text Coverage

Text coverage in this current study represents the percentage of running words known in the text corpus of the first-year undergraduate economics textbooks of the Bachelor of Economics International Program (BEIP).

(2) Word Frequency

The word frequency refers to the number of occurrences of each word or the rate at which words appear, or are repeated, in a corpus of the first-year undergraduate economics textbooks of the Bachelor of Economics International Program (BEIP).

(3) The BNC/COCA Word Family Lists

In this current study, the BNC/COCA word family lists refer to Nation's (2017) BNC/COCA frequency-based word family lists which contain only the headwords of the 25,000 word families in different word bands or frequency levels.

(4) The Academic Word List

The Academic Word List, or AWL, in this current study refers to Coxhead's (2000) Academic Word List, which has 570 word families that were chosen because they occur frequently in a wide variety of academic texts.

(5) The First-Year Undergraduate Textbooks of the BEIP

The First-Year Undergraduate Textbooks of the BEIP refer to the English-written textbooks of the first year of study of the Bachelor of Economics International Program at an autonomous university in Thailand. The textbooks are used as the core course textbooks. During the investigation, the textbooks of economics core courses are compiled into a corpus and subcorpora of the aforementioned corpus.

(6) Vocabulary Size

Vocabulary size refers to the number of word families that a reader knows in a text and, it has an effect on reading comprehension.

1.7 Significance of the Study

The present study is intended to provide findings from a corpus-based vocabulary analysis of the first-year undergraduate textbooks of the Bachelor of Economics International Program (BEIP) that would be predominantly beneficial to subject content lecturers in economics, English language lecturers, course designers, curriculum and materials developers, and even students themselves. The results are to be considered as another guided resource for developing English proficiency, in particular the vocabulary associated with the first-year undergraduate textbooks in an EMI context.

Firstly, subject content lecturers may benefit from a research-guided economics word list, ensuring that students understand their topics and content knowledge associated with the target words they need to know in the first-year undergraduate textbooks of the BEIP. In order to ensure that the words selected are most meaningful and beneficial for the students of economics, instructors should prioritize their efforts on focused words rather than expanding all vocabulary words for their students throughout class.

Secondly, English language instructors benefit from comprehending how vocabulary is used in the Bachelor of Economics International Program's first-year undergraduate textbooks in terms of text coverage, word frequency, and word distribution, taking into account the use of vocabulary of common words, academic words, and technical words. As English trained professionals, English language lecturers can assist students in learning more of the target words on the list by using a variety of vocabulary teaching strategies, such as comprehensive reading programs, vocabulary games, and paper or electronic flash cards.

Thirdly, course designers, curriculum developers, and materials developers can ensure that it provides criteria for choosing the appropriate materials or textbooks, syllabus selection, and grading target words in economics; therefore, it is a very practical and effective tool (Flowerdew, 1997). The samples of the study, moreover, can be employed as an alternative guideline for designing a preparation course in English for economics to raise a wider range of economics words in their subject content knowledge. Consequently, investigating the corpus-based vocabulary of the

textbooks in a specific-subject content course, regarding text coverage, word frequency, and word distribution, in particular, is worthwhile to see what vocabulary is taught to students and whether the materials are suitable for a specific group of learners.

Last but not least, first-year undergraduate students may gain information from research, particularly those who are English language learners. Regarding the student having doubts about economics vocabulary words during their lessons, they may be assisted by the informants on the list provided. They are also able to use the list to improve their vocabulary size for successful learning.

1.8 Organization of the Study

Chapter 1 provides an introductory overview of the study. It begins with the background of the study and follows with the problem statements. More details are then discussed in the following sections: the research objectives, the research questions, the scope of the study, the definition of terms, the significance of the study, the organization of the study, and the limitations.

Chapter 2 consists discussion of the theoretical frameworks, critical analysis of related research works, and a conclusion.

Chapter 3 describes the data collecting process is described, with a focus on the instruments and methods used to address the research questions.

Chapter 4 comprises the results to answer the research questions of the study.

Chapter 5 presents a summary of the findings, followed by a discussion of the findings, relating it to what other research works discovered, and recommendations based on the findings.

1.9 Limitations

This study has some limitations, as the findings contribute to corpus-based vocabulary analysis for vocabulary development. The researcher then excludes the multi-word units of the first-year undergraduate textbooks of the BEIP because the focus has been on word families as the counting unit of this study. A word family is simply defined as a set of words, for example, help, helper, helpful, unhelpful, helps, helped. Understanding word families can help explain how these words are related and

reflect the use of word families. It also requires the ability to recognize word family patterns. However, future research may investigate multi-word units or other multi-word expressions appearing in textbooks, such as lexical bundles and idiomatic expressions. Additionally, for analyzing the vocabulary in the textbooks, the research study will use data from the BEIP's first-year undergraduate textbooks. Specific words that are often used in texts that are not a part of the observed corpus might be excluded using this method. Although it is referred to as an economics word list, not all economics professionals may find it adequate. Since it only employed textbooks as corpus materials, this study's text corpus was relatively large when compared to those used in other studies, enabling the extraction of key vocabulary. Since many research studies mainly depended on journal articles as a source of corpus texts, it could be different text types in nature.

CHAPTER 2
REVIEW OF LITERATURE

This chapter aims to conceptually review the literature on corpus-based studies, followed by second language (L2) vocabulary learning, vocabulary size, text coverage, the BNC/COCA word family lists, academic vocabulary and the Academic Word List (AWL), and word frequency, word families, counting units. Relevant research studies on frequency-based family word lists and academic words are also discussed.

2.1 Corpus-Based Studies

Corpus linguistics studies predominantly assist in revealing the vocabulary that appears most frequently used in a given text corpus. The importance of vocabulary selection, according to Laufer (2014), is due to the actual fact that students of L2 do not have plenty of time or opportunities to learn the same way that native speakers do. Thus, L2 learning experts and educators have made an effort to choose a reasonable number of the most crucial vocabulary words for learners' needs.

There are various ways to describe a corpus, but the definition that is most often accepted is that A corpus is a collection of authentic and machine-readable texts collected to represent a particular language or variation (McEnery et al., 2006). A collection of written or spoken texts that have been saved on a computer is referred to as a corpus (O'Keeffe et al., 2007). A corpus, according to Biber et al. (1998), is a collection of texts that has been organized and is suitable for both qualitative and quantitative research. In addition, the way academics think about vocabulary research studies has changed as a result of corpus-based research. Researchers have been using corpus data explicitly in order to critically evaluate the curriculum and teaching resources, despite the fact that corpora have mostly been employed to give more precise representations of language use (McEnery et al., 2006). Corpora may also include frequency information which could possibly influence the input given during instruction. This is based on the notion that learners should be familiar with those words that are used the most frequently in a particular language.

According to Nation (2017), vocabulary may be classified into 25 frequency-based word family lists utilizing information from the BNC/COCA frequency-based word lists, which are composed of up formal written text and spoken language. According to Nation (2006: 63), these lists "represent the higher frequency end of a learner's vocabulary, largely in the order of its range and frequency." The lists, which include a wide range of word families, offer more detailed analyses of the vocabulary load of the texts. Word frequency may assist facilitate learning, nevertheless it should not be the most important factor when determining what to teach and in what sequence. Several word lists have been made using corpora in an effort to identify the most important and frequent words used in an academic context. Since corpus-based studies have been made possible, researchers and educators have been working hard to develop corpus-based learning materials that could assist students from any field of study, focusing on lexical elements that are used across all academic disciplines (Coxhead, 2000).

The BNC, comprising over 100 million words, is a balanced corpus that represents all of up-to-date British English. On the contrary, the AWL, which replaced prior lists as the new standard at the start of the new century, has been a reliable resource for vocabulary instruction in English language instruction for more than ten years (Gardner and Davies, 2013). Such research need to be supported by a more extensive corpus that contains a larger selection of university textbooks. It is crucial to understand that this is not the first investigations of this type to be conducted. Examples such as Chen and Ge (2007), Chung and Nation (2003), Martinez et al. (2009), Vongpumivitch et al. (2009), and Khani and Tazik (2013) have all conducted studies of a similar nature. Researchers carried out their research in a range of contexts, including publications from the field of agricultural research, applied linguistics, and medical journals.

2.2 Second Language (L2) Vocabulary Learning and English Language Skills

Vocabulary is central to all four foundational language skills, and vocabulary mastery enhances students' language skills, including listening, speaking, reading, and writing, as there is a significant relationship between vocabulary and

language skills (Gardner, 2013; Nation, 2008; Richard & Rodgers, 2001; Schmitt, 2008; 2010). In addition, Thai EFL students recognize the importance of vocabulary learning (Muensorn & Tepsuriwong, 2009), which substantially contributes to language skills.

Vocabulary learning is crucial in teaching and learning a second language. Factors such as context, morphological generalization, and direct teaching contribute to vocabulary growth. (Bauer & Nation, 1993). Communication in every subject is enhanced by a strong vocabulary. According to Vongpumivitch et al. (2009), one of the major challenges students encounter when reading and writing academic discourse in an EFL environment is vocabulary. In addition, students are more likely to have limited or poor vocabulary understanding, and they are more likely to agree that their academic performance can suffer from a lack of vocabulary development (Altalhab, 2019; Pookcharoen, 2007).

Vocabulary learning is a challenging task for language learners, as it involves mastering a vast set of thousands of words. This task is characterized by its quantity and the mastery of various features and patterns (Laufer & Nation, 2012). Research studies (Aizawa and Rose, 2020; Chuenjit, 2011) found that vocabulary size is regarded as one of the perceived challenges contributing to L2 academic listening. For instance, Thai hospital pharmacists' statistically significant English listening comprehension were affected by the breadth of their vocabulary knowledge. Wang (2015) conducted a recent study on 120 Chinese non-English major undergraduate students at a medical university in order to investigate the relationship between vocabulary knowledge and L2 listening comprehension. It was discovered that both vocabulary breadth and depth significantly affected the results. Regarding the possible pedagogical implications discussed in the study, teachers should encourage students to perceive the value of vocabulary learning for listening proficiency development. Moreover, advanced English proficiency students should be provided with more diverse reading materials to increase their vocabulary size. Overall, these results are consistent with those published by Afshari & Tavakoli (2017) and Teng (2016), who discovered that both breadth and depth of vocabulary knowledge were positively and substantially correlated with academic listening comprehension achievement. In line with previous studies, Li & Zhang (2019) examined 290 Chinese

undergraduate student's vocabulary knowledge after taking the IELTS listening test. The relationship between L2 listening comprehension and the three dimensions of L2 vocabulary knowledge—vocabulary breadth, depth, and fluency—was examined using auditory vocabulary knowledge assessments. Vocabulary knowledge in all three dimensions significantly predicts L2 listening comprehension. In addition, the strongest predictor is the amount of L2 auditory vocabulary knowledge. Due to the moderate correlation between vocabulary breadth and depth, students should be aware of both aspects of vocabulary knowledge. However, vocabulary has a substantial role in both reading and writing skills. The quantity of the receptive vocabulary has a significant correlation with reading and writing skills, whereas it is only modestly associated with speaking and listening abilities (Staehr, 2008).

Reading comprehension is severely restricted by inadequate vocabulary knowledge, as it hinders Taiwanese college students' ability to recognize words and understand unfamiliar words (Huang, 1997). Reading comprehension may be hindered by an excessive number of unfamiliar words, as vocabulary is a strong predictor of one's ability to comprehend what is being read. (Nation, 2001; Sternberg, 1987).

2.3 Vocabulary Size

Vocabulary knowledge has conventionally been conceptualized in terms of vocabulary breadth, or vocabulary size, vocabulary depth, and fluency. In this current study, the term 'vocabulary size' is used because it corresponds to the breadth of vocabulary knowledge or vocabulary breadth. It theoretically refers to the number of words a person knows. In other words, it refers to having surface-level knowledge of many vocabulary words. Students know a word or its collocate but are unable to incorporate it into a sentence. It is generally accepted that vocabulary size plays an important role in L2 acquisition. Learners' progress in second language acquisition is highly dependent on their progress in acquiring individual words and parts of speech (Barcroft, 2004). Meara and Fitzpatrick (2000) claimed that learners with a larger vocabulary can communicate more successfully than learners who have detailed mastery of a smaller vocabulary.

In the area of vocabulary learning, attempts have been made to systematically measure the vocabulary size of native English speakers, for which there

are a number of reasons or motivations (Nation & Waring, 1997). Vocabulary size may be due to an individual's educational background and other related influencing factors. In addition, it has been pointed out that the number of words an L2 learner needs to know depends on what the learner wants to do with the language (Nation & Waring, 1997, as cited in Benedict & Shabdin, 2021). Interestingly, vocabulary knowledge was found to be more important in reading for L2 learners than for L1 learners (Verhoeven, 2000). Nation (2006) tried to find out through questioning what is the reading and listening vocabulary knowledge (receptive vocabulary knowledge) or vocabulary size of English that a learner needs to know for typical uses of the language, e.g., reading authentic texts and literature such as a newspaper, a novel, a food label, or even a journal article for an academic purpose and experiencing authentic language environments such as watching a movie, listening to a song, etc. It is a widely held view that there are meaningful goals for vocabulary learning based on the following three different aspects: 1) how many words there are in English, 2) how many words a native speaker knows, and 3) how much vocabulary a learner needs to know for typical language use in order to decide how large a non-native learner's vocabulary of that language should be (Nation, 2006; Schmitt, 2008).

It is then a good idea to clarify key terms whose understanding is essential in any discussion of vocabulary and reading. According to research, an English learner needs to comprehend 98% of running words, equivalent to around 8,000 word families (Nation, 2006), in order to understand written English without assistance (Hu & Nation, 2000; Schmitt, Jiang, & Grabe, 2011), as it requires deliberate and incidental vocabulary learning (Nation & Anthony, 2013). Schmitt (2008) suggests a large vocabulary size of 8,000–9,000 word families for reading and 5,000–7,000 families for oral discourse. L2 learners must make efforts to expand their vocabulary to read a wide range of materials without interference from unfamiliar vocabulary. A large vocabulary is essential for functioning in English, with 8,000–9,900 word families for reading. Laufer and Ravenhorst-Kalovski (2010) and Nation (2006) suggest that textbooks require a minimal vocabulary size of 4,000–5,000 word families for 95% text coverage and 7,000–8,000 word families for 98%.

As a result, knowledge of a certain vocabulary list is needed to comprehend what is read. In addition, vocabulary forms in a very complicated way (Tseng &

Schmitt, 2008). Based on a range of vocabulary research studies, certain vocabulary size is required to achieve particular language goals. The following is a summary of the English lexical requirements: 3,000 word families are needed for authentic texts, 5,000–9,000 word families to read authentic texts without assistance, and 10,000 word families to be able to effectively use most of the language.

The vocabulary size of Thai students in primary and secondary school has been the focus of several studies. The vocabulary size of Thai university students who have studied English since elementary school for many years are mostly unknown. The vocabulary size of Thai first-year undergraduate students was estimated by Mungkonwong and Wudthayagorn's study (2017) to be approximately 4,200 word families. Additionally, it exceeded the basic core curriculum's recommended vocabulary size of 3,600–3,750 word families.

Due to the different text types of general and academic texts, Moon (2017) investigated if students needed different types of vocabulary for reading comprehension, and the findings revealed Vocabulary size was a significant predictor of reading comprehension, but academic vocabulary was more predictive. General and academic reading comprehension have different characteristics, so each type should be taught differently. Interestingly, vocabulary size reciprocally results from text coverage, whereas text coverage is not affected by text length (Fan, 2013).

2.4 Text Coverage

Text coverage is brought into question by the question of how many words readers should know to understand the text they are reading. The term ‘text coverage’, also known as lexical text coverage, is a fundamental concept that refers to the percentage of running words, also known as tokens, in the text a reader knows (Nation, 2006). It theoretically explains how many words are known in a text. Text coverage is of increasing interest and significance for measuring the intelligibility of reading materials in the field of vocabulary teaching and learning.

According to the study of Chanchanglek and Sriussadaporn (2011), Rangsit University, Thammasart University, and Rajamangala University of Technology's English for Engineering courses are highly impacted by words from the General Service List (GSL). GSL vocabulary comprised almost 80% of corpora, covering more than

80% of the text in the majority of reading texts. This is consistent with Nation's (2001) finding that high-frequency GSL words account for more than 80% of all written texts.

This indicates that from the viewpoint of the reader, familiarity with the AWL will ensure the understanding of 10% of the vocabulary in academic contexts. The percentage of words that a reader comprehends can also be used to describe coverage. For instance, if readers comprehend 95% of the running tokens in the text, they have achieved 95% text coverage (Laufer & Ravenhorst-Kalovski, 2010).

The estimated proportion of text coverage that contributes to reading comprehension was shown in earlier research studies, which also showed a high correlation between vocabulary knowledge and reading comprehension in written texts. According to Laufer (1989), for a reader to have an adequate comprehension of the text, they must be familiar with 95% of the vocabulary. According to Hu and Nation's (2000) study, 98% lexical coverage was required for adequate comprehension. In a more recent research, Schmitt et al. (2011) recommended 98% as the ideal coverage for an adequate comprehension of academic texts.

The lexical threshold is a crucial factor in determining the coverage of vocabulary in specialized academic word list studies. Minshall (2013) highlights the diminishing returns when setting 98% as the lexical threshold as vocabulary sizes increase. Hsu (2014) used an estimated coverage percentage of 95% for this research to develop an Engineering English Word List to fill the gap experienced by EFL students. 4,000 word families plus proper nouns included books and newspapers in 95% of text coverage, according to Nation (2006), whereas 8,000–9,000 word families plus proper nouns covered in 98%. EFL students must know twice as much vocabulary as they do with 95% coverage in order to achieve only a 3% increase in coverage. As a result, this study uses 95% as the minimum level of textbook reading comprehension for Chinese computer science students.

Text coverage, or the percentage of a text that a particular vocabulary size covers, is a measurement used in the teaching and learning of vocabulary. This standardized process assists researchers and educators in comprehending vocabulary learning objectives and the simplicity or complexity of a given text. (Brown, 2010).

2.5 The BNC/COCA Word Family Lists

According to Bauer and Nation (1993), reading vocabulary growth is crucial for developing readers, and learners' vocabulary knowledge is significantly influenced by their morphological knowledge. This knowledge increases over time, with individuals still growing in their capacity to understand derivational affixes.

A word family is a collection of words that includes a base word together with its derivational and inflected forms. Word families enable language learners to understand a language without having to learn each word separately. The word family grows in size as a learner's knowledge of affixation develops. Underlying word families are investigated on a fundamental basis that after the base word or its derivational form is recognized, more family members may be identified with minimal effort. The base's meaning in the derived forms must be relatively comparable to the base's meaning when it appears alone or in other derived forms.

Paul Nation created the BNC/COCA word family lists, which are frequency-based lists with 1000 word bands. The version employed in this current study ranges from 1k to 25k, giving the 25,000 English words that are used most commonly. 29 lists total, comprising 25 frequency and range data-based lists, four extra lists of proper names, marginal words, transparent compounds, and abbreviations, constitutes the BNC/COCA word family lists. Based on word frequency, 100 million words were selected for these lists, with each 1000-word family representing one level. The theoretical study on the 25 levels is concerned with how students master vocabulary in terms of word frequency. Proper noun and interjection vocabularies that can be employed in software analysis are included in the BNC/COCA corpus. The first and second 1000-word family lists were selected from a 10-million-word corpus consisting of 4 million words of written English and 6 million words of spoken British and American English. These spoken and written genres were used to create the frequency lists. This prevented the written corpus utilized for the following lists from excessively impacting the first two lists. Along with word sets like numbers, days of the week, and months of the year, common spoken words were also included. Additionally, there was a vocabulary list of travel-related survival words. After excluding the 1k and 2k words, the rankings from the BNC and COCA were used to develop the third 1000 list.

2.6 Academic Vocabulary and the Academic Word List (AWL)

Academic vocabulary, commonly referred to as sub-technical or semi-technical words, is a group of words that appear frequently in academic works but are uncommon in other genres, according to Coxhead and Nation (2001). In arranging an author's ideas, these words serve both organizational and rhetorical functions (Baker, 1988). As a result, academic words include running words that are substantially greater in academic texts than in other text genres. However, academic vocabulary has been a challenge for language learners due to its lack of frequency compared to general English words. This indicates the most common language issue for both non-native and native learners, which is unfamiliarity with academic vocabulary. Learning English for academic purposes requires a combination of knowledge related to the specific discipline and the academic language. High-frequency words, academic words, technical terms, and low-frequency words are the four categories of academic words in English (Nation, 2001). High-frequency words are necessary for common speech and running words in all writings, whereas technical words are required and used in specialized disciplines and are determined by various methods, such as expert opinions or specialized dictionary references.

Academic word lists have been developed to assist students with the learning and teaching of academic vocabulary in order to deal with this challenge. English academic vocabulary, which is inspired by two lists, must be employed in academic texts. The University Word List (UWL), encompassing 8.5% of the words used in academic texts, has been substituted by the Academic Word List (AWL) by Coxhead (2000). The AWL, which was developed using a corpus of texts from four different academic disciplines, has 570 items and covers the UWL in the majority of academic texts. Most textbooks now employ the AWL, ensuring a thorough knowledge of English in diverse academic materials, even though the GSL is still frequently used for such works (Lessard-Clouston, 2013).

The Academic Word List (AWL; Coxhead, 2000) is one of the most influential and well-known collections of academic words and was extracted from a corpus of 3.5 million running words (tokens). It is a corpus compilation of academic texts, including textbooks, journal articles, book chapters, and laboratory manuals,

aimed at representing first-year students' reading at the tertiary level (Coxhead, 2011, as cited in Coxhead, 1998, 2000), and extracted from four disciplines: arts, commerce, law, and science. The AWL consisted of all affixes and its stem, including Level Six of Bauer and Nation's (1993) scale. In addition, a word family is a set of words that may all have the same root word but differ in their morphological prefixes and suffixes. As mentioned earlier, a word family, usually containing a stem and its members, may include their family members as derivational affixes (added to forms to create separate words, for example, changing a verb to a noun) and inflectional affixes (added not to create separate words, for example, occurring due to changes in grammatical properties). For example, a word family of 'analyse' contains the following members: analysed, analyser, analysers, analyses, analyzing, analysis, analyze, analyzed, analyzes, and analyzing.

The average percentage of academic texts that employ AWL is determined to be 10%. The AWL established an ideal framework for the development of academic word lists and made significant contributions to EFL education. The AWL nevertheless faced some criticism as a result of its widespread use. Gardner and Davies (2014) identified weaknesses in the AWL and developed a revised list of 3,000 lemmas, or 2,000 word families, using a corpus of 120 million words of academic texts.

The focus of the study, according to Chen and Ge's (2007) investigation, is on the significance of academic words in medical research articles (RAs), with 292 (51.2%) of them being used often. The results support Coxhead's work on academic texts in different topic disciplines and reveal that academic vocabulary has a high text coverage at about 10.07% in medical RAs. Learners, especially those studying EMP (English for Medical Purpose), can perform better when academic assignments are completed by identifying AWL words in medical RAs. In terms of the frequently used academic words in medical RAs, the AWL list is far from being representative. Learners and users should pay close attention to the rhetorical purposes that some academic words fulfill since these functions are crucial to understanding academic vocabulary.

Ali (2010) investigated the relationship between students' vocabulary knowledge and their reading comprehension. It was conducted on university students in the English Department at a university in Jakarta. The research discovered a

significant connection between reading and vocabulary. The study's conclusions suggest that students with a larger vocabulary find it easier to learn newly acquired information, such as by means of stories published in English, and to expand their knowledge by reading a variety of books.

Hsu (2010) has compiled a corpus of General English (GE) reading materials used by Taiwanese educational institutions. The 2,000 basic English words needed by Taiwan's Ministry of Education for the Grades 1–9 Curriculum, as well as the elementary and intermediate vocabulary from the General English Proficiency Test (GEPT), were lemmatized into word families and added to the base words, the BNC high-frequency word lists, and the AWL analyzed in the Range software program. A GE textbook corpus contains 1.3%–6.54% of the total words in Coxhead's (2000) 570 academic word families.

A list of high-frequency content words from 64 open-access English social science research articles published in 11 journals was compiled by Chanasattru and Tangkiengsirisin (2016). AntWordProfiler 1.4.0 and AntConc 3.4.3 were used to compare the ranges and frequencies of words found in the social science corpus with those found in the New General Service List (NGSL) and the Academic Word List (AWL). 1,120 word members and 394 high-frequency content headwords were discovered.

Using a corpus-based and expert-judged technique, It-ngam and Phoocharoensil (2019) created the specialized academic word list from 11 natural scientific sub-disciplines. The Science Academic Journal (SAJ) Corpus, a 5.5 million-word corpus used for this study, was created. Around 10% of the AWL as an academic text corpus was covered by the SAJ corpus, and 568 AWL word families were found. Comparable to Coxhead's (2000) study, which discovered that the 570 words in AWL constitute 10% of the academic corpus, is this proportion.

Following Coxhead's (2000) corpus analysis method in significantly certain aspects, researchers have created several field-specific academic or technical word lists to satisfy the specialized vocabulary demands of EFL learners of different disciplines. Hsu (2013) employed 155 textbooks from 31 different medical topic areas as a corpus and extracted 595 word families to create the Medical Word List. In a pilot investigation, Coxhead and Hirsch (2007) found 318 word families that

had far higher coverage in a corpus devoted to science than in non-scientific corpora. For less experienced engineering undergraduates, there is also Ward's (2009) 299-word Basic Engineering List, Yang's (2015) Nursing Academic Word List with 676 word families, and Coxhead and Demecheleer's (2018) Plumbing Word List with 1,465 different types.

The Computer Science Word List (CSWL), created by Minshall (2013) to generate a list of 433 word families from a corpus of 3.66 million tokens gathered from journal articles and conference proceedings encompassing 10 sub-disciplines of computer science, is the most pertinent to the present study. With the help of the GSL, the AWL, and the CSWL, 95.11 percent of the tokens found in the corpus met Laufer's (1989) minimal reading comprehension threshold.

Selecting the general high-frequency words, as those in West's (1953) GSL, which are believed to be known by students aiming for university study, is a common technique employed in the production of specialized word lists. However, some academics have questioned the justification of including academic word lists to a general service list by pointing out that many words in the GSL have academic or technical meanings (e.g., Dang et al., 2017; Gardner & Davies, 2014; Lei & Liu, 2016; Neufeld et., 2011).

2.7 Word Frequency, Word Lists, and Counting Units

To count words in a text, there are some key concepts associated with corpus linguistics that help us understand word counting better. These are 'word frequency', 'word list', 'word type', 'running word (or token)', and 'concordance'.

A lot of words are used in English writing, but not all of them are essential. Word frequency refers to the number of times a word appears in a particular text. With its occurrence in approximately seven percent of written and spoken words in English, the word "is" is a very useful word. A text can be understood to a good extent with this frequency. The majority of running words in a written or spoken text are content words; therefore, second language learners and teachers can benefit from being familiar with them. The text can be easier to comprehend if these words are known.

A text in English uses a large number of words, but not all are equally useful. Word frequency refers to the notion that not all words are used with equal

frequency (Hirsh & Nation, 1992; Nation & Waring, 1997). Word frequency is a measure of usefulness, indicating how often a word occurs in normal use of the language. This frequency is crucial for second language learners and teachers, as it allows them to know a large proportion of the running words in a written or spoken text.

Nation (2001) divided vocabulary into high-, mid-, and low-frequency categories. High-frequency vocabulary is the foundation of language use and is used frequently regardless of text type. Learning these high-frequency words initially is strongly encouraged for beginning-level students, as it will considerably expand their vocabulary coverage (Nation, 2001). High-frequency, academic, technical, and low-frequency are the two systems of classification for English words that Nation (2001) and Schmitt (2000) proposed. Later, a frequency-based classification of high, mid, and low-frequency words was suggested by Schmitt and Schmitt (2014). The system was then restructured to classify the levels based on both frequency (high, mid, and low) and text type (academic and technical). Researchers have used the definitions provided by Nation for each level as the foundation for their vocabulary analyses.

Word lists, or word frequency lists, are an essential tool in ESL and EFL teaching materials because they can be used to determine the most useful vocabulary that ESL and EFL learners need. According to Nation and Waring (1997), word lists play a vital role in curriculum design and teaching, providing valuable insights for both teachers and learners. Course designers and teachers need reference lists to assess the relevance of words and class content. The GSL, or General Service List, is a classic list of high-frequency words, containing words within the word family with their own frequency. It was written to be used as a resource for compiling simplified reading texts into stages or steps. However, doubts have been cast on its adequacy due to its age, errors, and solely written base.

Here are some of the most well-known word lists that appear the most frequently in the English language. The General Service List (GSL) by West (1953), the Academic Word List (AWL) by Coxhead (2000), the Medical Word List by Hsu (2013), the Nursing Academic Word List by Yang (2015), the Academic Spoken Word List by Dang, Coxhead, and Webb (2017), and many other word lists have been created in previous studies using word families as the unit of counting.

A word family is created when a base word is combined with its inflectional and derivational forms, which consist of affixes from the affixation list for word families by Bauer and Nation (1993). For instance, the inflectional forms accessed, accesses, accessing as well as the derivational forms accessibility, accessible, inaccessibility, inaccessible are all members of the word access. Eight different word members are counted as one item in the word list. The rationale behind using word families as the unit of counting is that after learners have mastered the stem word and acquired a basic understanding of word-building techniques, they will naturally deduce the meanings of frequently inflected and derived forms of that word (Bauer & Nation, 1993). Lemmas were employed by other scholars as the unit of counting (Brezina & Gablasova, 2015; Dang & Webb, 2016; Gardner & Davies, 2014). In addition, only the headword and its inflectional forms are included in a lemma.

The counting unit in word list studies depends on the concentration of a particular skill. Generally, Lemmas are a more practical counting unit for productive usage compared to word families for receptive use (Brezina & Gablasova, 2015; Nation, 2016). One of the research objectives in the present study is to create an AWL academic word list for economics first-year undergraduate students to achieve better reading comprehension of textbooks. Consequently, the preferred unit of counting is word families rather than lemmas.

Based on the notion that too many words will increase students' learning burden and extend their learning time, understanding the ways to measure which words and how many words are in the first-year undergraduate textbooks of the BEIP is critical. This study will also be an alternative guideline for all the stakeholders who use and benefit from first-year undergraduate textbooks in an international program in economics.

CHAPTER 3
RESEARCH METHODOLOGY

This chapter aims to present a comprehensive, structured research methodology for conducting the research by describing and explaining the research procedures in detail and explaining how the methodological approach, methods, and techniques that the researcher systematically applied and carried out to answer the research questions. The chapter begins with the research design, which encompasses the research instruments. Then, the text corpus compilation of the textbooks, and data analysis are discussed.

Focusing on a corpus-based vocabulary analysis in the first-year undergraduate economics textbooks of the Bachelor of Economics International Program (BEIP), the research questions are thus formulated as follows:

(1) What text coverage do the BNC/COCA word family lists and the Academic Word List provide in the first-year undergraduate economics textbooks of the BEIP?

(2) What vocabulary size is needed to read the first-year undergraduate economics textbooks of the BEIP?

(3) What are the most frequently used academic words in the first-year undergraduate economics textbooks of the BEIP?

3.1 Research Design

The quantitative research methodology was employed to measure text coverage, vocabulary size, and the academic words appearing in the text data of the textbooks of the first-year undergraduate economics core courses taught in the BEIP at an autonomous university in Thailand. In other words, the methodology is typically concerned with the quantitative research analysis in the corpus text. The design of the corpus and the corpus compilation are then discussed.

3.2 Research Instruments

3.2.1 AntWordProfiler

The AntWordProfiler software program (Anthony, 2022), a freeware vocabulary analysis program, was selected as one of the research instruments in order to draw on a quantitative text analysis to yield results in answering the research questions of this study. It serves crucial functions that allow the researcher to investigate the text coverage and word frequency of the text corpus of the first-year undergraduate economics textbooks of the BEIP and contributes to the results of estimated vocabulary size and the most frequently occurring academic words.

The researcher compared the vocabulary of several texts (the textbooks) using the AntWordProfiler tool. In other words, it was used to compare the vocabulary of many text files or a single large corpus at the same time in order to investigate the text coverage of the base word lists, the vocabulary size, and how frequently an individual word appeared overall and in each file. Additionally, word lists based on word frequency and word distribution are created, text coverage by certain word lists is investigated. As a result, teachers can plan realistic sequences for mastering vocabulary in class, focus on particular vocabulary learning requirements, or evaluate and adjust the requirements for learning and opportunities presented by texts. Additionally, instead of being at random, it enables students to increase their vocabulary knowledge at a certain level (Cobb, 2010). For this study, the first 1,000 to twenty-five thousand BNC/COCA word families were used according to word frequencies. The word families in the lists were developed at Level 6, which, according to Bauer and Nation's (1993) definition of word families, comprises inflections and the high-frequency, regular, productive, and transparent derivational affixes. Thus, to compare the first-year undergraduate textbooks of the BEIP, the following two categories of word lists were used: (1) the BNC/COCA 1st to 25th 1,000 word families (Nation, 2017) and (2) Coxhead’s (2000) AWL containing 570 word families.

The word family is the most sensible unit to use for measuring receptive knowledge (reading and listening), according to Nation and Webb (2011). The rationale for adopting the concept of "word family" as a unit of counting is that if just one or two family members are recognized, then just a minimal amount of learning

is needed for their receptive usage (comprehension) by other members of their family. According to Bauer and Nation (1993), the idea of a word family is used to represent a collection of words whose meanings may be inferred when the learner is aware of the meaning of the base form in the group. In this respect, throughout the study, 'word family' is used as a unit for counting the most frequent words in the corpus.

3.2.2 The Selected Undergraduate Economics Textbooks

In order to achieve the purpose of this study, the researcher selected the compulsory economics textbooks used by first-year undergraduate students of the Bachelor of Economics International Program at an autonomous university in Thailand, hereafter referred to as the 'BEIP'. As the target textbooks significantly contribute to the research investigation procedures and results, they are mandatory textbooks that all the first-year undergraduate students in the BEIP read in each compulsory core course listed in the study plan of the BEIP. Students are required to earn academic credits in the economics core courses and the general education courses during the first year of study. In other words, the concentration is only on the textbooks in the first-year major area courses, excluding the list of minor areas or non-economics elective courses, free elective courses, and general education courses. In addition, the general education courses were excluded as follows: (1) civic engagement; (2) Thailand, ASEAN, and the world; (3) social life skills; (4) life and sustainability; (5) critical thinking, reading, and writing; (6) communication skills in English; (7) creativity and communication; and (8) fundamental mathematics. The researcher excluded the aforementioned list of courses because the other courses were different disciplines in nature, and the focus is only on specific textbooks in economics required to be read in the first year of study in an international program that this current study aims to investigate. The texts in the textbooks were of different lengths and were aimed at providing economics-major students with economic backgrounds, principles, and content-specific knowledge of each discipline.

According to the 2018 BEIP curriculum, the program requires students to complete a minimum of 131 credits following the course structure and requirements to graduate. The researcher purposefully investigated the first year of economics core courses in which BEIP students are accredited with 12 credits (see Appendix B for the BEIP education plan). In the first year of study, the BEIP introduces

the following economics core courses: (1) the course of ‘Principles of Microeconomics’ (EE211); (2) the course of ‘Principles of Macroeconomics’ (EE212); (3) the course of ‘Calculus for Social Science 1’ (MA216); and (4) the course of ‘Statistics for Social Science 1’ (ST216).

Principles of Microeconomics (EE211), an economics core subject, is specified as a compulsory course that the students are required to take during the first semester of their first year of study. It focuses on the fundamentals of microeconomics, including value, price, and resource allocation. It also introduces theories of consumption and production with a focus on the variables that affect supply and demand for goods and services, price determination, and the effectiveness of resource allocation in both perfect and monopoly markets. Finally, it covers the fundamentals of the competitive factor market and introduces the idea of market failures.

Principles of Macroeconomics (EE212), an economics core course, also emphasizes indicators, objectives, and issues in macroeconomics, as well as theories of aggregate consumption and investment, the accelerator principle, money markets, the theory of supply and demand for money, the joint equilibrium model of product and money markets (IS-LM model), the balance of payments, and fiscal and monetary policies as tools for stabilizing the economy. In addition, methods for gathering and organizing Thai macroeconomic data are provided for the analysis of economic circumstances.

Calculus for Social Science 1 (MA216), the core economics course, focuses on the limits and continuity of one-variable functions, derivatives of algebraic and transcendental functions, implicit differentiation, higher-order derivatives, Roll's theorem, the mean value theorem, applications of the derivative for determining the maximum and minimum of functions, differentials and their applications, antiderivatives, indefinite integrals and integration, definite integrals and application of area solving, and functions of several variables.

Statistics for Social Science 1 (ST216), the core economics course, focuses on the following topics: an introduction to descriptive statistics; index numbers; unconditional and conditional probability; random variables and probability distribution; unconditional and conditional expectations; elementary sampling and

sampling distribution; estimation and hypothesis testing for one population; and statistical package results interpretation.

3.3 The Text Corpus Compilation of the Textbooks

Due to the wide variety of in-class and outside-of-class materials used in the Bachelor of Economics International Program (BEIP), the researcher decided to select the first-year undergraduate textbooks of the BEIP for the text corpus compilation to be representative of the economics textbooks required to be read in the first year of study. Students in their first year of study in the BEIP are likely to continue their studies or not, and they are going to read English-medium textbooks that they are not familiar with, especially in a more specific discipline. Even students with a large vocabulary may encounter difficulties in learning or comprehending what they are reading in a text. As mentioned earlier, the text corpus comprises the first-year undergraduate economics textbooks of the BEIP listed in the course syllabuses of the BEIP curriculum. The researcher continued to compile a corpus, which was at a critical stage because various important issues should have been taken into consideration. The researcher officially asked for the faculty's permission to have a list of the compulsory textbooks used in the first-year undergraduate economics core courses. Developing the text corpus involved collecting texts in each specific textbook in an electronically readable format. If some of the first-year undergraduate textbooks of the BEIP are in hardcopy or printed format, the researcher will use a scanner to invert the files into an electronically readable format. However, all the text files were electronically readable.

3.4 Data Analysis

At the initial stage of corpus compilation, the researcher will prepare texts which include large, electronic data sets of corpus data of the first-year undergraduate textbooks of the BEIP. The corpus texts will be used to compile a corpus of the economics core courses. However, the aforementioned corpus is compiled to represent the first-year undergraduate economics textbooks of the BEIP. As mentioned earlier, a text with Microsoft Word (.doc) format file or any other format is not able to be processed on the AntWordProfiler software program. If some of the first-year undergraduate textbooks are printed textbooks or only available in a hardcopy textbook

version, the texts will be processed to electronically-readable data. Additionally, any pictures, tables, anchoring objects, numbers, symbols, acronyms, the table of contents, indices, appendices, and any glossaries are skimmed and scanned to remove unnecessary information.

The spelling-checking feature and the Find and Replace feature in a word processor software program were used to recheck the corpus data. In addition, an electronic version of the first-year undergraduate textbooks of the BEIP is converted from PDF (.pdf) format to Microsoft Word (.doc or .docx) format using AntFileConverter (2022), a software program employed in this study to convert either PDF format and Microsoft Word (.doc or .docx) format into plain text for use in corpus tools. After the texts have been completely stripped of irrelevant items, the researcher finally converts them into plain text (.txt) files for the next stage of data analysis. In addition, the corpus data are retrieved from the textbooks of the economics core courses. The texts were compiled into a corpus of the first-year undergraduate textbooks of the BEIP.

To answer the first research question, the AntWordProfiler software program (Anthony, 2022) was used to analyze the vocabulary of the BNC/COCA word family lists and the Academic Word List occurring in the first-year undergraduate textbooks of the BEIP. In addition, the software program is a vocabulary analysis program designed to inform how much and what vocabulary occurs in a particular text or a group of texts.

Familizer, a web-based vocabulary software program, is also used to assist the researcher in expanding a raw word list into a word family list. After the ECON corpus was analyzed and resulted in the most frequently occurring AWL academic headword. The next step was to allocate words to a family using Familizer (Cobb, n.d.) A headword and all of its derived and inflected forms constitute a word family. This may assist learners since understanding a headword makes it easier for them to understand its inflected and derived forms (Coxhead, 2000).

3.4 Validity and Reliability

As all Microsoft Office programs can check spelling and most can deal with grammar, the researcher can ensure the validity of the texts. It possibly causes some errors deviating from the validity because some sections or chapters were typed manually or edited for spelling using a spelling-checking feature and the Find and Replace feature in a word processor software program on the computer. To ensure the validity of the corpus, the researcher also removed all the bibliographies, tables, indices, photographic images, and appendices from the text corpus.

A text corpus's ability to correctly represent the corpus is one of the most important factors to take into consideration when evaluating its reliability. However, even a fairly large corpus should be manageable. For instance, the AWL text coverage in Cobb and Horst's (2002) research was 6.72%, and 2,024 running words from a corpus of a single medical text were examined in the study. Because it might be difficult to accept a relatively small sample as adequate and convincing, the representativeness and reliability of a small text corpus can be questioned (Chen & Ge, 2007). Sinclair (2005) argues that the best way to address this problem is to create a corpus that is large enough to incorporate even the longest texts. Text selection becomes crucial if this is not feasible and there is a potential that a lengthy text may have a significant impact on the entire corpus; this must be done using one's best or professional judgment. There is no apparent maximum or minimum number of texts that a corpus should include, but it is important to keep in mind that many writers have a propensity to endorse extreme perspectives when an estimated number is supplied.

CHAPTER 4
RESULTS

A corpus of first-year undergraduate economics textbooks used in the Bachelor of Economics International Program (BEIP) was compiled using a corpus-based vocabulary analysis. The 1,343,493-word ECON corpus has been created, comprising four subcorpora that were collected from the four economics course textbooks as follows:

(1) the course textbook of *Principles of Microeconomics,*

(2) the course textbook of *Principles of Macroeconomics,*

(3) the course textbook of *Calculus for Social Science 1,* and

(4) the course textbook of *Statistics for Social Science 1*.

As previously stated in Chapter 1, the research questions in the study were thus formulated as follows:

(1) 'What text coverage do the BNC/COCA frequency-based word family lists (Nation, 2017) and the Academic Word List (Coxhead, 2000) provide in the first-year undergraduate economics textbooks of the BEIP?'

(2) 'What vocabulary size is needed to read the first-year undergraduate economics textbooks of the BEIP?'

(3) 'What are the most frequently used academic words in the first-year undergraduate economics textbooks of the BEIP?'

The purpose of Chapter 4 is to present the results that were discovered and drawn from the data in order to answer the aforementioned research questions in the study. Consequently, the chapter is divided into the three major sections as follows:

4.1 the text coverage of the BNC/COCA frequency-based word family lists provided and the vocabulary size needed to read the textbooks,

4.2 the text coverage of the Academic Word List provided in the corpus of first-year undergraduate economics textbooks, and

4.3 the most frequently occurring AWL academic words in the textbooks.

4.1 The Text Coverage of the BNC/COCA Frequency-Based Word Family Lists Provided and the Vocabulary Size Needed to Read the Textbooks

The first two study objectives are to investigate how much of the university textbooks are covered by the BNC frequency-based lists and to predict the vocabulary size students should have in order to read the university textbooks. The ECON corpus and all the subcorpora were examined using the AntWordProfiler software program (Anthony, 2022) in order to obtain the answers to these questions.

In this section, the findings show (1) tokens, also known as running words (known and unknown), that appear in the whole corpus and its subcorpora of first-year undergraduate economics textbooks in this study; (2) word families, which refer to groups of words that share a common root word with different derivations (e.g., prefixes and suffixes) and inflections (different forms); and (3) text coverage, or the percentage of words known in the textbooks, which may indicate text comprehension. As stated in Chapter 2, text comprehension and vocabulary knowledge are necessary for reading comprehension. As a result, knowing the estimated percentage of known words in order to understand the texts being read is useful for materials developers and other related stakeholders.

The purpose of this section is to present the findings of the text coverage of the BNC/COCA frequency-based word family lists and the estimated vocabulary size in the textbooks as the whole corpus and subcorpora of the four different economics core courses as follows: principles of microeconomics (MIC), principles of macroeconomics (MAC), calculus for social science 1 (CAL), and statistics for social science 1 (STA).

4.1.1 The Text Coverage of the BNC/COCA Frequency-Based Word Family Lists and the Estimated Vocabulary Size Needed to Read in the Principles of Microeconomics (MIC) Subcorpus

One of the economics core course textbooks, Principles of Microeconomics of the Bachelor of Economics International Program (BEIP), was chosen to represent the economics textbook subcorpus, referred to as the MIC

subcorpus, which was the compulsory textbook required to read, as listed in the first year of study of the 2018 BEIP curriculum.

This part of the investigation is primarily aimed at answering the first two research questions addressing: (1) the text coverage provided by the BNC/COCA frequency-based word family lists in the BEIP's first-year undergraduate economics textbook (the MIC subcorpus); and (2) the vocabulary size needed to read the textbook (the MIC subcorpus). The MIC subcorpus was then analyzed using the AntWordProfiler software program (Anthony, 2022) to investigate as follows: (1) the tokens, also known as running words, in each word band in the MIC subcorpus; (2) the word families found in each word band of the BNC/COCA frequency-based word family lists (Nation, 2017) in the MIC subcorpus; (3) the text coverage of the BNC/COCA frequency-based word family lists in the MIC subcorpus; and (4) the cumulative text coverage of the BNC/COCA frequency-based word family lists.

Table 4.1

Word Levels, Tokens, Number of Word Families, Text Coverage, and Cumulative Text Coverage of the BNC/COCA Word Family Lists in the MIC Subcorpus

BNC/COCA Word List	Tokens	Number of Word Families	Text Coverage	Cumulative Text Coverage
Proper Nouns	2,734	-	1.20%	1.20%
1st 1,000	165,081	927	72.32%	73.51%
2nd 1,000	30,889	773	13.53%	87.05%
3rd 1,000	**19,671**	**783**	**8.62%**	***95.66%**
4th 1,000	4,620	443	2.02%	97.69%
5th 1,000	**2,280**	**253**	**1.00%**	****98.69%**
6th 1,000	679	200	0.30%	98.98%
7th 1,000	419	121	0.18%	99.17%
8th 1,000	361	105	0.16%	99.33%
9th 1,000	70	43	0.03%	99.36%
10th 1,000	109	57	0.05%	99.40%
11th 1,000	86	30	0.04%	99.44%
12th 1,000	42	27	0.02%	99.46%
13th 1,000	65	17	0.03%	99.49%
14th 1,000	123	14	0.05%	99.54%
15th 1,000	37	14	0.02%	99.56%
16th 1,000	9	7	0.00%	99.56%
17th 1,000	16	4	0.01%	99.57%
18th 1,000	19	6	0.01%	99.58%

BNC/COCA Word List	Tokens	Number of Word Families	Text Coverage	Cumulative Text Coverage
19th 1,000	2	2	0.00%	99.58%
20th 1,000	30	6	0.01%	99.59%
21st 1,000	7	1	0.00%	99.59%
22nd 1,000	6	1	0.00%	99.60%
23rd 1,000	3	2	0.00%	99.60%
24th 1,000	0	0	0.00%	99.60%
25th 1,000	1	1	0.00%	99.60%
Off-List	915	-	0.40%	100.00%
Total	**228,274**	**-**	**100.00%**	**100.00%**

(Remarks: *meaning that 95% of text coverage was reached, **meaning that 98% of text coverage was reached)

Table 4.1 demonstrates (1) the total tokens in the MIC subcorpus and the tokens in each BNC/COCA frequency-based word family list; (2) the number of word families; (3) the text coverage; and (4) the cumulative text coverage of the BNC/COCA frequency-based word family lists in the textbook of the economic core course entitled *Principles of Microeconomics*, representing the MIC subcorpus in this current study.

The data indicated that the MIC subcorpus comprised 228,274 tokens (running words) in total. The cumulative text coverage of the BNC/COCA frequency-based word family lists has shown that the 3rd 1,000 BNC/COCA word family list has already reached 95.66% of the MIC subcorpus in order to have minimal comprehension (where minimally acceptable comprehension can occur) at 95% as suggested by Laufer and Ravenhorst-Kalovski (2010), resulting in having a vocabulary size of around 2,500–3,000 word families. The 5th 1,000-word band of the BNC/COCA lists already reaches 98.69% text coverage in the MIC subcorpus and it can be inferred that having vocabulary knowledge of around 4,500–5,000 word families to have optimal comprehension of the texts at 98%.

4.1.2 The Text Coverage of the BNC/COCA Frequency-Based Word Lists and the Estimated Vocabulary Size Needed to Read in the Principles of Macroeconomics (MAC) Subcorpus

One of the economics core course textbooks entitled *Principles of Macroeconomics* of the Bachelor of Economics International Program (BEIP) was

selected to represent the economics textbook subcorpus, hereafter the MAC subcorpus, which was the compulsory textbook needed to read as listed in the first year of study of the 2018 BEIP curriculum.

This part of the investigation is primarily aimed at answering the first research question concerning the text coverage that the BNC/COCA frequency-based word family lists provide in the BEIP's first-year undergraduate economics textbook (the MAC subcorpus) and (2) the vocabulary size needed to read the textbook (the MAC subcorpus). The MAC subcorpus was analyzed to investigate (1) the number of tokens, also known as running words, in the MAC subcorpus, (2) the word families found in each word band of the BNC/COCA frequency-based word lists (Nation, 2017), (3) the text coverage of the BNC/COCA lists provided in the subcorpus, and (4) the cumulative text coverage of the BNC/COCA frequency-based word family lists.

Table 4.2

Word Levels, Tokens, Number of Word Families, Text Coverage, and Cumulative Text Coverage of the BNC/COCA Word Family Lists in the MAC Subcorpus

BNC/COCA Word List	Tokens	Number of Word Families	Text Coverage	Cumulative Text Coverage
Proper Nouns	3,794	-	1.47%	1.47%
1st 1,000	184,658	895	71.38%	72.85%
2nd 1,000	37,026	728	14.31%	87.16%
3rd 1,000	**22,863**	**748**	**8.84%**	***96.00%**
4th 1,000	3,403	411	1.32%	97.32%
5th 1,000	**3,145**	**276**	**1.22%**	****98.53%**
6th 1,000	578	183	0.22%	98.76%
7th 1,000	509	117	0.20%	98.95%
8th 1,000	638	99	0.25%	99.20%
9th 1,000	95	45	0.04%	99.24%
10th 1,000	108	53	0.04%	99.28%
11th 1,000	272	32	0.11%	99.38%
12th 1,000	85	22	0.03%	99.42%
13th 1,000	107	19	0.04%	99.46%
14th 1,000	61	15	0.02%	99.48%
15th 1,000	21	7	0.01%	99.49%
16th 1,000	34	12	0.01%	99.50%
17th 1,000	24	12	0.01%	99.51%
18th 1,000	4	4	0.00%	99.51%
19th 1,000	22	7	0.01%	99.52%

BNC/COCA Word List	Tokens	Number of Word Families	Text Coverage	Cumulative Text Coverage
20th 1,000	23	3	0.01%	99.53%
21st 1,000	3	2	0.00%	99.53%
22nd 1,000	19	1	0.01%	99.54%
23rd 1,000	22	2	0.01%	99.55%
24th 1,000	0	0	0.00%	99.55%
25th 1,000	1	1	0.00%	99.55%
Off-List	1,171	-	0.45%	100.00%
Total	**258,686**	**-**	**100.00%**	**100.00%**

(Remarks: *meaning that 95% of text coverage was reached, **meaning that 98% of text coverage was reached)

Table 4.2 demonstrates (1) the total tokens in the MAC subcorpus and the tokens in each BNC/COCA frequency-based word family list, (2) the number of word families, (3) the text coverage, and (4) the cumulative text coverage of the BNC/COCA frequency-based word family lists over the textbook of the economic core course Principles of Macroeconomics representing the MAC subcorpus in the study.

The data indicated that the MAC subcorpus comprised 258,686 tokens (running words) in total. The cumulative text coverage of the BNC/COCA frequency-based word family lists has shown that the 3rd 1,000 BNC/COCA word family list has already reached 96.00% text coverage in the MAC subcorpus in order to have minimal comprehension (where minimally acceptable comprehension can occur) at 95% as suggested by Laufer and Ravenhorst-Kalovski (2010), resulting in having a vocabulary size of around 2,500–3,000 word families. The 5th 1,000-word level of the BNC/COCA lists already reaches 98.53% text coverage in the MAC subcorpus, and it can be inferred that having vocabulary knowledge of approximately 5,500–6,000 word families is necessary to have optimal comprehension of the texts at 98%.

4.1.3 The Text Coverage of the BNC/COCA Frequency-Based Word Lists and the Estimated Vocabulary Size Needed to Read in the Calculus for Social Science 1 (CAL) Subcorpus

One of the economics core course textbooks entitled *Calculus for Social Science 1* of the Bachelor of Economics International Program (BEIP) was selected to represent the economics textbook subcorpus, hereafter the CAL subcorpus,

which was the compulsory textbook needed to read as listed in the first year of study of the 2018 BEIP curriculum.

This part of the investigation is primarily aimed at answering the first two research questions concerning (1) the text coverage provided by the BNC/COCA frequency-based word family lists in the BEIP's first-year undergraduate economics textbook (the CAL subcorpus) and (2) the vocabulary size needed to read the textbook (the CAL subcorpus). The CAL subcorpus was analyzed to investigate (1) the number of tokens, also known as running words, in the CAL subcorpus, (2) the word families found in each word band of the BNC/COCA word family lists developed by Nation (2016), (3) the text coverage of the BNC/COCA lists provided in the CAL subcorpus, and (4) the cumulative text coverage of the BNC/COCA frequency-based word family lists.

Table 4.3

Word Levels, Tokens, Number of Word Families, Text Coverage, and Cumulative Text Coverage of the BNC/COCA Word Family Lists in the CAL Subcorpus

BNC/COCA Word List	Tokens	Number of Word Families	Text Coverage	Cumulative Text Coverage
Proper Nouns	134,529	-	26.77%	26.77%
1st 1,000	263,479	821	52.43%	79.20%
2nd 1,000	34,665	615	6.90%	86.09%
3rd 1,000	29,610	539	5.89%	91.99%
4th 1,000	10,727	298	2.13%	94.12%
5th 1,000	**5,311**	**206**	**1.06%**	***95.18%**
6th 1,000	2,609	130	0.52%	95.70%
7th 1,000	844	105	0.17%	95.87%
8th 1,000	1,848	92	0.37%	96.23%
9th 1,000	1,961	62	0.39%	96.62%
10th 1,000	451	44	0.09%	96.71%
11th 1,000	524	38	0.10%	96.82%
12th 1,000	571	30	0.11%	96.93%
13th 1,000	645	31	0.13%	97.06%
14th 1,000	184	15	0.04%	97.10%
15th 1,000	141	20	0.03%	97.12%
16th 1,000	326	26	0.06%	97.19%
17th 1,000	58	16	0.01%	97.20%
18th 1,000	158	7	0.03%	97.23%
19th 1,000	31	6	0.01%	97.24%

BNC/COCA Word List	Tokens	Number of Word Families	Text Coverage	Cumulative Text Coverage
20th 1,000	159	8	0.03%	97.27%
21st 1,000	31	8	0.01%	97.28%
22nd 1,000	171	6	0.03%	97.31%
23rd 1,000	16	5	0.00%	97.31%
24th 1,000	9	3	0.00%	97.31%
25th 1,000	235	13	0.05%	97.36%
Off-List	13,261	-	2.64%	100.00%
Total	**502,554**	**-**	**100.00%**	**100.00%**

(Remarks: *meaning that 95% of text coverage was reached, **meaning that 98% of text coverage was reached)

Table 4.3 demonstrates (1) the total tokens in the CAL subcorpus and the tokens in each BNC/COCA word family list, (2) the number of word families, (3) the text coverage, and (4) the cumulative text coverage of the BNC/COCA frequency-based word family lists over the textbook of the economics core course entitled *calculus for social science 1,* representing the CAL subcorpus in this current study.

The data indicated that the CAL subcorpus comprised 502,554 tokens (running words) in total. The cumulative text coverage of the BNC/COCA frequency-based word lists has shown that the 5th 1,000 BNC/COCA word family list already constituted 95.18% of the CAL subcorpus in order to have minimal comprehension (where minimally acceptable comprehension can occur) at 95%, as suggested by Laufer and Ravenhorst-Kalovski (2010), and a vocabulary size of around 4,500 to 5,000 word families is needed to read the text. However, having optimal comprehension of the texts at 98% was unable to be identified from the BNC/COCA frequency-based word lists due to the relatively high percentage of off-list words.

4.1.4 The Text Coverage of the BNC/COCA Frequency-Based Word Lists and the Estimated Vocabulary Size Needed to Read in the Statistics for Social Science 1 (STA) Subcorpus

One of the economics core course textbooks entitled *Statistics for Social Science 1* of the Bachelor of Economics International Program (BEIP) was selected to represent the economics textbook subcorpus, hereafter the STA subcorpus, which was the compulsory textbook needed to read as listed in the first year of study of the 2018 BEIP curriculum.

This part of the investigation is primarily aimed at answering the first two research questions regarding: (1) the text coverage that the BNC/COCA frequency-based word lists provide in the BEIP's first-year undergraduate economics textbook (the STA subcorpus); and (2) the vocabulary size needed to read the textbook (the STA subcorpus). The STA subcorpus was analyzed to investigate (1) the number of tokens, also known as running words, in the STA subcorpus; (2) the word families found in each word band of the BNC/COCA word family lists developed by Nation (2016); (3) the text coverage of the BNC/COCA lists provided in the STA subcorpus; and (4) the cumulative text coverage of the BNC/COCA frequency-based word family lists.

Table 4.4

Word Levels, Tokens, Number of Word Families, Text Coverage, and Cumulative Text Coverage of the BNC/COCA Word Family Lists in the STA Subcorpus

BNC/COCA Word List	Tokens	Number of Word Families	Text Coverage	Cumulative Text Coverage
Proper nouns	8,086	-	2.35%	2.35%
1st 1,000	227,275	865	65.98%	68.33%
2nd 1,000	46,394	660	13.47%	81.80%
3rd 1,000	42,186	673	12.25%	94.04%
4th 1,000	**6,507**	**348**	**1.89%**	***95.93%**
5th 1,000	3,277	217	0.95%	96.88%
6th 1,000	1,618	136	0.47%	97.35%
7th 1,000	870	90	0.25%	97.61%
8th 1,000	618	85	0.18%	97.79%
9th 1,000	475	45	0.14%	97.92%
10th 1,000	197	50	0.06%	97.98%
11th 1,000	**242**	**29**	**0.07%**	****98.05%**
12th 1,000	164	20	0.05%	98.10%
13th 1,000	467	26	0.14%	98.24%
14th 1,000	335	17	0.10%	98.33%
15th 1,000	71	17	0.02%	98.35%
16th 1,000	50	10	0.01%	98.37%
17th 1,000	96	4	0.03%	98.40%
18th 1,000	14	8	0.00%	98.40%
19th 1,000	9	7	0.00%	98.40%
20th 1,000	5	2	0.00%	98.40%
21st 1,000	9	3	0.00%	98.41%
22nd 1,000	10	5	0.00%	98.41%
23rd 1,000	25	4	0.01%	98.42%
24th 1,000	3	2	0.00%	98.42%

BNC/COCA Word List	Tokens	Number of Word Families	Text Coverage	Cumulative Text Coverage
25th 1,000	10	3	0.00%	98.42%
Off-List	5,442	**-**	1.58%	100.00%
Total	**344,455**	**-**	**100.00%**	**100.00%**

(Remarks: *meaning that 95% of text coverage was reached, **meaning that 98% of text coverage was reached)

Table 4.4 includes (1) the total tokens in the STA subcorpus and the tokens in each BNC/COCA word family list; (2) the number of word families; (3) the text coverage; and (4) the cumulative text coverage of the BNC/COCA frequency-based word family lists over the *statistics for the Social Science 1* textbook, which represents the STA subcorpus in this current study.

According to the data, there were 344,455 tokens (running words) in the STA subcorpus. The cumulative text coverage of the BNC/COCA frequency-based word lists revealed that the 1st 4,000 most frequent word families of the English language already comprised 95.93% of the STA subcorpus in order to have minimal comprehension (where minimally acceptable comprehension can occur) at 95%, as suggested by Laufer and Ravenhorst-Kalovski (2010), and a vocabulary size of around 3,500 to 4,000 word families is needed to read the text. In addition, having a vocabulary size of approximately 10,500 to 11,000 word families is necessary to have optimal comprehension of the texts at 98%, as the 11th 1,000 band of the BNC/COCA word family list constituted 98.04% of text coverage in the STA subcorpus.

4.1.5 The Text Coverage of the BNC/COCA Frequency-Based Word Family Lists and the Estimated Vocabulary Size Needed to Read in the ECON Corpus

The textbooks, which were selected to represent the economics core course textbooks needed to read in the first-year study of the Bachelor of Economics International Program (BEIP), were as follows:

(1) the textbook of the course *Principles of Microeconomics*,

(2) the textbook of the course *Principles of Macroeconomics*,

(3) the textbook of the course C*alculus for Social Science 1*, and

(4) the textbook of the course *Statistics for Social Science 1*.

This part contributes to the investigation concerning the text coverage in which the BNC/COCA frequency-based word family lists appear in the aforementioned first-year undergraduate economics textbooks of the BEIP curriculum. A corpus of 1,343,493 words from the textbooks was compiled and analyzed to illustrate: (1) the text coverage of the BNC/COCA frequency-based word family lists in each subcorpus and the whole corpus; and (2) the cumulative text coverage of the BNC/COCA frequency-based word family lists in each subcorpus and the ECON corpus.

Table 4.5

Word Levels, Text Coverage, and Cumulative Text Coverage of the BNC/COCA Word Family Lists in Each Subcorpus and the ECON Corpus

BNC/COCA Word List	MIC Subcorpus	MAC Subcorpus	CAL Subcorpus	STA Subcorpus	ECON Corpus
	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative Text Coverage
Proper Nouns	1.20% 1.20%	5.06% 5.06%	26.77% 26.77%	2.35% 2.35%	8.85% 8.85%
1st 1,000	72.32% 73.51%	68.36% 73.42%	52.43% 79.20%	65.98% 68.33%	64.77% 73.62%
2nd 1,000	13.53% 87.05%	13.62% 87.04%	6.90% 86.09%	13.47% 81.80%	11.88% 85.50%
3rd 1,000	8.62% ***95.66%**	8.35% ***95.39%**	5.89% 91.99%	12.25% 94.04%	8.78% 94.27%
4th 1,000	2.02% 97.69%	1.27% 96.66%	2.13% 94.12%	1.89% ***95.93%**	1.83% ***96.10%**
5th 1,000	1.00% ****98.69%**	1.15% 97.81%	1.06% ***95.18%**	0.95% 96.88%	1.04% 97.14%
6th 1,000	0.30% 98.98%	0.23% ****98.04%**	0.52% 95.70%	0.47% 97.35%	0.38% 97.52%
7th 1,000	0.18% 99.17%	0.19% 98.23%	0.17% 95.87%	0.25% 97.61%	0.20% 97.72%
8th 1,000	0.16% 99.33%	0.25% 98.48%	0.37% 96.23%	0.18% 97.79%	0.24% 97.96%
9th 1,000	0.03% 99.36%	0.05% 98.53%	0.39% 96.62%	0.14% 97.92%	0.15% ****98.11%**
10th 1,000	0.05% 99.40%	0.05% 98.58%	0.09% 96.71%	0.06% 97.98%	0.06% 98.17%
11th 1,000	0.04%	0.11%	0.10%	0.07%	0.08%

BNC/COCA Word List	MIC Subcorpus	MAC Subcorpus	CAL Subcorpus	STA Subcorpus	ECON Corpus
	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative Text Coverage
	99.44%	98.69%	96.82%	**98.05%	98.25%
12th 1,000	0.02% 99.46%	0.04% 98.73%	0.11% 96.93%	0.05% 98.10%	0.06% 98.31%
13th 1,000	0.03% 99.49%	0.04% 98.77%	0.13% 97.06%	0.14% 98.24%	0.09% 98.39%
14th 1,000	0.05% 99.54%	0.07% 98.84%	0.04% 97.10%	0.10% 98.33%	0.07% 98.45%
15th 1,000	0.02% 99.56%	0.01% 98.85%	0.03% 97.12%	0.02% 98.35%	0.02% 98.47%
16th 1,000	0.00% 99.56%	0.02% 98.87%	0.06% 97.19%	0.01% 98.37%	0.02% 98.50%
17th 1,000	0.01% 99.57%	0.02% 98.89%	0.01% 97.20%	0.03% 98.40%	0.02% 98.52%
18th 1,000	0.01% 99.58%	0.01% 98.90%	0.03% 97.23%	0.00% 98.40%	0.01% 98.53%
19th 1,000	0.00% 99.58%	0.01% 98.91%	0.01% 97.24%	0.00% 98.40%	0.01% 98.54%
20th 1,000	0.01% 99.59%	0.01% 98.92%	0.03% 97.27%	0.00% 98.40%	0.01% 98.55%
21st 1,000	0.00% 99.59%	0.00% 98.92%	0.01% 97.28%	0.00% 98.41%	0.00% 98.55%
22nd 1,000	0.00% 99.60%	0.01% 98.93%	0.03% 97.31%	0.00% 98.41%	0.01% 98.56%
23rd 1,000	0.00% 99.60%	0.02% 98.95%	0.00% 97.31%	0.01% 98.42%	0.01% 98.57%
24th 1,000	0.00% 99.60%	0.00% 98.95%	0.00% 97.31%	0.00% 98.42%	0.00% 98.57%
25th 1,000	0.00% 99.60%	0.00% 98.95%	0.05% 97.36%	0.00% 98.42%	0.01% 98.58%
Off-List	0.40% 100.00%	1.05% 100.00%	2.64% 100.00%	1.58% 100.00%	1.42% 100.00%
Total	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**

(Remarks: *meaning that 95% of text coverage was reached, **meaning that 98% of text coverage was reached.)

Table 4.5 demonstrates (1) text coverage at each subcorpus of the BNC/COCA frequency-based word family lists in the textbooks across the first-year economics core courses of the Bachelor of Economics International Program (BEIP)

and (2) the cumulative text coverage of the BNC/COCA frequency-based word family lists in the textbooks.

In the MIC subcorpus, the results have shown that the 3rd 1,000 BNC/COCA word family list reached 95.66% and the 5th 1,000 BNC/COCA word family list reached 98.69% of text coverage. In the MAC subcorpus, the 3rd 1,000 BNC/COCA word family list reached 95.39%, and the 6th 1,000 BNC/COCA word family list reached 98.04%. In the CAL subcorpus, the 5th 1,000 BNC/COCA word family list reached 95.18%; however, 98% of text coverage was not reached. In the STA subcorpus, the 4th 1,000 BNC/COCA word family list reached 96.10% and the 11th 1,000 BNC/COCA word family list reached 98.05% of text coverage. As a result, in the ECON corpus, the 4th 1,000 BNC/COCA word family list reached 96.10%, resulting in having a vocabulary size of around 3,500–4,000 word families, and the 9th 1,000 BNC/COCA word family list reached 98.11%, resulting in having a vocabulary size of around 8,500–9,000 word families.

4.2 The Text Coverage of the Academic Word List in the Corpus of First-Year Undergraduate Economics Textbooks

In this section, the findings of the current study indicate (1) tokens, also known as running words, which are known and unknown words that appear in the text (corpus); (2) word families, which refer to groups of words sharing a common root word with different derivations (e.g., prefixes and suffixes) and inflections (different forms); and (3) text coverage, meaning the percentage of known words in the text, which may indicate text comprehension. As mentioned in the literature review chapter, text comprehension and vocabulary knowledge are necessary for reading comprehension. Knowing the approximate percentage of texts that readers may understand is therefore helpful for course designers and material developers.

This section aims to present the findings of the text coverage of the Academic Word List in the textbooks as the whole corpus (ECON corpus) and subcorpora of the four different economics core courses as follows: Principles of Microeconomics (MIC), Principles of Macroeconomics (MAC), Calculus for Social Science 1 (CAL), and Statistics for Social Science 1 (STA).

4.2.1 The Text Coverage of the Academic Word List in the Principles of Microeconomics (MIC) Subcorpus

This part of the investigation is primarily aimed at answering the first research question concerning the text coverage that the Academic Word List (hereafter the AWL) provides in the first-year undergraduate economics textbooks of the BEIP'. The researcher analyzed the MIC subcorpus and obtained the following data: (1) the tokens, also known as running words, in the MIC subcorpus; (2) the word families of the AWL (Coxhead, 2000) in the MIC subcorpus; (3) the text coverage of the AWL in the MIC subcorpus; and (4) the cumulative text coverage of the AWL in the MIC subcorpus.

Table 4.6

Word Levels, Tokens, Number of Word Families, Text Coverage, and Cumulative Text Coverage of the GSL and AWL in the MIC Subcorpus

Word List	Tokens (Running Words)	Number of Word Families	Text Coverage	Cumulative Text Coverage
General Service List (GSL)				
1st 1,000 GSL	184,025	1,034	80.62%	80.62%
2nd 1,000 GSL	13,216	659	5.79%	86.41%
Academic Word List	19,920	576	8.73%	95.13%
Off-list	11,113	-	7.42%	100.00%
Total	**228,274**	-	**100.00%**	**100.00%**

Table 4.6 shows the word lists of the 1st 1,000 General Service List (GSL), the 2nd 1,000 General Service List (GSL), and the Academic Word List (AWL), which are provided in the microeconomics textbook subcorpus, hereafter the MIC subcorpus.

The data indicated that the MIC subcorpus contained 228,274 tokens in total. The 1st 1,000 GSL constituted 80.62% (184,025 tokens), and the 2nd 1,000 GSL constituted 5.79% (13,216 tokens) of text coverage in the MIC subcorpus. Therefore, the cumulative text coverage of the 1st 1,000 GSL and the 2nd 1,000 GSL was 86.41% (197,241 tokens).

The researcher used the GSL (for general English words) and the AWL (for academic English words) as base lists. However, this part was primarily

aimed at focusing on the text coverage of the AWL in the MIC subcorpus. In other words, the AWL was used to estimate the percentage of the AWL words appearing in the textbook of the *Principles of Microeconomics* core course. It gave the text coverage of the AWL at 8.73% (19,920 tokens) in the MIC subcorpus. Lastly, the off-list words had 7.42% (11,113 tokens) of text coverage. Off-list words presented in this part refer to words that did not appear in any base lists of the GSL or the AWL. They did not belong to the aforementioned word lists.

4.2.2 The Text Coverage of the Academic Word List in the Principles of Macroeconomics (MAC) Subcorpus

This part of the investigation is primarily aimed at answering the first research question concerning the text coverage that the Academic Word List, hereafter the AWL, provides in the first-year undergraduate economics textbooks of the BEIP. The MAC subcorpus was analyzed through the AntWordProfiler software program (Anthony, 2022) to explore as follows: (1) the tokens, also known as running words, in the MAC subcorpus; (2) the word families of the AWL developed by Coxhead (2000) appearing in the MAC subcorpus; (3) the text coverage of the AWL providing in the MAC subcorpus; and (4) the cumulative text coverage of the AWL providing in the MAC subcorpus.

Table 4.7

Word Levels, Tokens, Number of Word Families, Text Coverage, and Cumulative Text Coverage of the GSL and AWL in the MAC Subcorpus

Word List	Tokens (Running Words)	Number of Word Families	Text Coverage	Cumulative Text Coverage
General Service List (GSL)				
1st 1,000 GSL	208,603	1,001	77.78%	77.78%
2nd 1,000 GSL	12,399	602	4.62%	82.40%
Academic Word List	25,552	578	9.53%	91.93%
Off-List	21,669	-	8.07%	100.00%
Total	**268,210**	**-**	**100.00%**	**100.00%**

Table 4.7 shows the word lists of the 1st 1000 General Service List (GSL), the 2nd 1000 General Service List (GSL), and the Academic Word List (AWL),

which appear in the macroeconomics textbook subcorpus, hereafter the MAC subcorpus.

The findings indicated that the MAC subcorpus consisted of 268,210 tokens in total. The 1st 1,000 GSL constituted 77.78% of text coverage, and there were 1,001 word families found in the MAC subcorpus. The 2nd 1,000 GSL constituted 4.62% of text coverage, with 602 word families found in the MAC subcorpus. Consequently, the GSL words, including the 1st 1,000 GSL and the 2nd 1,000 GSL, gave approximately 82.40% of text coverage in the MAC subcorpus. The AWL provided 9.53% of text coverage in the MAC subcorpus. Lastly, the off-list words had 8.07% of the text coverage.

The word lists of the GSL (for general English words) and the AWL (for academic English words) were used as base lists. However, in order to answer the first research question, the study was aimed at focusing on the text coverage of the AWL. In other words, the AWL was used to estimate the percentage of AWL words appearing in the textbook of the *Principles of Macroeconomics* core course.

4.2.3 The Text Coverage of the Academic Word List in the Calculus for Social Science 1 (CAL) Subcorpus

The textbook of the economics core course entitled *Calculus for Social Science 1* of the Bachelor of Economics International Program (BEIP) was selected to represent the economics textbook subcorpus, hereafter the CAL subcorpus, which was retrieved from the compulsory textbook needed to read, as listed in the first year of study of the 2018 BEIP curriculum.

This investigation part was intended to answer the first research question concerning the text coverage that the Academic Word List, hereafter the AWL, provides in the first-year undergraduate economics textbooks of the BEIP. Consequently, the CAL subcorpus was analyzed through the AntWordProfiler software program (Anthony, 2022) in order to explore the following: (1) the total number of tokens, also known as running words, in the CAL subcorpus; (2) the word families of the AWL developed by Coxhead (2000) appearing in the CAL subcorpus; and (3) the text coverage, including the text coverage in each word list and the cumulative text coverage, of the AWL in the CAL subcorpus.

Table 4.8

Word Levels, Tokens, Number of Word Families, Text Coverage, and Cumulative Text Coverage of the GSL and AWL in the CAL Subcorpus

Word List	Tokens (Running Words)	Number of Word Families	Text Coverage	Cumulative Text Coverage
General Service List (GSL)				
1st 1,000 GSL	335,736	915	66.81%	66.81%
2nd 1,000 GSL	15,583	490	3.10%	69.91%
Academic Word List	32,384	458	6.44%	76.35%
Off-List	118,851	-	23.65%	100.00%
Total	**502,554**	**-**	**100.00%**	**100.00%**

Table 4.8 illustrates data regarding (1) the tokens, also known as running words, appearing in total and in the word lists of the 1st 1,000 General Service List (GSL), the 2nd 1,000 General Service List (GSL), and the Academic Word List (AWL) in the calculus textbook subcorpus (the CAL subcorpus), (2) the numbers of word families of each base word list in the CAL subcorpus, (3) the text coverage, meaning the percentage of words known in the CAL subcorpus, and (4) the cumulative text coverage of the CAL subcorpus.

The findings indicated that the CAL subcorpus consisted of 502,554 tokens in total. The 1st 1,000 GSL constituted 66.81% of text coverage, and there were 915 word families found in the CAL subcorpus. The 2nd 1,000 GSL constituted 3.10% of text coverage, with 490 word families found in the CAL subcorpus. Consequently, the GSL words provided approximately 69.91% of the text coverage in the CAL subcorpus. In addition, the word lists of the GSL (for general English words) and the AWL (for academic English words) were used as base lists. According to the concentration on answering the first research question, the study was aimed at focusing on the text coverage of the AWL. In other words, the AWL was used to estimate the percentage of AWL words appearing in the textbook of the economics core course entitled *Calculus for Social Science 1*. The AWL gave 6.44% text coverage in the CAL subcorpus. Lastly, the off-list words had 23.65% of the text coverage. Off-list words in this study refer to the words that did not appear in any of the aforementioned word lists, i.e., the GSL and the AWL. In other words, they did not belong to the base lists.

4.2.4 The Text Coverage of the Academic Word List in the Statistics for Social Science 1 (STA) Subcorpus

The textbook of the economics core course entitled *Statistics for Social Science 1* of the Bachelor of Economics International Program (BEIP) was selected to represent the economics textbook subcorpus, hereafter the STA subcorpus, which was retrieved from the compulsory textbook needed to read, as listed in the first year of study of the 2018 BEIP curriculum.

This investigation part was intended to answer the first research question concerning the text coverage that the Academic Word List, hereafter the AWL, provides in the first-year undergraduate economics textbooks of the BEIP. Consequently, the STA subcorpus was analyzed through the AntWordProfiler software program (Anthony, 2022) in order to explore the following: (1) the total number of tokens, also known as running words, in the STA subcorpus; (2) the word families of the AWL developed by Coxhead (2000) appearing in the STA subcorpus; and (3) the text coverage, including the text coverage in each word list and the cumulative text coverage, of the AWL in the STA subcorpus.

Table 4.9

Word Levels, Tokens, Number of Word Families, Text Coverage, and Cumulative Text Coverage of the GSL and AWL in the STA Subcorpus

Word List	Tokens (Running Words)	Number of Word Families	Text Coverage	Cumulative Text Coverage
General Service List (GSL)				
1st 1,000 GSL	255,580	970	74.20%	74.20%
2nd 1,000 GSL	20,426	508	5.93%	80.13%
Academic Word List	47,544	550	13.80%	93.93%
Off-List	20,905	-	6.07%	100.00%
Total	**344,455**	**-**	**100.00%**	**100.00%**

Table 4.9 illustrates data regarding (1) the tokens, also known as running words, appearing in total and in the word lists of the 1st 1,000 General Service List (GSL), the 2nd 1,000 General Service List (GSL), and the Academic Word List (AWL) in the statistics textbook subcorpus (the STA subcorpus), (2) the numbers of word families of each base word list in the STA subcorpus, (3) the text coverage,

meaning the percentage of words known in the STA subcorpus, and (4) the cumulative text coverage of the STA subcorpus.

The findings indicated that the STA subcorpus comprised 344,455 tokens in total. The 1st 1,000 GSL constituted 74.20% of text coverage, and there were 970 word families found in the STA subcorpus. The 2nd 1,000 GSL constituted 5.93% of text coverage, with 508 word families found in the STA subcorpus. Consequently, the 2,000 GSL words provided approximately 80.13% of the text coverage in the STA subcorpus. In addition, the word lists of the GSL (for general English words) and the AWL (for academic English words) were used as base lists. According to the concentration on answering the first research question, the study was aimed at focusing on the text coverage of the AWL. In other words, the AWL was used to estimate the percentage of AWL words appearing in the textbook of the economics core course entitled *Statistics for Social Science 1*. The AWL gave 13.80% text coverage in the STA subcorpus. Lastly, the off-list words had 6.07% of the text coverage. Off-list words in this study refer to the words that did not appear in any of the aforementioned word lists, i.e., the GSL and the AWL. In other words, they did not belong to the base lists.

4.2.5 The Text Coverage of the Academic Word List in the ECON Corpus

The textbooks, which were selected to represent the economics core course textbooks needed to read in the first-year study of the Bachelor of Economics International Program (BEIP), were as follows:

(1) the textbook of the course *Principles of Microeconomics*,
(2) the textbook of the course *Principles of Macroeconomics*,
(3) the textbook of the course C*alculus for Social Science 1*, and
(4) the textbook of the course *Statistics for Social Science 1*.

This part contributes to the investigation concerning the text coverage of the Academic Word List in the aforementioned first-year undergraduate economics textbooks of the BEIP curriculum, and therefore a 1,343,493-word corpus of the textbooks was compiled and analyzed to report: (1) text coverage of the General Service List (GSL) and the Academic Word List (AWL) in each subcorpus and the whole corpus; and (2) cumulative text coverage of the GSL and AWL in each subcorpus and the ECON corpus.

Table 4.10

Word Levels, Text Coverage, and Cumulative Text Coverage of the GSL and AWL in Each Subcorpus and the ECON Corpus

Word List	MIC Subcorpus	MAC Subcorpus	CAL Subcorpus	STA Subcorpus	ECON Corpus
	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative Text Coverage	Text Coverage/ Cumulative text coverage
General Service List (GSL)	80.62%	77.78%	66.81%	74.20%	74.85%
1st 1,000 GSL	80.62%	77.78%	66.81%	74.20%	74.85%
2nd 1,000 GSL	5.79%	4.62%	3.10%	5.93%	4.86%
	86.41%	82.40%	69.91%	80.13%	79.71%
Academic Word List	**8.73%**	**9.53%**	**6.44%**	**13.80%**	**9.62%**
	95.13%	91.93%	76.35%	93.93%	89.33%
Off-List	7.42%	8.07%	23.65%	6.07%	10.67%
	100.00%	100.00%	100.00%	100.00%	100.00%
Total	**100.00%**	**100.00%**	**100.00%**	**100.00%**	**100.00%**

Table 4.10 shows (1) the text coverage of the General Service List (GSL), the Academic Word List (widely known as AWL) in the four subcorpora, and the ECON corpus representing the textbooks of the first-year undergraduate economics core courses of the Bachelor of Economics International Program (BEIP), and (2) the cumulative text coverage of the GSL and the AWL in the four subcorpora and the ECON corpus in the textbooks of the first-year undergraduate economics core courses of the BEIP. In addition, the following word lists were used as base lists in this study: (1) the General Service List (GSL) as the vocabulary list for general English words, and (2) the Academic Word List (AWL) as the vocabulary list for academic English words. To answer the first research question regarding the AWL found in the textbooks, this section primarily reported the text coverage and the cumulative text coverage of the AWL in each text subcorpus and the whole corpus.

According to the results of GSL's text coverage, the subcorpora in the study were as follows: (1) the subcorpus of the course *principles of microeconomics* (hereafter referred to as the MIC subcorpus) comprising 80.62% and 5.79% of the 1st 1,000 GSL and 2nd 1,000 GSL of text coverage and resulting in the cumulative text coverage of the GSL at 86.41%, (2) the subcorpus of the course *principles of*

macroeconomics (hereafter referred to as the MAC subcorpus) comprising 77.78% and 4.62% of the 1st 1,000 GSL and 2nd 1,000 GSL of text coverage with 82.40% of the cumulative text coverage, (3) the subcorpus of the course *calculus for social science 1* (hereafter referred to as the CAL subcorpus) comprising 66.81% and 3.10% of the 1st 1,000 GSL and 2nd 1,000 GSL of text coverage and resulting in the cumulative text coverage of the GSL at 69.91%, and (4) the subcorpus of the course *statistics for social science 1* (hereafter referred to as the STA subcorpus) comprising 74.20% and 5.93% of the 1st 1,000 GSL and 2nd 1,000 GSL of text coverage and the cumulative text coverage at 80.13%. According to the text coverage of GSL in the whole corpus (ECON corpus) in this study, the 1st 1,000 GSL and the 2nd 1,000 GSL constituted 74.85% and 4.86%, respectively, of the text coverage. Consequently, the GSL constituted 79.71% of the text coverage in the ECON corpus.

According to the results of the text coverage analysis of the AWL in each text subcorpus and the ECON corpus, the following subcorpora were analyzed as follows: (1) the AWL is given 8.73% of text coverage in the MIC subcorpus; (2) the AWL is given 9.53% of text coverage in the MAC subcorpus; (3) the AWL is given 6.44% of text coverage in the CAL subcorpus; and (4) the STA subcorpus is given 13.80% of the AWL text coverage. In the ECON corpus, the AWL therefore gave the text coverage of 9.62%.

However, there were some words that did not belong to the base word lists, i.e., the GSL and the AWL. Those words were categorized into the out-of-the-list group (also known as the off-list words). According to the tables, the off-list words in the ECON corpus in total were 10.67%. The off-list words in the following subcorpora: the MIC subcorpus, the MAC subcorpus, the CAL subcorpus, and the STA subcorpus, were at 7.42%, 8.07%, 23.65%, and 6.07%, respectively.

4.3 The Most Frequently Occurring AWL Academic Words in the First-Year Undergraduate Economics Textbooks

Identifying the most frequently occurring Academic Word List (AWL) academic words in the first-year BEIP textbooks was one of the main objectives of the present study. The software program AntWordProfiler (Anthony, 2022) was used to

identify the most common AWL academic words which can be used in academic contexts, especially written academic contexts, in the textbooks. The researcher explored some other previous studies on how to theoretically process the target vocabulary, and continued to study the subcorpora and the ECON corpus.

The Academic Word List, developed by Coxhead (2000), contains 570 word families selected for their frequency, more specifically, their high frequency in academic contexts. In addition, the words in the list do not belong to the General Service List, which contains the most frequent 2,000 English general words. The requirement for frequency uniformity according to Coxhead (2000) was 100 occurrences overall with at least ten in each of four fields in the 3.5-million-word corpus. Across the 1.5 million words in the corpus, Vongpumivitch et al. (2009) stipulated that a list of 475 AWL appeared at least 50 times overall and five times in each subcorpus. In a corpus of 508,802 words, Hajiyeva (2015) estimated that the word occurrences would appear 16 times overall and at least twice of which would appear in each subcorpus.

The third research question in this study will identify the AWL academic words used most frequently in the BEIP's first-year undergraduate textbooks. To begin addressing this question, a list of AWL academic words that meet the standardized criteria for frequency uniformity according to other studies, which include word frequency, needs to be compiled. Based on the Coxhead's (2000) criteria for word selection of AWL words, more specifically, and the occurrences were then determined. After the mathematical calculation, the frequency of occurrence of AWL academic words should be at least 38 times in the ECON corpus and at least 4 times in each subcorpus.

As shown in Table 4.11, it was discovered that there were 187 AWL academic words, ranked in frequency order and they appeared at least 38 times overall in the whole corpus and at least 4 times across the subcorpora. The most frequent AWL academic words were data (2,978 times), function (2,895 times), equation (2,808 times), chapter (2,398 times), income (2,053 times), series (1,575 times), distribution (1,549 times), area (1,306 times), economy (1,232 times), estimate (1,222 times), labor (1,187 times), section (1,063 times), percent (981 times), compute (965 times), functions (905 times) and so on.

Table 4.11

187 Most Frequently Occurring AWL Academic Words

No.	Vocabulary Item	Frequency
1	data	2,978
2	function	2,895
3	equation	2,808
4	chapter	2,398
5	income	2,053
6	series	1,575
7	distribution	1,549
8	area	1,306
9	economy	1,232
10	estimate	1,222
11	labor	1,187
12	section	1,063
13	percent	981
14	compute	965
15	functions	905
16	marginal	863
17	hypothesis	858
18	sum	814
19	analysis	797
20	economic	796
21	positive	788
22	method	772
23	consumer	740
24	constant	732
25	period	728
26	normal	719
27	equations	713
28	revenue	708
29	consumption	690
30	investment	679
31	economics	672
32	computer	666
33	estimated	660
34	region	645
35	evaluate	555
36	process	548
37	percentage	532
38	maximum	524
39	assume	495
40	required	492
41	minimum	480
42	negative	479

No.	Vocabulary Item	Frequency
43	theory	449
44	available	431
45	methods	419
46	initial	408
47	obtain	391
48	conclusion	373
49	resources	373
50	corresponding	370
51	factor	370
52	instance	369
53	procedure	366
54	approach	353
55	similar	344
56	individual	342
57	range	322
58	ratio	320
59	shift	311
60	annual	304
61	illustrate	303
62	affect	261
63	predict	258
64	factors	257
65	items	252
66	survey	247
67	appendix	245
68	assumption	240
69	alternative	236
70	conclude	236
71	areas	235
72	computed	231
73	source	230
74	occurs	228
75	indicate	224
76	estimates	218
77	similarly	210
78	summary	210
79	major	208
80	construct	207
81	appropriate	203
82	computing	194
83	identify	186
84	individuals	185
85	approximately	181
86	indicates	180
87	assuming	173

No.	Vocabulary Item	Frequency
88	potential	170
89	approaches	163
90	coordinate	163
91	role	163
92	hence	158
93	previous	155
94	plus	153
95	periods	150
96	occur	147
97	portion	145
98	distributed	143
99	technology	142
100	sector	137
101	located	136
102	evidence	135
103	energy	133
104	focus	131
105	impact	131
106	strategy	131
107	comment	127
108	capacity	126
109	chapters	126
110	location	125
111	labeled	123
112	assumed	120
113	assumptions	120
114	consists	120
115	input	119
116	affects	118
117	illustrates	118
118	corresponds	116
119	identical	115
120	normally	115
121	designed	114
122	exceeds	114
123	illustrated	114
124	analyze	111
125	defined	110
126	project	109
127	summarize	109
128	goal	108
129	concept	107
130	computers	106
131	preceding	106
132	principles	106

No.	Vocabulary Item	Frequency
133	sources	106
134	requires	105
135	require	102
136	cycle	100
137	maximize	100
138	definition	99
139	physical	98
140	inputs	95
141	involved	95
142	item	94
143	create	93
144	finally	88
145	team	87
146	derived	83
147	conclusions	82
148	complex	80
149	consistent	79
150	initially	78
151	substitute	77
152	involves	75
153	evaluating	74
154	primary	74
155	decline	73
156	maintain	73
157	previously	73
158	achieve	70
159	indicating	70
160	exceed	69
161	environment	67
162	expand	67
163	ensure	66
164	eventually	66
165	published	65
166	analyzing	62
167	construction	60
168	summarized	60
169	created	58
170	task	58
171	derive	55
172	equivalent	53
173	involve	52
174	structure	52
175	eliminate	49
176	medium	47
177	features	46

No.	Vocabulary Item	Frequency
178	nonetheless	46
179	somewhat	45
180	aid	43
181	restrictions	43
182	ignore	40
183	insight	40
184	consequences	39
185	creating	39
186	proceed	39
187	restricted	39

Since the AWL academic words in the ECON corpus matched the word frequency criteria based on Coxhead's (2000) study, the words were then analyzed through Familizer (Cobb, 2020) in order to obtain a list of word family headwords. Because a word family includes the headword and its inflectional and derived forms, it is necessary to list these individual academic words in the form of headwords.

As you can see on Table 4.12, there were 133 AWL word family headwords, ranked in alphabetical order, in the ECON corpus of this current study.

Table 4.12

133 Most Frequently Occurring AWL Word Family Headwords (Sorting in Alphabetical Order)

No.	Vocabulary Item
1	achieve
2	affect
3	aid
4	alternative
5	analyse
6	annual
7	appendix
8	approach
9	appropriate
10	approximate
11	area
12	assume
13	assumption
14	available
15	capacity
16	chapter
17	comment

No.	Vocabulary Item
18	complex
19	computer
20	concept
21	conclude
22	conclusion
23	consequence
24	consist
25	consistent
26	constant
27	construct
28	consume
29	consumption
30	coordinate
31	correspond
32	create
33	cycle
34	data
35	decline
36	define
37	derive
38	design
39	distribute
40	economy
41	eliminate
42	energy
43	ensure
44	environment
45	equation
46	equivalent
47	estimate
48	evaluate
49	eventually
50	evidence
51	exceed
52	expand
53	factor
54	feature
55	final
56	focus
57	function
58	goal
59	hence
60	hypothesis
61	identical
62	identify

No.	Vocabulary Item
63	ignore
64	illustrate
65	impact
66	income
67	indicate
68	individual
69	initial
70	input
71	insight
72	instance
73	invest
74	involve
75	item
76	label
77	labour
78	locate
79	maintain
80	major
81	margin
82	maximise
83	maximum
84	medium
85	method
86	minimum
87	negative
88	nonetheless
89	normal
90	obtain
91	occur
92	percent
93	period
94	physical
95	plus
96	portion
97	positive
98	potential
99	precede
100	predict
101	previous
102	primary
103	principle
104	procedure
105	proceed
106	process
107	project

No.	Vocabulary Item
108	publish
109	range
110	ratio
111	region
112	require
113	resource
114	restrict
115	revenue
116	role
117	section
118	sector
119	series
120	shift
121	similar
122	somewhat
123	source
124	strategy
125	structure
126	substitute
127	sum
128	summary
129	survey
130	task
131	team
132	technology
133	theory

This chapter (Chapter 4) answered the research questions of this present study and then yielded the results of the BNC/COCA and AWL text coverage in the Bachelor of Economics International Program (BEIP)'s first year undergraduate textbooks, the vocabulary size needed to read the textbooks, and the most frequently occurring AWL academic words in frequency order as well as word family headwords in alphabetical order. The next chapter (Chapter 5) is going to summarize and interpret the results of the corpus-based vocabulary analysis which were reported in this chapter. It also includes the objectives, data collection, and findings. The results are then compared with prior research studies, followed by the recommendations for practice (implications) and the recommendations for further research.

CHAPTER 5
CONCLUSIONS AND DISCUSSIONS

This chapter comprises (1) a summary of the study, (2) discussion, (3) conclusions, (4) implications, and (5) recommendations for further research.

5.1 Summary of the Study

5.1.1 Objectives of the Study

This current study aims to (1) investigate the text coverage provided by the BNC/COCA Word Family Lists (Nation, 2017) and the Academic Word List (Coxhead, 2000) in the first-year Bachelor of Economics International Program (BEIP) textbooks; (2) investigate the vocabulary size required for reading the first-year BEIP textbooks; and (3) identify the most frequently used AWL academic words in the first-year BEIP textbooks. The chosen textbooks are intended to be used as a representation of the written texts in English that students need to know or be familiar with in an English as a Medium of Instruction (EMI) setting. The results also aim to provide a pedagogical view of the vocabulary that is needed to read the textbooks used in the BEIP.

5.1.2 Data Collection

The researcher selected the required economics textbooks used by first-year undergraduate students in the Bachelor of Economics International Program at an autonomous institution in Thailand, referred to as the "BEIP," in order to achieve the study's objectives. In addition, the textbooks were the required readings for every core course included in the BEIP study plan for first-year undergraduate students enrolled in the BEIP. The first year of economics core courses, for which BEIP students earn 12 credits, was the focus of the researcher's investigation. The BEIP introduced the following core economics courses during the first year of study: (1) "Principles of Microeconomics" (EE211); (2) "Principles of Macroeconomics" (EE212); (3) "Calculus for Social Science 1" (MA216); and (4) "Statistics for Social Science 1" (ST216). The texts provided in the economics core course textbooks ranged in length and were designed to offer students majoring in economics a grounding in economics,

an understanding of its guiding principles, and mastery of each discipline's particular subject.

Thus, the researcher compiled a corpus of first-year economics core course textbooks from each individual textbook. The texts were gathered in an electronically readable format in order to create the text corpus. AntFileConverter (Anthony, 2022) is a software program used to convert PDF and Microsoft Word formats into plain text for use in corpus tools. In other words, it is used to convert first-year undergraduate textbooks for the BEIP. The researcher afterwards converted texts into plain text (.txt) files for the following stage after removing all unnecessary parts. Then, the research prepared the software program AntWordProfiler (Anthony, 2022) to analyze the text corpus in the next stage.

5.1.3 Summary of the Results

To answer the first and second research questions about text coverage and vocabulary size, a corpus of 1,343,493 words from textbooks was compiled and analyzed to show the text coverage of the BNC/COCA frequency-based word family lists and the Academic Word List (AWL) in each subcorpus and the entire corpus, and the cumulative text coverage of the BNC/COCA frequency-based word family lists and the AWL in each subcorpus and the entire corpus. Identifying the most often used AWL academic words in the BEIP's first-year undergraduate textbooks was the third research question in this study.

5.1.3.1 Text Coverage of the BNC/COCA Frequency-Based Word Family Lists and the Estimated Vocabulary Size in the Textbooks

The textbooks selected as representing the economics core course textbooks required for first-year study in the Bachelor of Economics International Program (BEIP) were: (1) the textbook of the course *principles of microeconomics*; (2) the textbook of the course *principles of macroeconomics*; (3) the textbook of the course *calculus for social science 1*; and (4) the textbook of the course *statistics for social science 1*. As a result, the investigation was conducted in order to answer the first research question about text coverage, in which the BNC/COCA frequency-based word family lists appear in the aforementioned first-year undergraduate economics textbooks of the BEIP curriculum. As a result, a corpus of 1,343,493 words from textbooks was compiled and analyzed to show: (1) the text

coverage of the BNC/COCA frequency-based word family lists in each subcorpus and the entire corpus; and (2) the cumulative text coverage of the BNC/COCA frequency-based word family lists in each subcorpus and the entire corpus (ECON corpus).

For selecting vocabulary for study, the BNC/COCA word family lists are frequency-based lists that are frequently used. According to the findings, the text coverage of BNC/COCA word family lists in the ECON corpus is as follows: 1) the most frequent 1,000 words, including proper nouns, comprised 73.62%; 2) the most frequent 2,000 words comprised 85.50%; 3) the most frequent 3,000 words comprised 94.27%; 4) the most frequent 4,000 words comprised 96.10%; 5) the most frequent 5,000 words comprised 97.14%; 6) the most frequent 6,000 words comprised 97.52%; 7) the most frequent 7,000 words comprised 97.72%; 8) the most frequent 8,000 words comprised 97.96%; 9) the most frequent 9,000 words comprised 98.11%; 10) the most frequent 10,000 words comprised 98.17%; 11) the most frequent 11,000 words comprised 98.25%; 12) the most frequent 12,000 words comprised 98.31%; 13) the most frequent 13,000 words comprised 98.39%; 14) the most frequent 14,000 words comprised 98.45%; 15) the most frequent 15,000 words comprised 98.47%; 16) the most frequent 16,000 words comprised 98.50%; 17) the most frequent 17,000 words comprised 98.52%; 18) the most frequent 18,000 words comprised 98.53%; 19) the most frequent 19,000 words comprised 98.54%; 20) the most frequent 20,000 words comprised 98.55%; 21) the most frequent 21,000 words comprised 98.55%; 22) the most frequent 22,000 words comprised 98.56%; 23) the most frequent 23,000 words comprised 98.57%; 24) the most frequent 24,000 words comprised 98.57%; and 25) the most frequent 25,000 words comprised 98.58%.

The most frequent 1,000 words and proper nouns theoretically enable readers to understand 73.62% of the running words. The following 2nd 1,000 words and the following 3rd 1,000 give an additional 11.88% and 8.78 of the text, whereas the following 4th 1,000 words give only 1.83%. Each of the four subcorpora of the BNC/COCA frequency-based word family lists has a different vocabulary level for the textbooks. As a result, in the ECON corpus, the 1st 4,000 high-frequency word families comprised 96.10% of text coverage, which was more than 95 percent of text coverage, according to Laufer and Ravenhorst-Kalovski (2010) in the ECON corpus. In other words, if readers have mastery of the 1st 4,000 BNC/COCA frequency-based

word families or a vocabulary size of around 3,500–4,000 word families, they are supposed to achieve the aforementioned text coverage and have a minimal comprehension of learning that lexical threshold of high-frequency word families. Moreover, having optimal comprehension of the texts at 98.11% was reached at the 1st 9,000 BNC/COCA word family list, as readers are suggested to have the lexical threshold of 98% of text coverage, and it was implied that a vocabulary size of around 8,500–9,000 families is needed.

Analyzed in this study, the lexical thresholds required to read the target textbooks as compiled into four subcorpora of the whole corpus appeared to be moderately different. The third 1,000 BNC/COCA word family lists achieved 95.66% text coverage in the MIC subcorpus, while the 5th 1,000 BNC/COCA word family lists achieved 98.69% text coverage. The 3rd 1,000 BNC/COCA word family lists in the MAC subcorpus reached 95.39%, and the 6th 1,000 BNC/COCA word family lists reached 98.04%. The 5th 1,000 BNC/COCA word family lists in the CAL subcorpus attained 95.18%, but 98% text coverage was not achieved. The text coverage in the STA subcorpus was 96.10% for the 4th 1,000 BNC/COCA word family lists and 98.05% for the 11th 1,000 BNC/COCA word family lists.

5.1.3.2 Text Coverage of the Academic Word List in the Textbooks

The following books were chosen to represent the economics core course readings for the Bachelor of Economics International Program (BEIP) first-year students: (1) the textbook for the course on principles of microeconomics; (2) the textbook for the course on principles of macroeconomics; (3) the textbook for the course on calculus for social science 1; and (4) the textbook for the course on statistics for social science 1. Thus, a corpus of 1,343,493 words from the textbooks was compiled and processed in order to report the findings on (1) text coverage of the General Service List (GSL) and the Academic Word List (AWL) in each subcorpus and the entire corpus, and (2) cumulative text coverage of the GSL and AWL in each subcorpus and the entire corpus. This resulted in the investigation to answer the first research question about text coverage, in which the Academic Word List appears in the aforementioned first-year undergraduate economics textbooks of the BEIP curriculum.

The subcorpora were investigated based on the results of AWL text coverage analysis in each text subcorpus and the ECON corpus. The AWL

comprised 8.73% of text coverage in the MIC subcorpus, 9.53% in the MAC subcorpus, 6.44% in the CAL subcorpus, and 13.80% in the STA subcorpus. As a result, the AWL provided text coverage of 9.62% in the ECON corpus.

5.1.3.3 The Most Frequently Occurring AWL Academic Words in the Textbooks

One of the primary goals of this study was to identify the most commonly occurring Academic Word List (AWL) academic vocabulary in first-year BEIP textbooks. AntWordProfiler (Anthony, 2022) was used to identify the most common AWL academic words in the target textbooks that can be employed in academic environments, particularly written academic contexts. Coxhead's (2000) Academic Word List has 570 word families selected for their frequency, notably their high frequency in academic contexts. Furthermore, the words on the list do not appear on the General Service List, which covers 2,000 high-frequency English general words. According to Coxhead (2000), the threshold for frequency uniformity was 100 occurrences overall, with at least ten in each of the four areas of the 3.5-million-word corpus. Vongpumivitch et al. (2009) mentioned that a list of 475 AWL words appeared at least 50 times overall and five times in each subcorpus across the 1.5 million words in the corpus. Hajiyeva (2015) estimated that the word occurrences would appear 16 times overall and at least twice in each subcorpus in a corpus of 508,802 words.

The third research question in this study aimed to identify the most frequently used AWL academic words in the BEIP's first-year undergraduate textbooks. To do this, a list of words that match Coxhead's (2000) criteria for word selection must be prepared. The frequency of occurrences of AWL academic words in the ECON corpus should be at least 38 times and at least 4 times in each subcorpus. The most frequently used AWL academic words were as follows: data (2,978 occurrences), function (2,895 occurrences), equation (2,808 occurrences), chapter (2,398 occurrences), income (2,053 occurrences), series (1,575 occurrences), distribution (1,549 occurrences), area (1,306 occurrences), economy (1,232 occurrences), estimate (1,222 occurrences), labor (1,187 occurrences), section (1,063 occurrences), percent (981 occurrences), compute (965 occurrences), functions (905 occurrences), etc.

5.2 Discussion

Vocabulary is crucial for language abilities, as it enhances basic skills like reading, writing, listening, and speaking. Laufer and Sim (1985) suggested 65–70% should be understood as an absolute minimum for comprehension, but Laufer (1989), as described in Hu and Nation (2000), suggests 95% should be known for solid comprehension. Hu and Nation (2000) suggest that 98% of words should be known for unaided comprehension. According to the previous studies done by Laufer and Ravenhorst-Kalovski (2010) and Nation (2006), the minimal vocabulary necessary to read and comprehend the textbooks is 4,000–5,000 word families for 95% text coverage and 7,000–8,000 word families for 98% coverage. The results of this current study suggest that a vocabulary size of 3,500–4,000 word families in an economics textbook is needed for reasonable reading comprehension as the BNC/COCA word family lists account for 95% of the running words in the ECON corpus, and a vocabulary size of approximately 9,000 word families is needed for optimal reading comprehension as the lists account for 98% of the running words in the ECON corpus. Nevertheless, it does not claim that reasonable comprehension can occur even if readers have reached a particular lexical threshold (word band), which will automatically lead to good reading comprehension (Laufer & Ravenhorst-Kalovski, 2010).

Based on the results of this current study, when the text coverage was at 95%, each word band after reaching 95% was smaller. According to BNC/COCA word family lists in the ECON corpus in this current study, the most frequent 4,000 words constituted 96.10%, the most frequent 5,000 words constituted 97.14%, the most frequent 6,000 words constituted 97.52%, the most frequent 7,000 words constituted 97.72%, and the most frequent 8,000 words constituted 97.96% of text coverage. These figures illustrated the smaller increases in cumulative text coverage: 1.04%, 0.38%, 0.20%, and 0.24%, respectively, compared to the word bands before 95% text coverage was reached (the 1st 1,000-word band: 73.62%, the 2nd 1,000-word band: 11.88%, and the 3rd 1,000-word band: 8.77%). If vocabulary in a word band was less frequent, the text coverage was supposed to be smaller. Laufer and Ravenhorst-Kalovski's (2010) study found that even a small increase in text coverage leads to the highest improvement in reading scores. However, low-frequency words, which are among a text's key words,

may be essential for comprehension due to their critical role in academic papers or the fact that learners with a greater vocabulary have better automaticity of decoding. Students also benefit from learning high-frequency words, which are produced and processed more quickly than low-frequency words, to expand vocabulary coverage for beginning-level students and make them more familiar with these words (Nation, 2001). Consequently, selecting words based on their frequency should focus on balanced and effective teaching and learning.

Comparing the text coverage of the Academic Word List (AWL) in this current study with prior studies in similar or different disciplines is essential in order to check if the results are consistent with other previous studies or not. Text coverage of the AWL word families different studies was reported, for example, 8.6% and 17.4% in two specialized texts: anatomy and applied linguistics as found in Chung and Nation (2003), 10.07% in medical research articles investigated in Chen and Ge (2007), 10.6% in the academic text corpus of biology, physics, computer science, mechanical and electronic engineering, sociology, business studies, and applied linguistics investigated in Hyland and Tse (2007), 9.06% in research articles of agriculture in Martínez et al. (2009), 11.17% in applied linguistics research articles as found in Vongpumivitch et al. (2009), 10.46% in a financial services corpus investigated in Li and Qian (2010), 11.96% as found in Khani and Tazik (2013), and 9.96% in chemistry research articles in Valipouri and Nassaji (2013).

According to Coxhead (2011), the AWL is present in roughly 10% of academic texts across all academic fields. In the corpus of first-year undergraduate economics textbooks of the Bachelor of Economics International Program (BEIP) investigated in this current study, it was found that the AWL covered about 9.62% of the cumulative text coverage in the ECON corpus. According to Coxhead's (2000) AWL, the text coverage of the AWL in cross-discipline academic texts was approximately 10% of the running words in the texts. In other words, this current study showed lower text coverage than Coxhead's (2000) AWL. However, the vocabulary in the following economics core courses: 1) *Principles of Macroeconomics* with 8.73% AWL text coverage; 2) *Principles of Microeconomics* with 9.53% AWL text coverage; and 3) *Statistics for Social Science 1* with 13.80% AWL text coverage, was relatively relevant to Coxhead's (2000) AWL, representing a general academic vocabulary. On

the other hand, the course *Calculus for Social Science 1* yielded an AWL text coverage of 6.44%, which was lower than other economics courses, and this might possibly result from the assumption that there were more numbers and specialized figures than academic vocabulary in the text and that the course was in a more specific discipline than any other course.

5.3 Conclusions

Corpora provide not only authentic English language usage but also guidance in developing quality course materials. Vocabulary mastery is crucial for students to develop language skills, including listening, speaking, reading, and writing. The relationship between vocabulary mastery and language skill performance is clear, but its contribution varies across different skills. However, English has a large vocabulary, which may cause a challenge for students, and they require advice on how to learn new words and use them purposefully. It is noteworthy that non-native speakers have a smaller vocabulary than native speakers when choosing words to teach L2 vocabulary (Laufer & Nation, 2012). Research with undergraduates and graduates has repeatedly shown that ESL/EFL students lack the necessary vocabulary knowledge for reading, listening, speaking, and writing in English.

Whether the context is general (English for Academic Purposes, EAP) or specific (English for Specific Purposes, ESP), analyzing English word lists requires analyzing the teacher's circumstances. Given their particular objectives for studying and using English, it is also essential for teachers to take into account the interests and demands of the students.

The researcher offers the following conclusions in light of this research. This study was conducted to answer three research questions regarding the text coverage of BNC/COCA and AWL word families in a corpus of first-year undergraduate economics textbooks from an international program in economics, the estimated vocabulary size needed to read the textbooks, and the most frequently used AWL academic words in the textbooks. A corpus of 1,343,493 words was analyzed using the AntWordProfiler software. In the ECON corpus, the 4th 1000 BNC/COCA

word family lists reached 96.10%, and the 9th 1000 BNC/COCA word family lists reached 98.11%.

Text coverage is crucial for stakeholders who use and design materials, for example, curriculum developers, instructional materials developers, and teachers, since it may be used to calculate the proportion of known and unknown words in a text. Words can be selected to teach or include in an exercise according to their frequency. Apart from an explicit focus on high-frequency words, teachers should also concentrate on words students may struggle with autonomously, such as polysemous words with unusual meanings, as it enhances students' understanding of complex vocabulary (Todd, 2017).

The AWL, developed by Coxhead (2000), is a 570-word family with a 10% token frequency in academic texts. It has stable coverage in various disciplines, accounting for around 10% of tokens in various academic disciplines. The results showed that AWL accounted for 9.62% of text coverage, while the first and second 1000-word bands of the GSL provided a text coverage of 79.71%. Then, the most common AWL academic words were identified. The selected words should appear at least 38 times in the text corpus, and there were 187 AWL academic words, for example, data, function, equation, chapter, income, series, distribution, area, economy, estimate, labor, section, percent, compute, and functions. The 133 frequently used academic word family headwords found in the ECON corpus can serve as a guide for prospective curricula and instructional materials. To read and comprehend university textbooks, first-year undergraduate students in economics may benefit from this list of vocabulary through either individual learning or instruction in class.

5.4 Implications

According to Laufer and Nation (2012), vocabulary instruction has been mostly overlooked and delegated to other units of analysis in curriculum design, such as grammar, topics, or tasks. The results of the current study can be useful for curriculum developers, materials designers, subject content lecturers, and English language lecturers. They can use the results to redesign or modify instructional materials so that vocabulary will be presented in a meaningful way.

Instructional materials are important resources for a course, so careful word selection is crucial for effective learning experiences. Words can be classified in many ways, but words can be beneficial when they are appropriately selected for a particular pedagogical purpose, its context, and students' needs. The CEFR guidelines categorize words using reference-level descriptors. The CEFR, the Common European Framework of Reference, can be used at two levels: in designing curricula and textbooks and in applying the CEFR in the classroom. However, many professionals find it difficult to use without guidance (Cambridge, 2011). The CEFR should be seen as a flexible tool, serving as a framework for creating better materials to meet diverse needs of language learners. Text Inspector, an online lexical profiling tool, was informally used to predict the CEFR level of each word of the AWL headwords and to show a few measures of vocabulary use in this study. However, the tool was not performed with a full-length text. Regardless of word frequency, the frequently-used AWL academic word family headwords in this current study were analyzed by Text Inspector (Cambridge University Press, 2015) and then categorized into groups corresponding to the six reference levels of the CEFR, ranging from A1 to C2, which are alphabetically listed and shown in Table 5.1.

Table 5.1

CEFR Levels, Tokens, Percentage, and Each Word by Level of 133 Most Frequently Occurring AWL Word Family Headwords

CEFR Levels	Token Count/ Percentage	Words by Level
A1	1 (0.75%)	computer
A2	9 (6.77%)	area, available, final, goal, negative, normal, plus, project, team
B1	37 (27.82%)	achieve, annual, approach, chapter, comment, conclusion, create, cycle, design, energy, environment, individual, initial, instance, involve, item, label, locate, maximum, medium, method, minimum, percent, period, positive, predict, previous, publish, range, region, require, role, section, series, similar, sum, technology
B2	57 (42.86%)	affect, alternative, analyse, appropriate, approximate, assume, capacity, complex, concept, consequence, constant, construct, consume, correspond, data, decline, define,

CEFR Levels	Token Count/ Percentage	Words by Level
		distribute, economy, ensure, estimate, eventually, evidence, expand, factor, feature, focus, function, identical, identify, ignore, illustrate, impact, income, indicate, input, invest, maintain, major, obtain, occur, physical, portion, potential, primary, procedure, process, resource, shift, source, strategy, structure, substitute, summary, survey, task, theory
C1	21 (15.79%)	aid, assumption, conclude, consumption, derive, eliminate, equation, equivalent, evaluate, exceed, hence, insight, labour, nonetheless, principle, proceed, ratio, restrict, revenue, sector, somewhat
C2	4 (3.01%)	consistent, hypothesis, margin, precede
Unlisted	4 (3.01%)	appendix, consist, coordinate, maximise
Total	133 (100.00%)	-

Table 5.1 shows the most frequently occurring AWL word family headwords, which are classified by each level of the CEFR as follows: 1 word (0.75%) at the A1 level, 9 words (6.77%) at the A2 level, 37 words (27.82%) at the B1 level, 57 words (42.86%) at the B2 level, 21 words (15.79%) at the C1 level, 4 words (3.01%) at the C2 level, and 4 words (3.01%) that are not in the list. The CEFR outlines six levels of language proficiency, ranging from A1 (beginner) to C2 (most advanced), catering to individuals with limited language experience, intermediate speakers, and proficient speakers.

Curriculum developers and materials designers need to support individuals with different needs. First-year economics students with limited proficiency in English or those who struggle in learning may take time and need extra attention and support from economics textbooks by adding extra pinpoints for them, for example, introducing words with their definitions both in a general meaning and in an economics context, vocabulary learning strategies with guidance in learning new vocabulary in a textbook, giving multiple opportunities for vocabulary instruction to enhance students' vocabulary practice and memory of important words, expanding vocabulary through semantic maps by visually connecting words and related concepts, and so on.

Alternatively, learners can learn new vocabulary by learning a word family, including a base word and its derived and inflected forms. They require knowledge of word bases to understand relationships between words in a word family and recognize

words that are related by sharing a common base (Bauer & Nation, 1993). Teaching word families involves incorporating inflected and derived forms, which can significantly impact learning. A few inflected forms can be more effective than many, making it easier to teach and learn various word families. The headwords in this study, function (2,895 occurrences) and economy (1,232 occurrences), should be suggested in a first-year economics textbook. For example, the root words of 1) function and 2) economy should be introduced first, followed by briefly mentioning the possibilities of 1) functional, functionalism, functionalist, functionalists, functionalities, functionality, functionally, functioned, functioning, and functions, and 2) economic, economical, economically, economics, economies, economise, economised, economises, economising, economist, economists, economize, economized, economizes, economizing, uneconomic, uneconomical, and uneconomically, respectively. This results in a less time-consuming method. In addition, students are suggested to apply rules in word formation so that their vocabulary knowledge is broadened, especially in derived words (Yang & Dai, 2011).

Word memorization, or rote memorization, is another form of explicit vocabulary instruction that students are familiar with. They found it beneficial and effective because it provides students with easy-to-understand definitions, examples, discussion, and practice opportunities to build their vocabulary knowledge. However, "the deliberate learning of vocabulary may contribute directly to implicit knowledge if the words learned are not complicated and if the learning is meaningful" (Nation, 2002, p. 271).

Word lists have been developed to help identify the most useful vocabulary required by ESL and EFL learners, and they serve as a crucial tool in ESL and EFL teaching materials. A word list can be general or specific, such as the General Service List (GSL), developed by West (1953), which is a 2,000-word list of English words considering word frequency and general use in written English, arranged alphabetically. Specific frequency lists include the most common words in different genres or fields. These words can be classified into high-frequency, mid-frequency, and low-frequency words. According to Laufer and Nation (2012), frequent words are useful for all purposes, while infrequent words are useful for specific needs, such as academic reading or tour guiding. Stakeholders have to ensure that vocabulary is accurate and

useful for effective language learning. Consequently, word frequency is one factor that should undoubtedly be taken into consideration when selecting words for teaching and learning. High-frequency words should be prioritized in subject content for an economics core course to improve text comprehension, as they contribute to reading comprehension. These lists are crucial for reaching specific word bands in a text in order to have minimal or adequate reading comprehension. English teachers are advised to adopt a variety of approaches to teaching students vocabulary in order to help them recall new vocabulary words. Teachers commonly use a variety of strategies to communicate a particular vocabulary word, ensuring that students retain the knowledge. To keep students from forgetting, teachers require students to revisit and practice vocabulary on an ongoing basis.

Curriculum developers and instructional materials developers in Economics or Economics-related disciplines can incorporate these words into exercises. The selection of words in the BNC/COCA word families and AWL academic word family headwords can be used as a guideline when designing a course syllabus or curriculum and integrating vocabulary activities, practices, and tests in a lesson to ensure that the percentage of known words (text coverage) that students know is sufficient for them to comprehend an economics text. On the assumption that it is not necessary to comprehend or know every single word in a text for text comprehension, it is essential to know the text coverage, or percentage of known and unknown words in a text.

In a preparation course for economics undergraduate students, English language lecturers who help with preparing teaching materials may benefit from a research-based economics word list to make sure that students comprehend the topics and have the necessary content knowledge for the target words they must know in the BEIP first-year undergraduate textbooks. Regarding a range of objectives in vocabulary learning and teaching, word frequency can be used to design syllabuses and materials so that words can be learned purposefully. English language lecturers, as they teach and give lectures on the language, can help students gain knowledge of the target words in the first-year undergraduate textbooks. If they have some questions, they may ask the content lecturers for help.

The most frequently-used AWL academic words in this current study are intended to provide teachers, instructors, and course designers with academic words that are ready to use in the first-year undergraduate textbooks for the program's core economics courses. To guarantee that the chosen words are the most relevant and helpful for the students of economics, teachers may focus their efforts on subject-specific vocabulary in a meaningful way rather than instinctively adding vocabulary for their students throughout class.

This can be done by using vocabulary teaching techniques such as paper or electronic flashcards, vocabulary games, and extensive reading programs. Regarding the expected return on time and effort invested for students who have limited vocabulary, flashcard learning is an efficient and flexible method of direct (deliberate) vocabulary teaching and learning. Flashcards might be plain or come in a variety of colors. They also serve as a visual resource. They can be physically written (hard copies) or electronically stored (soft copies). Both teachers and students can purposefully concentrate on words that are highly significant yet difficult to learn through context or dictionary use (Nation, 2001; Nation & Meara, 2002, as cited in Chanasattru & Tangkiengsirisin, 2016). The majority of students find flashcards motivating since the images are captivating and the language is simple or easy to understand. In this current study, headwords or word families should be incorporated into learning activities by introducing them to students on one side of flashcards, while the other side can have word meanings.

It may be beneficial for undergraduate students who want to prepare themselves for learning the vocabulary in an economics course in their first year of study. However, students will find it challenging to read or comprehend a text if they are unfamiliar with difficult or unknown words. In other words, students with a strong vocabulary have better reading comprehension. To make it easier for language learners, lists of words appearing most frequently in general text have been created. Mastering the core vocabulary in the specific word lists of economics covers the majority of words in a text, enabling students to comprehend and use the majority of the text effectively. Since they keep students relaxed and foster an engaging environment, flashcards are an alternative tool for teaching English. Some words are more beneficial to some groups of language users than others. In this study, first-year undergraduate students must take

specialized and academic courses that they are not familiar with, and time and attention are required. Vocabulary learning is consequently essential for overall English proficiency development and vocabulary development itself. Vocabulary items in English are not equally relevant to learners with different proficiency levels and vocabulary knowledge, so reading BEIP textbooks can be daunting. According to West's (1953) General Service List, there were significant factors in selecting the word, i.e., frequency, ease of learning, the usefulness of the concept, and the register of the word. The lists were undertaken based on well-considered criteria.

Further research on different skill levels of learners considering different aspects is needed to achieve a more thorough understanding of different text types and vocabulary words.

5.5 Recommendations for Further Research

This study examined first-year economics undergraduate textbooks for core courses of an international program in Thailand, providing researchers, teachers, and related economics major stakeholders with frequency data that may be used to improve vocabulary input in class or in a preparation course for first-year students who need more practice and struggle with learning vocabulary.

Designing a corpus for research purposes requires considering factors like size, balance, and representativeness. In addition, being too small or too large in size can impact representativeness. This current study compiled a corpus from four economics courses, resulting in a 1,343,493-word corpus. However, the size of the text corpus used is large enough to be a good representative in order to answer the research questions.

Since the emphasis of this study has been on word families as the counting unit and the focus is on single-word units, it did not analyze multi-word units from the BEIP's first-year undergraduate textbooks. The concept of word families might help readers to comprehend how these words are related and to understand more about the words and their meanings. As a result, recognizing word family patterns is critical to comprehending the complexity of English. Further research could analyze multi-word

units or other multi-word expressions found in textbooks, such as lexical bundles, idiomatic expressions, and so on.

Further research of this nature should be conducted with economics textbooks in other years of study, i.e., the second year, the third year, and the fourth year. Regardless of the microeconomics and macroeconomics textbooks used in the current study, researchers who need to identify some gaps for new knowledge in a particular area of interest can investigate textbooks in other fields that are related to economics, such as labor economics, financial economics, econometrics, international economics and development economics, health economics, public and institutional economics, and other social sciences, such as sociology and politics. In addition, further studies in other different pedagogical and non-pedagogical genres are also suggested for investigation, as follows: journal articles, research abstracts, research proposals, conference papers, theses, and essays.

REFERENCES

Afshari, S., & Tavakoli, M. (2017). The relationship between depth and breadth of vocabulary knowledge and Iranian EFL learners' listening comprehension. *International Journal of Research Studies in Language Learning, 6*(3), 13–24.

Aizawa, I., & Rose, H. (2020). High school to university transitional challenges in English-medium instruction in Japan. *System, 95.*

Ali, I. R. I. (2010). The correlation between students' vocabulary mastery and reading comprehension. *(Unpublished bachelor's thesis). Universitas Islam Negeri Syarif Hidayatullah.*

Al-Jarf, R. (2007). Teaching Vocabulary to EFL College Students Online. *Call-EJ Online, 8*(2).

Almutairi, N. D. (2016). The effectiveness of corpus-based approach to language description in creating corpus-based exercises to teach writing personal statements. *English Language Teaching, 9*(7), 103–111. http://dx.doi.org/10.5539/elt.v9n7p103

Altalhab, S. (2019). The vocabulary knowledge of Saudi EFL tertiary students. *English Language Teaching, 12*(5), 55–65.

Anthony, L. (2022). AntFileConverter (Version 2.0.2) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Anthony, L. (2022). AntWordProfiler (Version 2.0.1) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

Baker, M. (1988). Sub-technical vocabulary and the ESP teacher: An analysis of some rhetorical items in medical journal articles. *Reading in a Foreign Language, 4*(2), 91–105.

Bauer, L., & Nation, I. S. P. (1993). Word families. *International Journal of Lexicography, 6*(4), 253–279.

Benedict, M. C., & Shabdin, A. A. (2021). Word frequency level and lexical coverage in the reading comprehension texts of the Malaysian university English test. *PASAA: Journal of Language Teaching and Learning in Thailand, 61*, 33–60.

Besthia, W. (2018). A survey on vocabulary learning strategies: A case of Indonesian EFL university students. *IOSR Journal of Research & Method in Education, 8*(5), 29–34.

Biber, D., Conrad, S., and Reppen, R. (1998). *Corpus Linguistics: Investigating language structure and use.* Cambridge: Cambridge University Press.

Briggs, J. G., Dearden, J., & Macaro, E. (2018). English medium instruction: Comparing teacher beliefs in secondary and tertiary education. *Studies in Second Language Learning and Teaching, 8*(3), 673–696.

Boonkit, K. (2010). Enhancing the development of speaking skills for non-native speakers of English. *Procedia - Social and Behavioral Sciences, 2*(2), 1305–1309. https://doi.org/10.1016/j.sbspro.2010.03.191

Boonnoon, S. (2019). Vocabulary learning strategies employed by Thai university students across four academic profiles. *Theory and Practice in Language Studies, 9*(8), 902–910. http://dx.doi.org/10.17507/tpls.0908.02

Boonpattanaporn, P. (2017). Grammar and communicative English language teaching in the Thai educational context. *University of the Thai Chamber of Commerce Journal, 37*(2), 152–169.

Bosuwan, T. (2013). Common syntactic errors in interlanguage writing output: A case study of EFL undergraduates. University of the Thai Chamber of Commerce Journal, 33(special issue), 229–246.

Bradford, A., & Brown, H. (2016). *CBI, CLIL & EMI Differing Approaches and Goals* (poster). https://doi.org/10.13140/RG.2.2.19525.04325

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics, 36*(1), 1–22.

Bukve, T. (2020). Students' perspectives on English medium instruction within higher education – exploring gender differences in Norway and Finland. *Apples – Journal of Applied Language Studies, 14*(1), 7–24.

Bunwirat, N. (2017). English language teaching in AEC era: A case study of universities in the upper northern region of Thailand. *FEU Academic Review, 11*(2), 282–293.

Burapha University. (n.d.). *International Admission Burapha University.* Retrieved November 9, 2020, from http://admission.buu.ac.th/index.php/MTJ8fHByb2dyYW0vaW5kZXg

Cambridge, E. S. O. L. (2011). Using the CEFR: Principles of good practice. *Cambridge ESOL*, 1–48.

Cambridge University Press. (2015). Text Inspector. [Web Tool]. https://languageresearch.cambridge.org/wordlists/text-inspector

Chai-ngam, R., Waikham K., & Chai-ngam, H. (2018). Problems and Solutions to Improve Students' English Skills of the Professional Learning Community of English Teachers of Khon Kaen Provincial Administrative Organization. *Journal of Education, Mahasarakham University, 12*(2), 7.

Chalapati, S. (2007). The internationalisation of higher education in Thailand: Case studies of two English-medium business graduate programs (Doctoral dissertation, RMIT University, Australia). Retrieved from https://www.researchgate.net/publication/228396117_The_internationalisation_of_higher_education_in_Thailand_case_studies_of_two_English-medium_business_graduate_programs

Chanakan, T., & Patanasorn, A. T. (2018). Errors in Adjective-Noun Order by Thai and Chinese EFL Learners: Roles of L1 and Language Proficiency. School of Liberal Arts, King Mongkut's University of Technology Thonburi.

Chanchanglek, S., & Sriussadaporn, N. (2011). A corpus analysis of English vocabulary input in course materials used for Engineering students. *Humanities Journal, 18*(1), 141–149.

Chatpunnarangsee, K. (2015). *Corpora, Concordancing and Collocations*. Thammasat University.

Chatsawat, T. (2013). English proficiency development as a preparation for the ASEAN Economic Community (AEC) 2015: a study of students at Thammasat University. Language Institute, Thammasat University.

Chawwang, N. (2008). *An Investigation of English Reading Problems of Thai 12th Grade Students in Nakhonratchasima Educational Regions 1, 2, 3, and 7* (Master's thesis). Srinakharinwirot University, Bangkok, Thailand.

Chen, H., Han, J., & Wright, D. (2020). An investigation of lecturers' teaching through English medium of instruction—A case of higher education in China. *Sustainability, 12*(10), 4046.

Chen, Q., & Ge, G-C. (2007). A corpus-based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs). *English for Specific Purposes, 26*, 502–514. DOI: 10.1016/j.esp.2007.04.003

Chiang Mai University. (n.d.). *International Program Admission System.* Retrieved November 9, 2020, from http://www1.reg.cmu.ac.th/reg-ipas/main/index.php

Chuang, Y. T. (2015). An EMI pedagogy that facilitates students' learning. *English Language Teaching, 8*(12), 63–73.

Chuenjit, T. (2011). *The Influence of Vocabulary Knowledge on the English Listening Comprehension of Thai Hospital Pharmacists* (Master's thesis). Thammasat University, Bangkok, Thailand.

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

Chulalongkorn University. (n.d.). *International Programs.* Retrieved November 9, 2020, from https://www.chula.ac.th/en/admissions/international-programs/

Cobb, T. (2010). Learning about language and learners from computer programs. *Reading in a Foreign Language, 22*(1), 181–200.

Cobb, T. (2020). *Familizer + Lemmatizer v.2.2* [computer program]. Available at: https://lextutor.ca/familizer.

Coffin, P. (2014). Challenges and needs of professional development for lecturers of English-medium instruction. *MFU Connexion, 3*(1), 42–53.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238.

Coxhead, A. (2011). The academic word list 10 years on: Research and teaching implications. *TESOL Quarterly, 45*(2), 355–362.

Coxhead, A., & Demecheleer, M. (2018). Investigating the technical vocabulary of plumbing. *English for Specific Purposes, 51*, 84–97.

Coxhead, A., & Hirsh, D. (2007). A pilot science-specific word list. *Revue Française de Linguistique Appliquée, 12*(2), 65–78.

Coxhead, A., & Nation, P. (2001). The Specialized Vocabulary of English for Academic Purposes. In J. Flowerdew, & M. Peacock (Eds.), Research Perspectives on English for Academic Purposes, Chapter: The Specialised Vocabulary of English for Academic Purposes (pp. 252–267). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139524766.020

Crawford, W., and Csomay, E. (2015). *Doing corpus linguistics*. Routledge.

Cushing, S. T. (2017). Corpus linguistics in language testing research. *Language Testing*, 34(4), 441–449.

Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The academic spoken word list. *Language Learning, 67*(4), 959–997.

Dimova, S. (2020). English language requirements for enrolment in EMI programs in higher education: A European case. *Journal of English for Academic Purposes, 47*, 1–13. https://doi.org/10.1016/j.jeap.2020.100896

Eak-in, S. (2015). Effects of a Corpus-Based Instructional Method on Students' Learning of Abstract Writing: A Case Study of an EAP Course for Engineering Students. Thammasat University.

Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin Press.

Ekoç, A. (2020). English Medium Instruction (EMI) from the perspectives of students at a technical university in Turkey. *Journal of Further and Higher Education, 44*(2), 231–243.

Fan, F. (2013). Text length, vocabulary size and text coverage constancy. *Journal of Quantitative Linguistics, 20*(4), 288–300. https://doi.org/10.1080/09296174.2013.830550

Firdaus, M. S. R. (2019). The correlation between students' vocabulary knowledge and the writing skill. *Language-Edu, 8*(1).

Foley, J. A. (2019). Adapting CEFR for English language education in ASEAN, Japan, and China. *The New English Teacher, 13*(2), 101–117.

Folse, K. (2010). Is explicit vocabulary focus the reading teacher's job? *Reading in a Foreign Language, 22*, 139–160.

Forti, L. (2017). Data-Driven Learning and the Acquisition of Italian Collocations: From Design to Student Evaluation. *Research-Publishing.Net*, 2017.

Galloway, N., & Ruegg, R. (2020). The provision of student support on English medium instruction programmes in Japan and China. *Journal of English for Academic Purposes, 45.* https://doi.org/10.1016/j.jeap.2020.100846

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics, 35*(3), 305–327.

Goodman, B. A. (2014). Implementing English as a medium of instruction at a Ukrainian University: Challenges, adjustments, and opportunities. *International Journal of Pedagogies and Learning, 9*(2), 130–141. https://doi.org/10.1080/18334105.2014.11082026

Hajiyeva, K. (2015). A corpus-based lexical analysis of subject-specific university textbooks for English majors. *Ampersand, 2*, 136–144.

Hakim, L. (2016). A Study on the Problems Faced by Thai Students in Learning English Speaking at Boarding House.

Hengsadeekul, C., Koul, R., & Kaewkuekool, S. (2014). Motivational orientation and preference for English-medium programs in Thailand. *International Journal of Educational Research, 66*, 35–44.

Hiranburana, K. (2017). Use of English in the Thai workplace. *Kasetsart Journal of Social Sciences, 38*(1), 31–38.

Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language, 8*(2), 689–696.

Hsu, W. (2013). Bridging the vocabulary gap for EFL medical undergraduates: The establishment of a medical word list. *Language Teaching Research, 17*(4), 454–484.

Hsu, W. (2014). Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes, 33*(1), 54–65.

Hu, G., Li, L., & Lei, J. (2013). English-medium instruction at a Chinese university: Rhetoric and reality. *Language Policy, 13*(1), 21–40.

Hu, J., & Wu, P. (2020). Understanding English language learning in tertiary English-medium instruction contexts in China. *System, 93*, 1–11. https://doi.org/10.1016/j.system.2020.102305

Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403–430.

Hua, T-L. (2019). Understanding the learning challenges of English-medium instruction learners and ways to facilitate their learning: A case study of Taiwan psychology students' perspective. *Latin American Journal of Content and Language Integrated Learning, 12*(2), 321–340.

Huang, T. L. (1997). On the importance of emphasizing vocabulary teaching. In *Proceedings of the 6th International Symposium on English Teaching*, 322–331, Taipei: Crane Publishing.

Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary"? *TESOL Quarterly, 41*(2), 235–253.

Ibrahim, E. H. E., Sarudin, I., & Muhamad, A. J. (2016). The relationship between vocabulary size and reading comprehension of ESL learners. *English Language Teaching, 9*(2), 116–123.

Intarasombat, P. (2002). *The Effect of Vocabulary Development Approach on Mathayomsuksa 4 Students' English Reading Comprehension* (Master's thesis). Khon Kaen University, Thailand.

It-ngam, T. & Phoocharoensil, S. (2019). The development of science academic list. *Indonesian Journal of Applied Linguistics, 8*, 657–667. doi: 10.17509/ijal.v8i3.15269

Jaiprasertsook, A. (2014). The survey study of English language awareness in high school students towards AEC: A case study of Benjamarachutit Ratchaburi School. Language Institute, Thammasat University.

Jiang, L., Zhang, L. J., May, S. (2019). Implementing English-medium instruction (EMI) in China: Teachers' practices and perceptions, and students' learning motivation and needs. *International Journal of Bilingual Education and Bilingualism, 22*(2), 107–119.

Jindapitak, N. (2019). English as an ASEAN lingua franca and the role of nativeness in English education in Thailand: Moving toward the ASEAN Economic Community (AEC). *English Today, 35*(2), 36.

Jindathai, S. (2013). A case study of problems in learning English as a Foreign Language among university students: Thai-Nichi Institute of Technology. *Journal of Business and Languages, 1*(1), 33–36.

Joe, Y., & Lee, H.-K. (2013). Does English-medium instruction benefit student in EFL context? A case study of medical students in Korea. *The Asia-Pacific Education Researcher, 22*(2), 201–207.

Johns, T., King, P. (1991). *Classroom concordancing.* Centre for English Language Studies, the University of Birmingham, pp. 1–16.

Jones, C., & Waller, D. (2015). *Corpus linguistics for grammar: A guide for research*. Routledge.

Kaewphanngam, C., & Tharmmapornphilas, R. (2013). Education policy and human resource development in Thailand under creative economy. *NIDA Economic Review, 7*(1), 251–306.

Karademir, C. A., & Gorgoz, S. (2019). English Teachers' Problems Encountered in Teaching Four Basic Language Skills. *International Education Studies, 12*(4), 118–127.

Kasetsart University. (n.d.). *Undergraduate Degree Programs.* Retrieved November 10, 2020, from http://www.interprogram.ku.ac.th/newsite/index.php?option=com_content&view=article&id=2&Itemid=35

Kaur, A., & Young, D. (2016). English education policy in Thailand: Why the poor results? In R. Kirkpatrick (Ed.), *English language education policy in Asia* (pp. 345–361). Springer International Publishing. 10.1007/978-3-319-22464-0_16

Kennedy, G. (2014). *An introduction to corpus linguistics.* Routledge.

Khan, R. M. I., Razuan, N. R. M., Shahbaz, M., Ibrahim, A. H., & Mustafa, G. (2018). The role of vocabulary knowledge in speaking development of Saudi EFL learners. *Arab World English Journal, 9*(1), 406–418.

Khon Kaen University. (n.d.). *International Programs.* Retrieved November 10, 2020, from https://www.kku.ac.th/academics/international-programs

Kilic, M. (2019). Vocabulary knowledge as a predictor of performance in writing and speaking: A case of Turkish EFL learners. *PASAA, 57*, 133–164.

Kim, E. G., Kweon, S.-O., & Kim, J. (2017). Korean engineering students' perceptions of English-medium instruction (EMI) and L1 use in EMI classes. *Journal of Multilingual and Multicultural Development, 38*(2), 130–145.

King Mongkut's Institute of Technology Ladkrabang. (n.d.) *Laksut Lae Phaen Kansueksa* [Curriculum and Study Plan]. Retrieved November 10, 2020, from http://www.reg.kmitl.ac.th/curriculum/index.php

King Mongkut's Institute of Technology North Bangkok. (n.d.) *International Program.* Retrieved November 10, 2020, from https://kmutnb.ac.th/academics/program_inter.aspx

King Mongkut's Institute of Technology Thonburi. (n.d.). *Laksut Radap Parinyatri* [Bachelor Degree Programs]. Retrieved November 10, 2020, from https://admission.kmutt.ac.th/bachelor

Kirkpatrick, A. (2010). *English as a lingua franca in ASEAN: a multilingual model*. Hong Kong University Press.

Kongthong, J. (2007). A survey study of L2 vocabulary learning strategies of 4th year students, the Faculty of Fine Arts, at Rajamangala University of Technology Tawan-ok Chakrabongse Bhuvanarth Campus (Master's thesis). Thammasat University, Bangkok, Thailand.

Lam, R. (2016). Assessment as learning: examining a cycle of teaching, learning, and assessment of writing in the portfolio-based classroom. *Studies In Higher Education*, *41*(11), 1900–1917. doi:10.1080/03075079.2014.999317

Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren, & M. Nordman (Eds.), *Special language: From human thinking to thinking machines* (pp. 316–323) Clevedon: Multilingual Matters.

Laufer, B. & Nation, I. S. P. (2012). Vocabulary. In S. M. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 163–176). London, England: Routledge.

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.

Laufer, B. (2014). Vocabulary in a second language: Selection, acquisition, and testing: A commentary on four studies for JALT vocabulary SIG. *Vocabulary Learning and Instruction, 3*(2), 38–46.

Lei, L., & Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes, 22*, 42–53.

Lei, Z. (2018). Effectiveness of college student English writing teaching based on data-driven learning. *International Journal of Emerging Technologies in Learning*, *13*(4), 106. doi:10.3991/ijet.v13i04.8474

Leki, I., & Carson, J. G. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL Quarterly, 28*, 81–101.

Lessard-Clouston, M. (2013). Word lists for vocabulary learning and teaching. *CATESOL Journal*, *24*(1), 287–304.

Li, Y., & Qian, D. D. (2010). Profiling the Academic Word List (AWL) in a financial corpus. *System, 38*(3), 402–411.

Li, Y., & Zhang, X. (2019). L2 vocabulary knowledge and L2 listening comprehension: A structural equation model. *Canadian Journal of Applied Linguistics, 22*(1), 85–102.

Liangpanit, C. (2010). *The development of a corpus-based vocabulary package for Business English majors* (Doctoral dissertation, Suranaree University of Technology, Nakhon Ratchasima, Thailand). Retrieved from http://sutir.sut.ac.th:8080/jspui/bitstream/123456789/3469/1/Fulltext.pdf

Liu, D., & Ping, J. (2009). Using a Corpus-Based Lexicogrammatical Approach to Grammar Instruction in EFL and ESL Contexts. *Modern Language Journal, 93*(1), 61–78.

Luo, Q. (2016). The effects of data-driven learning activities on EFL learners' writing development. *Springerplus, 5*(1), 1–13. doi:10.1186/s40064-016-2935-5

Macaro, E., & Han, S. (2020). English medium instruction in China's higher education: Teachers' perspectives of competencies, certification, and professional development. *Journal of Multilingual and Multicultural Development, 41*(3), 219–231.

Mala, D. (2018, November 5). Thai English proficiency drops. *Bangkok Post.* https://www.bangkokpost.com/thailand/general/1570042/thai-english-proficiency-drops

Mayer, R. E. (2008). *Learning and Instruction.* Upper Saddle River, New Jersey: Pearson Education, Inc.

Mae Fah Luang University. (n.d.). *Bachelor's Degree.* Retrieved November 12, 2020, from https://en.mfu.ac.th/admission/programmes/bachelors-degree.html

Maejo University. (n.d.). *Laksut Thi Poet Son* [Programs]. Retrieved November 10, 2020, from http://www.education.mju.ac.th/www/programInfo.aspx

Mahachulalongkornrajavidyalaya University. (n.d.). *Undergraduate Programs.* Retrieved November 12, 2020, from https://www.mcu.ac.th/pages/Undergradeate

Mahamakut Buddhist University. (n.d.). *Samnakngan Khanakammakan Kan Udomsueksa Dai Rapsap Kan Anumat Laksut Mahawitthayalai Maha Makut Rat Witthayalai* [The Office of the Higher Education Commission Has Verified the Programs of Mahamakut Buddhist University]. Retrieved November 12, 2020, from http://mbu.ac.th/index.php/2012-11-14-02-12-21/2018-03-07-06-40-52#.X60HJmgzbIV

Mahidol University. (n.d.). *Introduction International Program for Undergraduate Level.* Retrieved November 12, 2020, from https://op.mahidol.ac.th/ea/program_profile-en/

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. Taylor & Francis.

Mehring, J. (2006). Developing vocabulary in second language acquisition: From theories to the classroom. *Hawaii Pacific University TESOL Working Papers Series, 4*(2), 3–9.

Minshall, D. (2013). *A computer science word list.* Unpublished master's thesis. Swansea: Swansea University.

Moon, Y. (2017). Are different types of vocabulary needed for comprehending general and academic texts? *English Teaching, 72*(3), 47–68.

Mungkonwong, P., & Wudthayagorn, J. (2017). An investigation of vocabulary size of Thai freshmen and its relationship to years of English study. *LEARN Journal: Language Education and Acquisition Research Network, 10*(2), 1–9.

Muensorn, A., & Tepsuriwong, S. (2009). *Vocabulary Learning Strategies by Thai Primary School Students.* School of Liberal Arts, King Mongkut's University of Technology Thonburi.

Nation, I. S. P. (1993). Vocabulary size, growth, and use. *The bilingual lexicon, 6,* 115–134.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge University Press.

Nation, I. S. P. (2002). Best practice in vocabulary teaching and learning. In J. C. Richards, & W. A. Renandya, (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 267–272). Cambridge University Press.

Nation, I. S. P. (2005). Teaching Vocabulary. *Asian EFL Journal, 7*(3): 47–54.

Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *The Canadian Modern Language Review, 63*(1), 59–82.

Nation, I. S. P. (2008). *Teaching Vocabulary: Strategies and Techniques.* MA: Heinle, Cengage Learning.

Nation, I. S. P. (2015). Principles guiding vocabulary learning through extensive reading. *Reading in a Foreign Language*, 27(1), 136–145.

Nation, I. S. P. (2016). *Making and Using Word Lists for Language Learning and Testing*. Amsterdam: John Benjamins.

Nation, I. S. P. (2017). The BNC/COCA Level 6 word family lists (Version 1.0.0) [Data file]. Available from http://www.victoria.ac.nz/lals/staff/paul-nation.aspx

Nation, I. S. P., & Anthony, L. (2013). Mid-frequency readers. *Journal of Extensive Reading, 1,* 5–16.

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

Nation, I. S. P., & Webb, S. A. (2011). *Researching and analyzing vocabulary*. Boston, MA: Heinle, Cengage Learning.

Nirattisai, S., & Chiramanee, T. (2014). The contribution of vocabulary learning strategies to university students' vocabulary size. *Silpakorn University Journal of Social Sciences, Humanities, and Arts, 14*(2), 45–67.

Noom-ura, S. (2013). English-teaching problems in Thailand and Thai teachers' professional development needs. *English Language Teaching, 6*(11), 139–147. doi:10.5539/elt.v6n11p139

Neufeld, S., Hancioglu, N., & Eldridge, J. (2011). Beware the range in RANGE, and the academic in AWL. *System, 39*(4), 533–538.

Nuttawat, P. (2008). A survey study of needs and problems in learning to speak English of the third year Tourism Industry students at Suratthani Rajabhat University. Language Institute, Thammasat University.

O'Keeffe, A., McCarthy, M. J. and Carter, R. A. (2007). *From Corpus to Classroom: Language Use and Language Teaching.* Cambridge: Cambridge University Press.

Oya, T., Manalo, E., & Greenwood, J. (2009). The influence of language contact and vocabulary knowledge on the speaking performance of Japanese students of English. *The Open Applied Linguistics Journal, 2*, 11–21.

Padermprach, N. (2017). The problems of English Language Education at the upper secondary level in Thailand: the perceptions of Thai EFL teachers and students in Bangkok.

Panduangkaew, R. (2018). An analysis of vocabulary learning strategies employed by Thai EFL undergraduates: Dictionary use. *rEFLections, 25*(1), 116–125.

Park, K. (2014). Corpora and Language Assessment: The State of the Art. *Language Assessment Quarterly, 11*(1), 27–44. doi:10.1080/15434303.2013.872647

Pengnate, W. (2014). The Determinant factors of Engineering Students towards the Significant of English Proficiency in the ASEAN Economic Community

(AEC): A Case of 3rd year TNI Students. *TNI Journal of Business Administration and Languages, 2*(1), 50–53.

Phoocharoensil, S. (2013). Cross-Linguistic Influence: Its Impact on L2 English Collocation Production. English Language Teaching, 6(1), 1–10.

Phoocharoensil, S. (2014). Exploring Learners' Developing L2 Collocational Competence. *Theory & Practice in Language Studies, 4*(12), 25–33.

Pookcharoen, S. (2007). Exploring How Teaching Morphemic Analysis Foster Thai EFL Students' Vocabulary Knowledge. *Journal of English Studies, 3*, 85–105.

Pookcharoen, S. (2009). *Metacognitive online reading strategies among Thai EFL university students* (Doctoral dissertation). Indiana University. Retrieved from ProQuest LLC. (UMI No. 3390322)

Pookcharoen, S. (2011). Thai EFL Undergraduates' Vocabulary Learning Strategies: Perception and Practice. *Journal of Studies in the English Language, 6*, 31–53.

Pookcharoen, S. (2016). Thai EFL university teachers' beliefs and practices about vocabulary learning strategies. *LEARN Journal: Language Education and Acquisition Research Network, 9*(2), 155–172.

Prapphal, K. (2003). English Proficiency of Thai Learners and Directions of English Teaching and Learning in Thailand. *Journal of English Studies, 1*(1).

Prince of Songkla University. (n.d.). *International and English Programs at Prince of Songkla University.* Retrieved November 13, 2020, from https://en.psu.ac.th/degrees-and-admission/international-and-english-programs

Rajaraman, A., & Ullman, J. D. (2011). *Mining of massive datasets.* Cambridge: Cambridge University Press.

Richards, J. C. (1976). The Role of Vocabulary Teaching. *TESOL Quarterly, 10*(1), 77–89. https://doi.org/10.2307/3585941

Richard, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching.* Cambridge: Cambridge University Press.

Roche, T. & Harrington, M. (2013). Recognition vocabulary skill as a predictor of academic English performance and academic achievement in English. *Language Testing in Asia, 3*(12), 133–144. http://dx.doi.org/10.1186/2229-0443-3-12

Rodngam, S. (2011). *The use of authentic materials to enhance vocabulary for listening comprehension of Mathayomsuksa 6 students in the English and Japanese Program at Winitsuksa School, Lopburi Province, Thailand* (Master's thesis). Thammasat University, Bangkok, Thailand.

Rohmatillah, R. (2014). A Study on Students' Difficulties in Learning Vocabulary. *English Education: Jurnal Tadris Bahasa Inggris*, *6*(1), 75–93.

Rojananak, K, & Vitayapirak, J. (2015). Comparison of English vocabulary learning strategies for good and weak learners at King Mongkut's Institute of Technology. *International Journal of Language, Literature and Linguistics, 1*(1), 1–5.

Saengpakdeejit, R. (2014). Strategies for Dealing with Vocabulary Learning Problems by Thai University Students. *Humanities, Arts and Social Sciences Studies, 14*(1), 147–168.

Sahatsathatsana, S. (2017). Pronunciation Problems of Thai Students Learning English Phonetics: A Case Study at Kalasin University. *Journal of Education, Mahasarakham University*, (4), 67.

Sanonguthai, S. (2014). Ready or not? The state of Thai schools in response to the ASEAN English language policy. *Thammasat Review, 16*(2), 128–142.

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

Schmitt, N. (2008). Review article: Instructed second language vocabulary learning. *Language Teaching Research, 12*(3), 329–363. https://doi.org/10.1177/1362168808089921

Schmitt, N. (2010). Researching vocabulary: A vocabulary research manual. New York, NY: Palgrave Macmillan.

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.

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

Scott, M. (1999). *Wordsmith Tools.* Software. Oxford: Oxford University Press.

Sheehan, B. (2015). Teaching in English to non-English first language students in some Thai university programs in business and management. *PSAKU International Journal of Interdisciplinary Research, 4*(2), 1–11.

Shen, Z. (2008). The roles of depth and breadth of vocabulary knowledge in EFL reading performance. *Asia Social Science, 4*(12), 135–137.

Sinclair, J. (1991). Corpus, concordance, collocation: Describing English language. Oxford University Press, Oxford.

Sinclair, J. (2005). Corpus and Text – Basic Principles. In M. Wynne (ed.), *Developing Linguistic Corpora: a Guide to Good Practice*, pp. 1–21. Oxford Text Archive. Available at: http://ahds.ac.th/linguistic-corpora/.

Sridhanyarat, K. (2018). Thai Learners' Acquisition of L2 Collocations: An Interlanguage Perspective. *GEMA Online Journal of Language Studies, 18*(1), 1.

Sripicharn, P. (2010). How can we prepare learners for using language corpora. *The Routledge handbook of corpus linguistics*, 371–384.

Sriphicharn, P. (2002). Introducing Data-Driven Learning: Theory, Practice, and Evaluation. *Journal of Liberal Arts, 2*(2), 100–111.

Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *The Language Learning Journal, 36*(2), 139–152. https://doi.org/10.1080/09571730802389975

Sternberg, R.J. (1987). Most Vocabulary Is Learned from Context. In McKeown, M.G., and Curtis, M.E. (eds.), *The nature of vocabulary acquisition.* Hillside, NJ: Lawrence Erlbaum, 89–105

Sudathip, P., Udomsiang, P., & Sripoomsuriyakan, W. (2017). Enhancing Students' Vocabulary Learning through Storytelling for Grade 8 Students at Demonstration School Khon Kaen University. *Journal of Education Khon Kaen University, 40*(2), 9–16.

Syafitri, W., & Artika, F. S. (2019). English as a Lingua Franca and ASEAN Economic Community (AEC): A perspective of Islamic Economy Faculty in building strong Muslim students in Iain Bukittinggi. *ISLAM REALITAS: Journal of Islamic & Social Studies, 5*(1), 71–81.

Tang, K. N. (2021). Implementing English-medium instruction (EMI) in Thailand: University students' perspectives. *KKU Research Journal (Graduate Studies) Humanities and Social Sciences, 9*(2), 95–108.

Teng, F. (2016). An in-depth investigation into the relationship between vocabulary knowledge and academic listening comprehension. *TESL-EJ, 20*(2), 1–17.

Thammasat University. (n.d.). *Undergraduate Courses.* Retrieved November 12, 2020, from https://tu.ac.th/en/undergraduate

Thawarom, T., & Chaimongkol, S. (2013). Problems of Thai Students in Comprehending English Mathematics. *Journal of Liberal Arts Ubon Ratchathani University, 9*(2), 347–365.

Timmis, I. (2015). *Corpus linguistics for ELT: research and practice*. London: Routledge, Taylor & Francis Group, 2015.

Todd, R. W. (2017). An opaque engineering word list: Which words should a teacher focus on?. *English for Specific Purposes, 45*, 31–39.

Tseng, P.-H., Richards, K., & Pilcher, N. (2018). Constructing English-medium instruction indicators in the shipping courses of Taiwan's higher education. *Maritime Business Review, 3*(1), 20–35. https://doi.org/10.1108/MABR-07-2017-0020

Turhan, B., & Kirkgöz, Y. (2018). Motivation of engineering students and lecturers toward English medium instruction at tertiary level in Turkey. *Journal of Language and Linguistic Studies, 14*(1), 261–277.

Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *Journal of English for Academic Purposes, 12*(4), 248–263.

Verhoeven, L. (2000). Components in early second language reading and spelling. *Scientific Studies of Reading, 4*(4), 313–330.

Vidal, K., & Jarvis, S. (2020). Effects of English-medium instruction on Spanish students' proficiency and lexical diversity. *Language Teaching Research, 24*(5), 568–587.

Viroonphan, N. (2010). *World-class standard school.* Ministry of Education. Retrieved from http://www.moe.go.th/moe/th/news/detail.php?NewsID=15525&Key=news12

Vongpumivitch, V., Huang, J., & Chang, Y. (2009) Frequency analysis of the words in the Academic Word List (AWL) and non-AWL content words in applied linguistic research papers. *English for Specific Purposes, 28,* 33–41. http://dx.doi.org/10.1016/j.esp.2008.08.003.

Waluyo, B. (2019). Thai first-year university students' English proficiency on CEFR levels: A case study of Walailak University, Thailand. *The New English Teacher, 13*(2), 51–71.

Wang, S. (2015). An empirical study on the role of vocabulary knowledge in EFL listening comprehension. *Theory and Practice in Language Studies, 5*(5), 989–995.

Ward, J. (1999). How large a vocabulary do EAP engineering students need? Reading in a Foreign Language, 12(2), 309–323.

Ward, J. (2009). EAP reading and lexis for Thai engineering undergraduates. *Journal of English for Academic Purposes*, *8*(4), 294–301.

Wee, L., Pakir, A., & Lim, L. (2010). *English in Singapore: Modernity and Management.* Hong Kong University Press.

Wen-Ming, H. and Hsien-Chin, L. (2008). A Case Study of Corpus-Informed Online Academic Writing for EFL Graduate Students. *CALICO Journal*, (1), 28.

West, M. (1953). *A general service list of English words.* London: Longman, Green, & Co.

Xin, J. F., & L. Affrunti, R. (2019). Using iPads in vocabulary instruction for English language learners. *Computers in the Schools*, *36*(1), 69–82.

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

Yang, W., & Dai, W. (2011). Rote memorization of vocabulary and vocabulary development. *English Language Teaching*, *4*(4), 61–64. http://dx.doi.org/10.5539/elt.v4n4p61

Yang, M. (2015). A nursing academic word list. *English for Specific Purposes, 37*, 27–38.

Yang, M., O'Sullivan, P. S., Irby, D. M., Chen, Z., Chun, L., & Lin, C. (2019). Challenges and adaptations in implementing an English-medium medical program: A case study in China. BMC Medical Education, 19(15), 1–8. https://doi.org/10.1186/s12909-018-1452-3

Yeh, C.-Y., & Wang, Y.-H. (2004). An investigation into vocabulary learning strategies used by senior high school students in Taiwan. *Taiwan Journal of TESOL, 1*(2), 1–44.

Zhao, Q., & Li, H. (2018). Application of New Media Technology in College English Vocabulary Teaching. *Educational Sciences: Theory & Practice, 18*(6), 3591–3598. http://dx.doi.org/10.12738/estp.2018.6.271

Zhang, D., & Yang, X. (2016). Chinese L2 learners' depth of vocabulary knowledge and its role in reading comprehension. *Foreign Language Annals*, *49*(4), 699–715.

APPENDICES

APPENDIX A

International and English Undergraduate Degree Programs of Thai Higher Education Institutions (As of July 2023)

No.	Name of University	Number of International Programs	Number of English Programs	Fields of Study
1	Burapha University	10	1	*International Program* 1) Digital and Creative Marketing 2) Finance 3) Global Business Communication 4) International Business Management 5) International Human Resources 6) Business Administration 7) Holistic Health and Wellness Management 8) Hospitality, Tourism and MICE Management 9) Human Resources and Communication 10) Smart Logistics and Supply Chain Management *English Program* 1) Nursing Science *Bilingual Program* 1) Applied Artificial Intelligence and Smart Technology
2	Chiang Mai University	15	-	*International Program* 1) Accountancy 2) Digital Innovation 3) Economics 4) Environmental Science 5) Food Science and Technology 6) Humanities and Sustainability 7) Industrial Engineering and Logistics Management 8) Information Systems and Network Engineering 9) Integrated Design in Emerging Architecture 10) Mechanical Engineering 11) Modern Management and Information Technology 12) Nursing Science

No.	Name of University	Number of International Programs	Number of English Programs	Fields of Study
				13) Social Science and Sustainable Development 14) Software Engineering 15) Thai as a Foreign Language
3	Chulalongkorn University	17	-	*International Program* 1) Aerospace Engineering 2) Applied Chemistry 3) Architectural Design 4) Arts and Science in Integrated Innovation 5) Automotive Design and Manufacturing Engineering 6) Biotechnology 7) Business Administration 8) Chemical and Process Engineering 9) Communication Design 10) Communication Management 11) Economics 12) Information and Communication Engineering 13) Language and Culture 14) Nano-Engineering 15) Politics and Global Studies 16) Psychological Science 17) Robotics and Artificial Intelligence Engineering
4	Kasetsart University	16	8	*International Program* 1) Aerospace Engineering 2) Agro-Industrial Innovation and Technology 3) Bioscience and Technology 4) Business Administration 5) Communicative Thai Language for Foreigners 6) Digital Manufacturing and Robotics Integration Engineering 7) Electrical Engineering 8) Entrepreneurial Economics 9) Industrial Engineering 10) Integrated Chemistry 11) Marketing 12) Mechanical Engineering

No.	Name of University	Number of International Programs	Number of English Programs	Fields of Study
				13) Polymer Science and Technology 14) Software and Knowledge Engineering 15) Tourism Management 16) Tropical Agriculture English Program 1) Accounting 2) Economics 3) Financial and Investment 4) Hospitality Industry Management 5) International Business 6) Logistics Management 7) Management 8) Strategic Digital Marketing and Branding
5	Khon Kaen University	13	1	*International Program* 1) Accountancy 2) Chemical Process Engineering 3) Creative Media Technology 4) Logistics Engineering 5) Digital Media Engineering 6) Economics 7) Global Business 8) International Affairs 9) International Entrepreneurship 10) International Journalism 11) International Marketing 12) Telecommunications Engineering 13) Tourism Management *English Program* 1) Tourism Innovation Management *Bilingual Program* 1) Hospitality and Event Management
6	King Mongkut's Institute of Technology Ladkrabang	23	-	*International Program* 1) Aeronautical Engineering and Commercial Pilot 2) Aerospace Engineering 3) Applied Microbiology 4) Architecture

No.	Name of University	Number of International Programs	Number of English Programs	Fields of Study
				5) Biomedical Engineering 6) Business Administration 7) Business Information Technology 8) Chemical Engineering 9) Civil Engineering 10) Computer Innovation Engineering 11) Creative Arts and Curatorial Studies 12) Culinary Science and Foodservice Management 13) Electrical Engineering 14) Energy Engineering 15) Engineering Management and Entrepreneurship 16) Financial Engineering 17) Global Entrepreneurship 18) Industrial and Engineering Chemistry 19) Industrial Engineering and Logistics Management 20) Logistics Management 21) Mechanical Engineering 22) Robotics and AI Engineering 23) Software Engineering
7	King Mongkut's Institute of Technology North Bangkok	6	2	*International Program* 1) and Aerospace Engineering 2) Chemical Engineering 3) Informatics for Digital Economy 4) Innovative Materials Engineering 5) International Trade and Business Logistics 6) Mechanical Engineering *English Program* 1) Electrical Engineering 2) Robotic Engineering and Automation System
8	King Mongkut's University of	11	1	*International Program* 1) Architecture 2) Automation Engineering 3) Chemical Engineering

No.	Name of University	Number of International Programs	Number of English Programs	Fields of Study
	Technology Thonburi			4) Civil Engineering 5) Communication Design 6) Computer Engineering 7) Creative Technology 8) Digital Design 9) Design Innovation 10) Electrical Communication and Electronic Engineering 11) Environmental Engineering 12) Interior Architecture 13) Landscape Architecture 14) Multiple Intelligences for Design Integration *English Program* 1) Computer Science
9	Maejo University	-	-	No records found
10	Mahachulalongkornrajavidyalaya University	-	-	No records found
11	Mahamakut Buddhist University	-	-	No records found
12	Mahidol University	30	-	*International Program* 1) Actuarial Science 2) Applied Mathematics 3) Bioinnovation 4) Biological Sciences 5) Biomedical Engineering 6) Biomedical Science 7) Bioresources and Environmental Biology 8) Business Economics 9) Chemical Engineering 10) Chemistry 11) Communication Design 12) Computer Engineering 13) Computer Science 14) Environmental Science 15) Food Science and Technology 16) Industrial Engineering 17) Industrial Mathematics 18) Information and Communication Technology

No.	Name of University	Number of International Programs	Number of English Programs	Fields of Study
				19) Information Systems 20) Intercultural Studies and Languages 21) International Business 22) International Hospitality Management 23) International Relations and Global Affairs 24) Finance 25) Marketing 26) Materials Science and Nano Engineering 27) Media and Communication 28) Physics 29) Prosthetics and Orthotics 30) Prosthetics and Orthotics (Blended Distance Learning)
13	Prince of Songkla University	13	1	*International Program* 1) Rubber Engineering and Technology 2) Business Innovation Management 3) Chinese for Business Communication 4) Creative Media and Digital Technologies 5) Digital Business 6) Digital Engineering 7) Innovation Engineering and Management 8) International Business: China 9) International Studies 10) Islamic Studies 11) Hospitality Management 12) Management and Entrepreneurship 13) Tourism Management *English Program* 1) Nursing Science
14	Princess Galyani Vadhana	-	-	No records found

No.	Name of University	Number of International Programs	Number of English Programs	Fields of Study
	Institute of Music			
15	Thammasat University	31	2	*International Program* 1) Accounting 2) British and American Studies 3) Business English Communication 4) Business Law 5) Finance 6) Traditional Chinese Medicine 7) Marketing 8) Cardiovascular and Thoracic Technology 9) Industrial Engineering and Logistics Systems (SIIT) 10) Innovative Digital Design 11) International Studies (ASEAN-China) 12) Chemical Engineering 13) Chemical Engineering (SIIT) 14) Computer Engineering (SIIT) 15) Digital Engineering (SIIT) 16) Mechanical Engineering 17) Mechanical Engineering (SIIT) 18) Management Technology (SIIT) 19) Engineering Management (SIIT) 20) Civil Engineering 21) Civil Engineering (SIIT) 22) Electrical Engineering 23) Electrical Engineering (SIIT) 24) Chinese Studies 25) Optometry 26) Social Policy and Development 27) Philosophy, Politics, and Economics 28) Mass Media Studies 29) Thai Studies 30) Economics 31) Politics and International Relations *English Program*

No.	Name of University	Number of International Programs	Number of English Programs	Fields of Study
				1) Design, Business, and Technology Management 2) Medicine *Bilingual Program* 1) Dental Surgery

APPENDIX B

Education Plan of the Bachelor of Economics International Program

Year 1, Semester 1

Course	Credits
General Education Course, Part 1	9
General Education Course, Part 2	3
Principles of Microeconomics (EE211) or Principles of Macroeconomics (EE212)	3
Calculus for Social Science 1 (MA216) or Statistics for Social Science 1 (ST216)	3
Total Credits	**18**

Year 1, Semester 2

Course	Credits
General Education Course, Part 1	12
Principles of Macroeconomics (EE212) or Principles of Microeconomics (EE211)	3
Statistics for Social Science 1 (ST216) or Calculus for Social Science 1 (MA216)	3
Total Credits	**18**

Year 2, Semester 1

Course	Credits
General Education Course, Part 2	3
Microeconomic Theory (EE311) or Macroeconomic Theory (EE312)	4
Introductory Mathematical Economics (EE320) or Introductory Econometrics (EE325)	3
Cluster-Specific Economics Course	3
Minor or Non-Economics Elective Course	3
Free-Elective Course	3
Total Credits	**19**

Year 2, Semester 2

Course	Credits
Microeconomic Theory (EE311) or Macroeconomic Theory (EE312)	4
Introductory Mathematical Economics (EE320) or Introductory Econometrics (EE325)	3
Compulsory Cluster-Specific Course	6
Minor or Non-Economics Elective Course	6
Total Credits	**19**

Year 3, Semester 1

Course	Credits
Compulsory Cluster-Specific Course	6
Elective Cluster-Specific Course	3
Minor or Non-Economics Elective Course	6
General Education Course, Part 2	3
Total Credits	**18**

Year 3, Semester 2

Course	Credits
Compulsory Cluster-Specific Course	9
Minor or non-economics elective course	6
Free-elective course	3
Total Credits	**18**

Year 3, Summer

Course	Credits
EE300 Internship or EE366 Local Study and Development	3 6
Total Credits	**3 or 6**

Option 1: Seminar

Year 4, Semester 1

Course	Credits
EE460 Thai Economy or EE404 History of Economic Thought or EE406 Contemporary Economic Issues	3
Elective cluster-specific courses	3
Minor or non-economics elective course	3
Free-elective course	3
Total Credits	**12**

Year 4, Semester 2

Course	Credits
Seminar	3
Free-elective course	3
Total Credits	**6**

Option 2: Honors Thesis (EE500)

Year 4, Semester 1

Course	Credits
Thai Economy (EE460) or History of Economic Thought (EE404) or Contemporary Economic Issues (EE406)	3
Seminar Honor Thesis (EE400)	3
Minor or non-economics elective course	3
Free-elective course	3
Total	**12**

Year 4, Semester 2

Course	Credits
EE500 Honors Thesis	3
Free-elective course	3
Total	**6**

APPENDIX C

The Most Frequently Occurring AWL Word Family Headwords

1. **achieve**: achievable, achieved, achievement, achievements, achiever, achievers, achieves, achieving, unachievable

2. **affect**: affected, affecting, affects, unaffected

3. **aid**: aided, aider, aiders, aiding, aids, unaided

4. **alternative**: alternatively, alternatives

5. **analyse**: analysed, analyser, analysers, analyses, analysing, analysis, analyze, analyzed, analyzes, analyzing

6. **annual**: annually, annuals

7. **appendix**: appendices, appendixes

8. **approach**: approachable, approached, approaches, approaching, unapproachable

9. **appropriate**: appropriacy, appropriately, appropriateness, inappropriacy, inappropriate, inappropriately, inappropriateness

10. **approximate**: approx, approximated, approximately, approximates, approximating, approximation, approximations

11. **area**: areas

12. **assume**: assumed, assumes, assuming, unassuming, unassumingly

13. **assumption**: assumptions

14. **available**: availability, unavailability, unavailable

15. **capacity**: capacities, incapacities, incapacity

16. **chapter**: ch, chapt, chapters

17. **comment**: commentaries, commentary, commented, commenting, comments

18. **complex**: complexed, complexes, complexities, complexity

19. **computer**: computable, computation, computational, computationally, computations, compute, computed, computerisation, computerise, computerised, computerises, computerising, computerization, computerize, computerized, computerizes, computerizing, computers, computes, computing

20. **concept**: conception, conceptions, concepts, conceptual, conceptualisation, conceptualisations, conceptualise, conceptualised, conceptualises, conceptualising, conceptualization, conceptualizations, conceptualize, conceptualized, conceptualizes, conceptualizing, conceptually

21. **conclude**: concluded, concludes, concluding, unconcluded

22. **conclusion**: conclusions

23. **consequence**: consequences

24. **consist**: consisted, consisting, consists

25. **consistent**: consistently, inconsistent

26. **constant**: constantly, constants, inconstantly

27. **construct**: constructed, constructing, construction, constructional, constructions, constructive, constructively, constructivist, constructivists, constructor, constructors, constructs, reconstruct, reconstructed, reconstructing, reconstruction, reconstructions, reconstructs, unreconstructed

28. **consume**: consumable, consumables, consumed, consumer, consumerism, consumerist, consumers, consumes, consuming

29. **consumption**

30. **coordinate**: coordinated, coordinates, coordinating, coordination, coordinator, coordinators, ordinated, ordinating, ordinator, ordinators, uncoordinated,

31. **correspond**: corresponded, correspondence, correspondences, corresponding, correspondingly, corresponds

32. **create**: created, creates, creating, creation, creations, creative, creatively, creativity, creator, creators, recreate, recreated, recreates, recreating

33. **cycle**: cycled, cycles, cyclic, cyclical, cycling, cyclist, cyclists

34. **data**: datum

35. **decline**: declined, declines, declining

36. **define**: definable, defined, defines, defining, definition, definitional, definitions, indefinable, indefinably, predefined, redefine, redefined, redefines, redefining, redefinition, redefinitions, undefinable, undefined

37. **derive**: derivation, derivations, derivative, derivatives, derived, derives, deriving

38. **design**: designed, designer, designers, designing, designs

39. **distribute**: distributable, distributed, distributes, distributing, distribution, distributional, distributions, distributive, distributor, distributors, redistribute, redistributed, redistributes, redistributing, redistribution, redistributive

40. **economy**: economic, economical, economically, economics, economies, economise, economised, economises, economising, economist, economists, economize, economized, economizes, economizing, uneconomic, uneconomical, uneconomically

41. **eliminate**: eliminated, eliminates, eliminating, elimination, eliminations, eliminator, eliminators

42. **energy**: energetic, energetically, energetics, energies, energise, energised, energiser, energisers, energises, energising, energize, energized, energizer, energizers, energizes, energizing

43. **ensure**: ensured, ensures, ensuring

44. **environment**: environmental, environmentalism, environmentalist, environmentalists, environmentally, environments

45. **equation**: equations

46. **equivalent**: equivalently, equivalents

47. **estimate**: est, estimated, estimates, estimating, estimation, estimations, estimator, estimators

48. **evaluate**: evaluated, evaluates, evaluating, evaluation, evaluations, evaluative, evaluator, evaluators

49. **eventually**: eventual, eventualities, eventuality

50. **evidence**: evidenced, evidences, evidential

51. **exceed**: exceeded, exceeding, exceedingly, exceeds

52. **expand**: expandability, expandable, expanded, expanding, expands

53. **factor**: factored, factoring, factors

54. **feature**: featured, featureless, features, featuring

55. **final**: finalise, finalised, finalises, finalising, finalist, finalists, finality, finalize, finalized, finalizes, finalizing, finally, finals, semifinalist, semifinalists

56. **focus**: foci, focused, focuses, focusing, focussed, focusses, focusing, refocus, refocused, refocuses, refocusing, refocussed, refocusses, refocussing, unfocused, unfocussed

57. **function**: functional, functionalism, functionalist, functionalists, functionalities, functionality, functionally, functioned, functioning, functions,

58. **goal**: goalless, goals

59. **hence**

60. **hypothesis**: hypotheses, hypothesise, hypothesised, hypothesises, hypothesising, hypothesize, hypothesized, hypothesizes, hypothesizing

61. **identical**: identically

62. **identify**: identifiable, identification, identifications, identified, identifier, identifiers, identifies, identifying, identities, identity, unidentifiable, unidentified

63. **ignore**: ignored, ignores, ignoring

64. **illustrate**: illustrated, illustrates, illustrating, illustration, illustrations, illustrative, illustrator, illustrators, unillustrated

65. **impact**: impacted, impacting, impacts

66. **income**: incomes

67. **indicate**: indicated, indicates, indicating, indication, indications, indicative, indicator, indicators

68. **individual**: individualise, individualised, individualism, individualist, individualistic, individualists, individuality, individualize, individualized, individually, individuals

69. **initial**: initialization, initialisations, initialise, initialised, initialises, initialising, initialization, initializations, initialize, initialized, initializes, initializing, initialled, initialling, initially, initials

70. **input**: inputs, inputting

71. **insight**: insightful, insights

72. **instance**: instanced, instances

73. **invest**: invested, investing, investment, investments, investor, investors, invests, reinvest, reinvested, reinvesting, reinvestment, reinvests

74. **involve**: involved, involvement, involvements, involves, involving, uninvolved

75. **item**: itemisation, itemise, itemised, itemises, itemising, itemization, itemize, itemized, itemizes, itemizing, items

76. **label**: labeled, labeling, labelled, labeller, labellers, labelling, labels, unlabeled, unlabelled

77. **labour**: labor, labored, laborer, laborers, laboring, labors, labored, labourer, labourers, labouring, labourism, labours

78. **locate**: located, locater, locaters, locates, locating, location, locational, locations, locator, locators, relocate, relocated, relocates, relocating, relocation, relocations

79. **maintain**: maintainability, maintainable, maintained, maintaining, maintains, maintenance, unmaintained

80. **major**: majored, majoring

81. **margin**: marginal, marginalisation, marginalise, marginalised, marginalising, marginality, marginalization, marginalize, marginalized, marginalizing, marginally, marginals, margined, margins

82. **maximise**: maximisation, maximised, maximises, maximising, maximization, maximize, maximized, maximizes, maximizing

83. **maximum**: max, maximums

84. **medium**: mediums

85. **method**: methodical, methodically, methodological, methodologically, methodologies, methodology, methods

86. **minimum**: minimums

87. **negative**: negatively, negatives, negativism, negativity

88. **nonetheless**

89. **normal**: normalisation, normalise, normalised, normalises, normalising, normality, normalization, normalize, normalized, normalizes, normalizing, normally

90. **obtain**: obtainable, obtained, obtaining, obtains, unobtainable

91. **occur**: occurred, occurrence, occurrences, occurring, occurs, reoccur, reoccurred, reoccurring, reoccurs

92. **percent**: percentage, percentages

93. **period**: periodic, periodical, periodically, periodicals, periodicity, periods

94. **physical**: physicality, physically

95. **plus**: pluses

96. **portion**: portioned, portioning, portions

97. **positive**: positively, positives, positivism, positivist, positivistic, positivists, positivity

98. **potential**: potentialities, potentiality, potentially, potentials

99. **precede**: preceded, precedence, precedent, precedents, precedes, preceding, unprecedented, unprecedentedly

100. **predict**: predictability, predictable, predictably, predicted, predicting, prediction, predictions, predictive, predictor, predictors, predicts, unpredictability, unpredictable, unpredictably

101. **previous**: previously

102. **primary**: primaries, primarily

103. **principle**: principled, principles, unprincipled

104. **procedure**: procedural, procedures

105. **proceed**: proceeded, proceeding, proceedings, proceeds

106. **process**: processed, processer, processers, processes, processing, processor, processors, reprocess, reprocessed, reprocesses, reprocessing, reprocessings, unprocessed

107. **project**: projected, projecting, projection, projectionist, projectionists, projections, projective, projects

108. **publish**: publishable, published, publisher, publishers, publishes, publishing, republish, republished, republishes, republishing, unpublished

109. **range**: midrange, ranged, ranges, ranging

110. **ratio**: ratios

111. **region**: regional, regionalism, regionalist, regionalists, regionally, regions

112. **require**: required, requirement, requirements, requires, requiring

113. **resource**: resourced, resourceful, resourcefulness, resources, resourcing, underresourced, unresourceful

114. **restrict**: restricted, restricting, restriction, restrictions, restrictive, restrictively, restricts, unrestricted, unrestrictive

115. **revenue**: revenues

116. **role**: roles

117. **section**: sectional, sectioned, sectioning, sections, subsection, subsections

118. **sector**: sectoral, sectors

119. **series**

120. **shift**: shifted, shifting, shifts

121. **similar**: dissimilar, dissimilarities, dissimilarity, similarities, similarity, similarly

122. **somewhat**

123. **source**: sourced, sources, sourcing, unsourced

124. **strategy**: strategic, strategically, strategies, strategist, strategists

125. **structure**: poststructuralism, poststructuralist, restructure, restructured, restructures, restructuring, structural, structuralism, structuralisms, structuralist, structuralists, structurally, structured, structures, structuring, unstructured

126. **substitute**: substituted, substitutes, substituting, substitution, substitutions

127. **sum**: summed, summing, sums

128. **summary**: summaries, summarily, summarisation, summarisations, summarise, summarised, summarises, summarising, summarization, summarizations, summarize, summarized, summarizes, summarizing

129. **survey**: surveyed, surveying, surveyor, surveyors, surveys, unsurveyed

130. **task**: tasked, tasking, tasks

131. **team**: teamed, teaming, teams

132. **technology**: tech, technological, technologically, technologies, technologist, technologists

133. **theory**: theories, theorist, theorists

BIOGRAPHY

Name	Mr. Pongsathon Wasikarat
Date of Birth	Thursday, January 28th, 1988
Educational Attainment	Academic Year 2019: Bachelor of Education (B.Ed.) in Educational Measurement and Evaluation Sukhothai Thammathirat Open University, Thailand Academic Year 2013: Master of Arts (M.A.) in Teaching English as a Foreign Language Thammasat University, Thailand Academic Year 2010: Bachelor of Arts (B.A.) in Business English Khon Kaen University, Thailand
Work Position	English Language Lecturer Faculty of Liberal Arts and Science Nakhon Phanom University
Scholarship	Fiscal Year 2017–2019: Employee Scholarship Fund (Full-Time Study Leave)
Work Experiences	2014–Present: English Language Instructor Nakhon Phanom University 2013–2014: English Program (EP) Teacher Anuban Khon Kaen School