THE DEVELOPMENT OF
A SCIENCE ACADEMIC WORD LIST
(SAWL)

BY

MR. TODSAPORN IT-NGAM

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY PROGRAM IN
ENGLIH LANGUAGE TEACHING
LANGAUGE INSTITUTE
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2018

THE DEVELOPMENT OF
A SCIENCE ACADEMIC WORD LIST
(SAWL)

BY

MR. TODSAPORN IT-NGAM

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY PROGRAM IN
ENGLIH LANGUAGE TEACHING
LANGUAGE INSTITUTE
THAMMASAT UNIVERSITY
ACADEMIC YEAR 2018

THAMMASAT UNIVERSITY
LANGUAGE INSTITUTE

DISSERTATION

BY

MR. TODSAPORN IT-NGAM

ENTITLED

THE DEVELOPMENT OF
A SCIENCE ACADEMIC WORD LIST (SAWL)

was approved as partial fulfillment of the requirements for
the degree of Doctor of Philosophy (English Language Teaching)

on December 27, 2018

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

Member and Advisor _____
(Assistant Professor Supakorn Phoocharoensil, Ph.D.)

Member _____
(Assistant Professor Passapong Sripicharn, Ph.D.)

Member _____
(Assistant Professor Jirada Wudthayagorn, Ph.D.)

Member _____
(Chanika Gampper, Ph.D.)

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

Dissertation Title	THE DEVELOPMENT OF A SCIENCE ACADEMIC WORD LIST (SAWL)
Author	Mr. Todsaporn It-ngam
Degree	Doctor of Philosophy Program in English Language Teaching
Major Field/Faculty/University	English Language Teaching (ELT) Language Institute Thammasat University
Dissertation Advisor	Assistant Professor Supakorn Phoocharoensil, Ph.D.
Academic Years	2018

ABSTRACT

English has become the academic language of science. EFL science students need to acquire academic vocabulary used in science disciplines in order to gain academic success. The present study was conducted to investigate English academic words frequently occur in journal articles in natural science disciplines and to establish the Science Academic Word List (SAWL), which consists of 853 word families covering 11 disciplines of natural sciences. To make the list easier to use, this study was also aimed to align the SAWL words with the Common European Framework of Reference for Languages. The development of the SAWL employed three main approaches: the corpus-based approach, the expert-judged approach, and the CEFR alignment approach.

First, the corpus-based approach was used to select specialized academic words frequently occurs in the Science Academic Journal (SAJ) corpus, which is the corpus of 1,062 research articles review papers in natural science disciplines. The SAJ corpus comprises 5.5 million running words divided into 11 subcorpora. The specialized academic words were chosen based on 3 word selection criteria: special occurrence, range, and frequency. The potential SAWL contains 899 word families. Second, the 899 potential SAWL words were reviewed through the expert-judged approach. They were rated by 3 experts in natural sciences using modified Chung and Nation's (2004) rating scale in order to measure their strength of relationship with the

natural science disciplines. Forty-six word families were removed from the list because their meanings were judged to be too specific. The final SAWL consists of 853 word families – 421 AWL words and 432 non-AWL words. The validity tests showed that the SAWL accounts for 15.5% coverage of the running words in the SAJ corpus – higher than Coxhead's (2000) AWL and Coxhead and Hirsh's (2007) SWL. Third, to facilitate word list users, the SAWL words were aligned with the CEFR levels: A1, A2, B1, B2, C1, C2, and Unlisted. The CEFR alignment criteria were established. It was revealed that the majority of SAWL words are at B2 (350 words), C1 (171 words), and B1 (158 words) respectively. The findings were validated against the SAJ corpus and Ishikawa's (2013) ICNALE Written Essay corpus.

Additionally, academic collocations and other lexical information of SAWL words were highlighted to help EFL science students increase their proficiency and to support ESP teachers in designing their courses. Finally, this study provides an online research tool for further investigations into specialized academic vocabulary learning.

Keywords: specialized academic words, natural science, frequency word list, corpus-based analysis, CEFR, collocations

ACKNOWLEDGEMENTS

While working on this dissertation, I received tremendous supports from a number of people. Firstly, I would like to express my deepest gratitude to my supervisor, Assistant Professor Dr. Supakorn Phoocharoensil for his continuous support, patience, encouragement, and valuable advice throughout my Ph.D. journey. Without him, this dissertation would not have been completed successfully.

I would like to thank my thesis committee: Associate Professor Dr. Supong Tangkiengsirisin, Assistant Professor Dr. Jirada Wudthayagorn, Assistant Professor Dr. Passapong Sripicharn, and Dr. Chanika Gampper for their invaluable guidance, thought-provoking questions, and constructive comments, which have led to the completion of thish thesis. I would especially like to thank Professor Paul Nation and Professor Averil Coxhead for being my inspiration and source of helpful comments given to me during the special occasions held by LITU. I would like to thank all the experts from the Faculty of Science, Burapha University, who assisted in evaluating the words in SAWL. Definitely, I would like to thank Mr. Jeffrey Siegfried who patiently proofread my thesis again.

My sincere thanks go to all lecturers and staff at the Ph.D. in ELT program. The knowledge gained through the program was very profound and comprehensive. Certainly, I would like to thank all of my Ph.D. classmates. The gaps in my knowledge have been filled after discussing and learning together with them.

My gratitude also goes to Assistant Professor Dr. Punnee Pimapunsri, Director of Burapha University Language Institute (BUULI), who allowed me to take full academic leave to complete my dissertation, and Assistant Professor Dr. Charan Chakandang, the former director of BUULI , who allowed me to pursue my dream and allocated the 3-year fund for my study. Additionally, I would like to thank all colleagues at BUULI who always encourage and help me.

In addition, I would like to thank everyone in my family for their great patience and wholehearted support. Without their help, my journey would have been extremely tiring and tortuous. Particularly, I dedicate this dissertation to Suparuthai It-ngam, my wife. It was an amazing chance that we had to begin our Ph.D. journey together and complete it in a similar timeframe. Finally, I am more than grateful for my

adorable and lively children – Kane and Kori. You two keep my life energized every day.

Mr. Todsaporn It-ngam

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Symbols/Abbreviations	Terms
^	The highest frequency word in family
*	The coinciding word with AWL (Coxhead, 2000)
ACL	Academic Collocation List (Ackermann & Chen, 2013)
AVL	Academic Vocabulary List (Gardner & Davies, 2014)
AWL	Academic Word List (Coxhead, 2000)
BNC	British National Corpus
BUU	Burapha University
CAWL	Chemistry Academic Word List (Valipouri & Nassaji, 2013)
CEFR	Common European Framework of Reference
COCA	Corpus of Contemporary American English
EAWL	Environmental Academic Word List (Liu & Han, 2015)
EEWL	Engineering English Word List (HSU, 2014)
EFL	English as a Foreign Language
ELT	English Language Teaching
EMI	English Medium Instruction
ENC	English News Corpus (Goldhahn, Eckart, & Quasthoff, 2012)
ESP	English for Specific Purposes
EST	English for Science and Technology
ESTCWL	English for Science and Technology Coursebook Word List
EVP	English Vocabulary Profile
ICNALE	International Corpus Network of Asian Learners of English (Ishikawa, 2013)
ICT Word List	A word list for Information and Communication Technology (Pugsee, Limgomolvilas, Wudthayagorn, & Janpugdee, 2017).

Symbols/Abbreviations	Terms
GSE	Global Scale of English
GSL	General service List (West, 1953)
MAVL	Medical Academic Vocabulary List (Lei & Liu, 2016)
μAWL	Microbiology Academic Word List (Boonyos, 2014)
NAWL	Nursing Academic Word List (Yang, 2015)
OHEC	Office of Higher Education Commission
POS	Part of Speech
RLDs	Reference Level Descriptions
SAJ Corpus	Science Academic Journal Corpus
SAWL	Science Academic Word List
SWL	Science Word List (Coxhead & Hirsh, 2007)
UWL	University Word List (Xue & Nation, 1984)
WEC	Written Essay Corpus (Ishikawa, 2013)

CHAPTER 1
INTRODUCTION

English has become the accepted global language. It is used among various groups of people for different communication purposes as a lingua franca. The importance of English has resulted in the growing number of English language learners around the world. According to *The English Effect* (British Council, 2013), there were 1.75 billion English learners worldwide and it is predicted that the number of people who learn English will be around two billion by 2020. This number is comprised of students learning English throughout formal education and older learners studying English in their workplace or in their free time. Computer technologies nowadays have been changing the way people learn and use English language. Corpus analysis and corpus-based frequency word lists are integral parts of the learning innovations. They reveal the authentic uses of the language, which are especially helpful for those who are learning and using English for academic purposes (Motteram, 2013).

1.1 English as the global academic language

English has achieved another status as the 'global academic language' (Crystal, 2012; Northrup, 2013) or the 'academic lingua franca' (Mauranen, Hynninen, & Ranta, 2010). It has a significant role as the medium of accessing the world's knowledge. More institutions offer courses and programs taught in English for international students. Publications circulated in English are viewed as prestigious works. Especially in science and technology areas, English is the dominant academic and scientific language (Philip, 2007). This shift has been witnessed throughout the world, especially at higher education institutes because the internationalization of universities has become a priority for them (Galloway, Kriukow, & Numajiri, 2017). An institution which wishes to become a center of international excellence needs to attract teachers and researchers from around the world. Moreover, it needs to encourage international students to enroll in its programs, thereby enhancing its reputation, income, and academic climate (Graddol, 2006).

The global English movement also alters the ways English is taught and learned worldwide. David Graddol's English Next (2006) reviews the trends of English learning and teaching for the near future. Previously, English education was seen as an isolated discipline which had its own identity. English as a foreign language (EFL) and English as a second language (ESL) have been the dominant approaches of teaching English. In recent years, however, many more models of English education have emerged in order to meet the specific needs of the learners in their local use of English.

The focus of global English is on international intelligibility rather than a specific variety of the language such as American or British (Graddol, 2006). English as a lingua franca (ELF), for example, is the approach arising from this trend of intercultural communication. In other words, the aim of ELF is not to become a native speaker, but rather to achieve the level of a fluent bilingual speaker who has pragmatic strategies to communicate with another non-native speaker. English for young learners (EYL) has also been introduced as another approach for young learners to learn the foreign language. The fact is that the age of English learners has steadlily been decreasing worldwide as many countries promote EYL to show their political and economic aspirations to create bilingual citizens. According to Graddol (2006), English is being taught in primary schools across the world. At secondary schools and higher institutions, several programs or courses have been taught through the English language. This trend brought about another approach of English education – English medium instruction (EMI). Both curriculum content and English collaborate in this approach. These new approaches reflect the fact that the specific needs and aspirations of the language users are the departures of English learning.

1.2 The English-medium instruction in Thailand

One of the important English education policies of Thailand is to support the EMI programs in schools and universities. In 2002-2003, there were only two types of EMI programs at the primary and secondary school levels: English Program (EP) and Mini-English Program (MEP) (Office of the National Education Commission, 2003). Later, apart from EP and MEP, a wider variety of EMI programs have been added such as International Programs (IP), English Bilingual Education (EBE), and

English for Integrated Studies (EIS) (Office of The Basic Education Commision, 2014). The predominant differences among these school programs are their curricula and the number of subjects taught in English. For higher education institutions, both Thai public and private universities offer EMI programs (also known as international programs) at undergraduate and graduate levels. The EMI programs have gained much popularity and more EMI programs have been offered each year as shown in Table 1.1. In 2004, Thai universities had only 465 EMI programs. In 2014, Office of the Higher Education Commission (2014) reports that there were 768 EMI programs taught in 156 institutions across the country. While the EMI programs at primary and secondary schools aim to enhance the students' English competency, the main purpose of EMI programs at universities is to further academic excellence and improve the international reputation. The students who want to enroll in the university EMI programs are expected to have a good command of English. For example, a minimum IELTS band score of 5.0 or 5.5 is required in order to apply for most international bachelor's degree programs in Thailand.

Table 1.1
EMI programs offered by Thai public and private universities (OHEC, 2009; OHEC, 2017)

Years	Undergraduate	Master Degree	Doctoral Degree	Others	Total
2004	253	203	109	0	456
2006	214	290	178	18	727
2008	296	350	215	23	884
2010	342	389	225	25	981
2012	344	394	249	30	1,017
2014	249	290	224	6	769

A significant gap, however, is that Thai school students are not well equipped with English language for learning at the university level. In other words, the students at school spend much time studying English grammar and structures, which are areas of central importance in university entrance examinations. Academic English, therefore, is out of their focus. Even though most university students in Thailand are required to study English as a compulsory subject for one year, the primary goal of such English courses is general communication. These students need to be prepared

explicitly for English for learning at university, especially in their specific fields of study.

1.3 The Impact of the Common European Framework of Reference for Languages

The Common European Framework of Reference for languages (CEFR) is the standard framework of language learning and teaching. It has turned English education worldwide, including Thailand, in new directions. Originally, the CEFR was published in 2001 as a part of the Council of Europe's language policy to promote the learning of several European languages and to encourage a positive attitude towards linguistic diversity among European countries (Council of Europe, 2001). Nowadays, it has gained global recognition. A study on the impact of the CEFR (Figueras, 2012) reports that the CEFR has been translated into more than 40 languages, including sign language. The Council of Europe also reports that the CEFR is frequently referred to by many countries around the world for the development of their foreign language curricula and policies (Martyniuk & Noijons, 2007).

The CEFR pinpoints that learner's communicative needs and aspirations should become the center of teaching, learning, and assessment. Although the CEFR aims to promote plurilingualism (or multilingualism) among European countries, the six reference levels and descriptions of the CEFR (see Figure 1.1) have been adapted to fit the specific context and needs of the English language teachers and leaners around the world. The CEFR is viewed as an innovation for language learning in the global English era.

Similar to other countries in Asia, Thailand has shifted to the global English movement. The national Thai language is the primary language for communication, but the English language is a key part of the country's economic and social development. English has been used as a business language in multinational companies across the country. The government of Thailand has realized the importance of English language and has subsequently tried to enhance the English competency of Thai citizens (Office of the Education Council, 2017). A great attempt was seen in 1996 when the English syllabus was introduced as the policy to start English at Grade 1. Although the policy

failed (Graddol, 2006), later educational reforms have continued to prioritize English language education.

One of the attempts is to align the national curriculum with the international English education framework. In 2014, the Ministry of Education announced that the CEFR would be used as the "main principles for the enacting of English language teaching and learning in Thailand" (Office of The Basic Education Commision, 2014, p. 1). The six levels of the CEFR have been referred to as the end goal of each education level as shown in Table 1.2. When students are at Grade 6, they should achieve the A1 level. At Grade 9, they should achieve the A2 level. At Grade 12 (including students in High Vocational Diploma programs), they should achieve the B1 level. The policy implies that university students should achieve the B1 or B2 levels. Due to the fact that university students must be able to use language to elaborate on their thoughts in more specialized areas, it seems that the government policy in 2014 well matches the descriptors of the B2 level in that the learners at the B2 level *can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialization*. Therefore, the B2 level becomes the desirable proficiency level for undergraduate and graduate students.

In 2017, this policy was explicitly emphasized again in the National Education Plan B.E. 2560 - 2579. According to this 20-year plan, in Table 1.3, Thai undergraduates should achieve the B2 level. Towards the end of the plan in 2036, those undergraduates should be able to communicate in English above the C1 level (Office of the Education Council, 2017). Obviously, at least until 2036, the CEFR will have a significant role in Thailand's English education.

Table 1.2
The goal of English proficiency development for Thai students (Office of The Basic Education Commision, 2014).

Education Levels	Proficiency Levels	CEFR Levels
1. Primary Education (Grade 6)	Basic User	A1
2. Compulsory Education (Grade 9)	Basic User	A2
3. Fundamental Education (Grade 12 / Vocational Cert.)	Independent User	B1

Table 1.3
The achievement index of the National Education Plan – The average English proficiency levels (Office of the Education Council, 2017).

Years	Lower Secondary School Students	Upper Secondary School Students	Undergraduates
1. Present	A1	A2	B2
2. 2017 – 2021	A1	A2	B2
3. 2022 – 2026	A2	B1	B2+
4. 2027 – 2031	B1	B1+	C1
5. 2032 – 2036	B2	B2	C1+

Figure 1.1
The CEFR descriptors (Council of Europe, 2001)

PROFICIENT USER	C2	Can understand with ease virtually everything heard or read. Can summarize information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
	C1	Can understand a wide range of demanding, longer texts, and recognize implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organizational patterns, connectors and cohesive devices.
INDEPENDENT USER	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialization. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans.
BASIC USER	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.

Figure 1.1 (Cont.)
The CEFR descriptors (Council of Europe, 2001)

BASIC USER	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

1.4 Science studies for social development

Science is one of the predominant academic disciplines with a strong establishment in most Thai universities. Due to the fact that scientific progress is important for economic growth and greater social well-being, the government of Thailand has given priority to science education since the First National Economic and Social Development Plan in the 1960s (Office of the National Economic and Social Development Board, 1961). Recent national economic development policies, known as Thailand 4.0 and Thailand Digital Economy, have stressed the significance of science education as being fundamental for other cutting-edge technology and economic development.

Table 1.4
University students enrolling in the Faculty of Science in Thai Higher Institutions, Academic Year of 2016 (Higher Educational Information, 2017).

Institutions	Undergrad-uate	Master Degree	Doctoral Degree	Others	Total
1. Public Universities	36,404	3,981	2,067	887	43,329
2. Rajabhat Universities	15,782	213	-	554	16,549
3. Open Universities	3,866	125	16	2	4,009
4. Private Universities	951	12	-	-	963
5. Private Colleges	83	-	-	-	83
Total	**57,086**	**4,331**	**2,083**	**1,433**	**64,933**

Thai universities and other higher institutions play an important role in producing well-equipped graduates in science to serve the country's development. According to the statistical information (Table 1.4), in 2016, there were almost 65,000

university students enrolling in the Faculty of Science in 34 institutions across the country (Higher Educational Information, 2017). Students from other countries are also welcome to join these scientific institutions. The report in 2016 shows that there were international students from 40 nationalities enrolling in scientific programs in Thai universities. Therefore, the English language is established as the academic language of the science discipline despite the fact that the majority of the students are Thai. Moreover, in the scientific area, English is commonly used in order to acquire the world's knowledge and to gain global recognition. In several Thai universities, English has been used as the medium of instruction for some programs and courses. However, the way English language is used in this discipline is unique (Halliday, 2004; McComas, 2014; Reeves, 2005). These facts imply the specific needs of English learning for science studies. Unfortunately, as discussed earlier in this chapter, university students in Thailand are not well equipped with academic and specialized English language used in their field of study. Science students and English teachers dealing with scientific language need to have learning materials that support the specific requirements. The uniqueness of English in science disciplines can be revealed through a corpus-based approach to study language.

1.5 Corpus linguistics and word list development

The language of science is different from that of several other academic disciplines. Reeves (2005) describes that the scientific language is a simple, descriptive system. The language in scientific reports must be "as free as possible from connotations that reflect or create cultural biases and emotional attachment" (p.10) because the goal of scientists is to impartially report facts. However, like other academic disciplines, scientific English has its own specialization. According to Halliday (2004), scientific English has many technical features developed over time by experts. These features could cause difficulty for non-English speaking science students and turn science into "the prerogative of an elite" (p.viii). In other words, only experts in this field could perceive and use the specialized language with ease. One way that language educators can help these students is to make them understand how English

naturally works in the scientific field. Corpus linguistics could be employed to illustrate the authentic use of English to the students.

Corpus linguistics is an approach to study language in use based on corpora (McEnery & Wilson, 1996). Corpus linguistics does not aim to study particular aspect of language, but it is a systematic analysis of the naturally occurring language. Language scholars who employ corpus linguistics must combine it with other related methods and theories (Lindquist, 2009). However, the basic essential resource for corpus linguistics is a corpus, a large principled collection of texts. Language teachers can utilize quantitative and qualitative findings from the corpus approach to elicit information about words, phrases, grammar, situations of use, and nuances of language in a specific field. This approach, as a result, is tremendously useful for appropriate syllabus design, especially for teaching learners who have specific needs, namely English for Specific Purposes (ESP) courses (Bennett, 2010).

Corpus-based frequency word lists are the outcome of the studies in corpus linguistics. These word lists are increasingly important for designing the ESP syllabus. Gardner and Davies (2014, p. 306) point out that the frequency word lists have many pedagogical functions:

- establishing vocabulary learning goals;
- assessing vocabulary knowledge and growth;
- analyzing text difficulty and richness;
- creating and modifying reading materials;
- designing vocabulary learning tools;
- determining the vocabulary components of academic curricula; and,
- filling many other crucial academic needs.

West's (1953) *A General Service List of English Words* (GSL) is one of the early lists of frequency words for language teaching and it still remains influential in designing syllabi nowadays. The GSL contains around 2,000 words with their frequencies. Nation and Kyongho (1995) report that the coverage of the GSL is around 75% of the running words in non-fiction texts and around 90% of the running words in fiction text. They also suggest that the GSL is essential for all learners, no matter for

general or special purposes. There has been criticism and several attempts have been made to replace the GSL, but most of the attempts have failed (Nation, 2016).

Coxhead's (2000) *Academic Word List* (AWL) is the corpus-based word list of 570 academic word families. The list was built for non-English speaking university students in English-speaking countries based on the assumption that they were already familiar with the words in the GSL (Nation, 2016). It has become widely known and used as a guideline of learning and teaching English for academic purposes and its creating principle has been used to develop other academic word lists (Coxhead, 2011, 2016). In recent years, there has been a strong interest in making new word lists of general academic vocabulary (e.g. Gardner and Davies's (2014) *New Academic Vocabulary List*; Dang, Coxhead, and Webb's (2017) *The Academic Spoken Word List*) and more specialized vocabulary (e.g. Hsu's (2014) *An Engineering English Word List*; Yang's (2015) *A Nursing Academic Word List*; Lei and Liu's (2016) *A New Medical Academic Word List*).

In scientific discipline, corpus linguistics has been employed to develop word lists for pedagogical purposes. The *Science Word List* (SWL) (Coxhead & Hirsh, 2007) has been developed based on a pilot Science corpus of written academic English. The corpus covers 14 subject areas as majors in a science degree (see Table 1.5). The *Chemistry Academic Word List* (CAWL) (Valipouri & Nassaji, 2013) has been made exclusively for EFL chemistry students. The *Environmental Academic Word List* (EAWL) (Liu & Han, 2015) was established as the first academic word list for environmental science. These word lists have been proven useful for the students in these specific fields.

However, more corpus-based word lists in scientific fields have been called for because the existing ones might not serve the needs of other groups of word list users. It is worth noting that none of the mentioned word lists has been aligned with the national education curricula or the framework of language learning, like the CEFR. As a result, the existing word lists do not seem to appropriately serve EFL science students' needs.

Table 1.5
The fourteen subject areas of the pilot science corpus (Coxhead & Hirsh, 2007).

Subject Area	Token
1. Agricultural Science	129,492
2. Biology	125,898
3. Chemistry	124,400
4. Computer Science	124,589
5. Ecology	123,759
6. Engineering and Technology	128,561
7. Geography	125,833
8. Geology	125,144
9. Horticultural Science	128,124
10. Mathematics	127,234
11. Nursing and Midwifery	124,218
12. Physics	123,136
13. Sport and Health Science	124,488
14. Veterinary and Animal Science	126,504
Total	**1,761,380**

1.6 English for science at Burapha University

Established in 1955, Burapha University (BUU) is a comprehensive university located in the eastern region of Thailand. The university offers a wide variety of degrees in many disciplines in the sciences, technology, health science, and humanities and social sciences. At present, there is an enrollment of over 38,000 students spread among the three campuses of Burapha University, including Chon Buri, Sa Keao, and Chanthaburi.

Faculty of Science is one of the oldest faculties of Burapha University. It was founded in 1955, together with the university, to produce undergraduates in science disciplines. At present, the faculty encompasses 11 undergraduate degrees, and 21 graduate degrees in 9 departments. The approximate number of students is 2,000 students. The undergraduate degrees are comprised of aquatic science, biology, biochemistry, biotechnology, chemistry, food chemistry, mathematics, statistics, microbiology, physics, and applied physics.

According to the current university curriculum, undergraduate students are required to enroll in three compulsory English courses, namely *English for Communication*, *Collegiate English,* and *English Writing for Communication*. The

students in the Faculty of Science are further required to enroll in another ESP course, *English for Science and Technology*. All of these language courses are under the supervision of Burapha University Language Institute, the university's central unit for foreign language education. The courses are also aligned with the CEFR as shown in Table 1.6.

Table 1.6
The alignment of the BUU English courses with the CEFR.

Course Titles	CEFR Levels
English for Communication	A2+
Collegiate English	B1
English Writing for Communication	B1
English for Science and Technology	B1/B1+

To promote an international academic environment, the faculty has applied the EMI approach to its programs. Teachers are encouraged to use English as much as possible in class. Most of the teaching materials and reading texts are written in English. Furthermore, in their third or fourth year, science students have to take one seminar course, for which only English is used as a medium of instruction. Students have to read many journal articles and listen to academic presentations in English.

According to the CEFR descriptors (see Figure 1.1), being able to understand the main ideas of complex texts in the field of specialization is the description of the B2 users, but the science students are only equipped with the B1 ability. Several students, consequently, cannot perform well when they read the specialized articles for the seminar course. This challenge implies that the ESP, *English for Science and Technology*, needs to be reconsidered. The course should be redesigned to overcome the students' difficulties. To solve this problem, the scientific academic word lists can be used as the basis for the new course design.

Unfortunately, existing word lists of science disciplines cannot serve the purpose well. First, the SWL (Coxhead & Hirsh, 2007) seems to cover many sub-disciplines, but applying the SWL to the ESP course might put some science students at Burapha University at a disadvantage. Although the SWL and the BUU's science majors share some sub-disciplines, many majors at BUU are not included in the SWL as shown in Table 1.7. On the contrary, the SWL contains several sub-disciplines that

are not offered as majors in BUU's Faculty of Science. Biber (2006) also explains that the specialized vocabulary in natural science (i.e. biology, chemistry, mathematics, and physics) is different by nature from other academic disciplines, even engineering. This implies that if the SWL is used, many words in the list might become an unnecessary burden of vocabulary learning.

Table 1.7
The comparison between the sub-disciplines in SWL and the BUU's scientific majors.

SWL and BUU	SWL only	BUU only
Biology	Agricultural Science	Applied Physics
Chemistry	Computer Science	Aquatic Science
Mathematics	Ecology	Biochemistry
Physics	Engineering and Technology	Biotechnology
	Geography	Food Chemistry
	Geology	Microbiology
	Horticultural Science	Statistics
	Nursing and Midwifery	
	Sport and Health Science	
	Veterinary and Animal Science	

Second, the SWL was built on the assumption that the students have learned the words in the GSL (West, 1953) and AWL (Coxhead, 2000). In fact, English education in Thailand does not rely on this system. Thai students might be familiar with the GSL words as the list has been applied in many English coursebooks, but the AWL has not been introduced explicitly at the undergraduate level. Using the SWL with the undergraduate students might inadvertently create a wider gap for English learning.

Finally, as the SWL and other academic word lists of science disciplines are not aligned with the CEFR, adopting those frequency word lists might not support the national and institutional education curricula. Instead of moving up to the higher level, the students might be bored with or confused about vocabulary learning because they are learning technical words that are too high for their actual proficiency level. Using the new word list of science aligned with the CEFR could help teachers make decisions about the appropriate words to focus on and allow students to achieve their learning goals.

In conclusion, it is necessary to develop a new corpus-based academic word list of science discipline and it should be aligned with the CEFR. This new word list

will help a teacher design an appropriate syllabus and allow the science students at Burapha University to use it as a guideline for self-study. With the appropriate instructions and support, students will be able to read academic texts more effectively. Additionally, the new word list can also be beneficial for other university students who are majoring in natural science disciplines.

1.7 Research questions

The research questions addressed in the study are:

1.7.1. Which academic words are frequently found in the natural science disciplines but are not among the GSL (West, 1953)?

1.7.2. Which academic words are frequently found in the natural science disciplines and in the AWL (Coxhead, 2000)?

1.7.3. To what extent do the lexical items in the new science academic word list align with the CEFR?

1.7.4. To what extent does the new science academic word list differ from the SWL (Coxhead & Hirsh, 2007) and the EST Coursebook Word List?

1.8 Objectives of the study

The objectives of this study are:

1.8.1. To identify academic words frequently used in the natural science disciplines.

1.8.2. To distinguish academic words which can be found in general academic disciplines and in the natural science disciplines.

1.8.3. To map the CEFR levels to the words in the new science academic word list.

1.8.4. To find differences between the new science academic word list and the SWL (Coxhead & Hirsh, 2007).

1.9 Definition of terms

The definition of key terms used in the study are provided as follows:

1.9.1. **Academic vocabulary** refers to English words that are frequently used in a wide range of academic disciplines. They are not usually used in a general context.

1.9.2. **Academic Word List (AWL)** refers to a corpus-based frequency word list of 570 word families. It has been compiled from written academic texts by Coxhead (2000).

1.9.3. **Corpus** refers to a principled collection of text used for making word lists.

1.9.4. **Natural Science Corpus** refers to a corpus of journal articles (review articles and research articles) related to the 11 sub-disciplines of natural science.

1.9.5. **Natural science disciplines** refer to a branch of science that deals with the study of the physical world. Particularly in this study, natural science disciplines contain 11 branches of science degrees offered by Faculty of Science, Burapha University (i.e. aquatic science, biology, biochemistry, biotechnology, chemistry, food chemistry, mathematics, statistics, microbiology, physics, and applied physics).

1.9.6. **The Common European Framework of Reference (CEFR)** refers to a standard framework of language proficiency developed by the Council of Europe in 2001. The CEFR divides language proficiency into 6 levels, namely A1, A2, B1, B2, C1, and C2.

1.9.7. **Word list** refers to a vocabulary list used for the teaching and learning of a foreign language.

1.9.8 **English News Corpus (ENC)** refers to a corpus of online English news developed by Goldhahn, Eckart, and Quasthoff (2012) and available online at http://wortschatz.uni-leipzig.de/en/download/. This corpus used the text material randomly collected from news websites published during 2005 – 2015. They were selected without considering their content in detail. The ENC is divided into nine sub-corpora and each sub-corpus contains around one million tokens. In the present study, the five latest sub-corpora (2009, 2010, 2013, 2014, and 2015) were used.

1.9.9 **English Essay Corpus (ESC)** refers to a corpus of written essays containing approximately 5,600 samples or 1.3 million running words. The ESC is a part of the International Corpus Network of Asian Learners of English (ICNALE) (Ishikawa, 2013). The samples of the ESC have been classified into CEFR levels according to the participants' scores on standardized tests (e.g. TOEIC, TOEFL, IELTS, and STEP). This corpus is freely available online at http://language.sakura.ne.jp/icnale/ index.html#6.

1.10 Significance of the study

The present study is significant for the following reasons:

1.10.1. As this study aims to establish the Science Academic Word List (SAWL) as a list of core academic words used across several science disciplines, the SAWL will contribute to the vocabulary learning and teaching in the field of science in many aspects. First, ESP teachers, course designers, and material developers of science disciplines can use the SAWL as a guideline for preparing their teaching, syllabi, or materials. Second, due to the fact that the SAWL shows the size and the CEFR levels of specialized vocabulary for EFL science students, the SAWL suggests the vocabulary learning paths from the curriculum perspective. That is, all words in SAWL can be divided according to the CEFR levels and be spread across the undergraduate curriculum. This further implies that the collaboration between English and science teachers is preferred. Third, the SAWL can also be used by test developers to examine the specialized vocabulary learning growth. Finally, the SAWL provides a series of sublists for different purposes, e.g. a frequency-based sublist, a CEFR-mapped sublist, and a collocation list. These will guide the learners to develop the appropriate strategies for learning specialized vocabulary.

1.10.2. There has probably been little or no study to date that assigns the CEFR levels to the items in the corpus-based list of academic words. The present study will shed light on a new method of pedagogical word list development.

1.11 Limitations of the study

This study was limited in the following aspects:

1.11.1. Journal articles were used to build the corpus for this study. Journal articles refer to both research articles and review articles. They were downloaded from online academic journals in scientific fields. The journal titles were recommended by science professors at Burapha University. The articles were limited to the most cited articles and articles published from October to December 2017. Textbooks and other reading materials were not in the scope of the study.

1.11.2 This study focuses on developing a scientific academic word list for undergraduate students at Burapha University. The Faculty of Science has 11 majors (mentioned in section 1.6), mostly related to natural science. Therefore, this study is limited to these nine disciplines. The results from this study may be more applicable to natural science than other scientific sub-disciplines.

CHAPTER 2
REVIEW OF LITERATURE

This chapter is divided into five main parts. The first part deals with vocabulary knowledge for English as a foreign language (EFL) learners: vocabulary size, vocabulary types, and text coverage. The second part discusses word lists for language learning: the history of word lists, Coxhead's (2000) *Academic Word List,* specialized word lists, different ways to identify technical words, and a method of making a word list. The third part focuses on the Common European Framework of References (CEFR): its brief history, the CEFR for English language teachers, the CEFR-based vocabulary profiling, and the CEFR vocabulary tools. The fourth part concerns the English of science disciplines. The last part reviews previous studies that are relevant to the development of specialized academic word lists.

2.1 Vocabulary knowledge for EFL learners

Vocabulary knowledge is defined as the amount of interrelated sub-knowledge of a word consisting of many levels, starting with familiarity with the word and ending with the ability to use it correctly in varied and multiple communicative contexts (Caro & Mendinueta, 2017). Acquiring vocabulary knowledge is an essential component of language learning because vocabulary ties into most language skills, especially writing, reading, and grammar accuracy (Milton, 2009).

Vocabulary knowledge frameworks have been suggested by several experts. For example, Nation (2013) states that the things learners should know about a word include its from, its meaning, and its use. Koda (2005) points out that the nature of vocabulary knowledge comprises knowing word meanings, knowing a word's properties, and leaning words form context. Matsuoka (2012) discusses in relationship to reading proficiency that vocabulary size, vocabulary types (or vocabulary levels), and text coverage can have an effect on vocabulary knowledge and word list development. Since the present study focuses on EFL university student's vocabulary

knowledge for academic reading proficiency, Matsuoka's (2012) framework is taken into account.

2.1.1 Vocabulary size

Vocabulary size, or the number of words, can be used as learning goals when designing a language course. To set the boundaries for vocabulary learning, it is primarily necessary to figure out how many words there are in English. Nation and Waring (1997) reviewed two related studies and estimated that there are around 54,000 word families in English. They suggested that the figure is far beyond the reaches of EFL learners and most native speakers. More recently, Nation (2013) has analyzed the British National Corpus (BNC) and found that there are around 70,000 word families in English. This figure is surely not a reachable learning goal as it is even further beyond.

For the EFL learners who need to study or work in English medium universities or workplaces, it is reasonable to consider how many words native speakers know (Nation & Waring, 1997). In general, it is estimated that educated native speakers of English know about 20,000 word families or about 70,000 individual words (Folse, 2010; Nation, 2013). This figure might be used as a long-term benchmark for EFL learners. In fact, these words are not equally useful for the learners.

Table 2.1
Levels of vocabulary size

Levels	Word Families
1. Over all English words	70,000
2. Native speakers	20,000
3. EFL learners	3,000 – 5,000

It is worth considering the extent to how many words learners need to know. Several frequency-based studies (Laufer & Ravenhorst-Kalovski, 2010; Nation, 2006; Schmitt, Cobb, Horst, & Schmitt, 2015; van Zeeland & Schmitt, 2013; Stuart Webb & Rodgers, 2009) show that a small number of English words occur very frequently. The findings imply that the learners with a limited vocabulary size (between 3,000 – 5,000 word families) will be successful in reading or listening to many texts (Biber, 2006; Nation & Waring, 1997). Based on the mentioned information, the

vocabulary size of English language can be summarized in Table 2.1. However, Matsuoka (2012) argues that, instead of having a large vocabulary size, the EFL learners need to know enough of the 'right' words. Put another way, they should know the necessary vocabulary types used in their specific context.

2.1.2 Vocabulary levels

Corpus-based studies, as mentioned in Chapter 1, shed light on the division of vocabulary based on frequency of occurrence in a corpus or corpora. Nation's (2012) *The BNC/COCA Word Family Lists* are the predominant example of such a division. The BNC/COCA lists are developed based on the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA). The lists consist of 25 lists of one thousand word families based on frequency and range data. According to the notion that not all English words have equal usefulness for learners, learners and teachers should spend considerable time on high-frequency words because these words cover a large proportion of the running words and occur in most kinds of uses (Nation, 2013). In addition, to aid learning and teaching, English vocabulary can be categorized into five levels: high-frequency words, academic words, technical words, mid-frequency words, and low-frequency words.

2.1.2.1 High-frequency words

The high-frequency words are words in the most frequent 2,000 word families and most of the 176 function words in English. They are essential for EFL learners because these words provide a coverage about 80-90% of the running words in most written texts (Matsuoka, 2012). However, Schmitt and Schmitt (2012) argue that high-frequency words should be extended to 3,000 word families because the 3,000 word family level gains 95% coverage in most texts (Nation & Anthony, 2013). The classic collection of these words is Michael West's (1953) *A General Service List of English Words* (GSL), which contains around 2,000 word families of most useful function words and content words. According to Nation and Kyongho (1995), this list is very useful for learning and teaching and it gives a good return for the learners. Generally, the coverage of the GSL is around 75% of running words in non-fiction texts and 90% of running words in fiction texts. Despite several criticisms of the GSL, it is still the highly influential word list for language learning (Gilner, 2011).

2.1.2.2 Academic words

When EFL learners learn the high-frequency words and intend to go on with academic study, their vocabulary learning should be directed to the vocabulary needed for academic study (Coxhead & Nation, 2001). Several studies have investigated vocabulary in academic texts (e.g. Gardner and Davies (2014); Coxhead (2000); Dang, Coxhead, and Webb (2017); Xue and Nation (1984); Ackermann and Chen (2013)) and have found that there are many words that are common, appearing in the texts across different academic disciplines. These words are usually referred to as *academic words*, *academic vocabulary*, *general academic vocabulary*, or *special purses vocabulary*. They are defined as "lexical items [that] occur frequently and uniformly across a wide range of academic materials" (Coxhead, 2000, p. 218) and might be described as "middle frequency words" (Matsuoka, 2012, p. 154). Nation (2013) asserts that the academic words provided up to 9% coverage of the running words in the academic texts.

The academic words can also be used as a useful learning goal because they are very important for learners with academic purposes. With this in mind, English teachers can effectively help the learners. To date, there are some pedagogical lists of academic words. Coxhead's (2000) *Academic Word List* (AWL) is the best-known list. It is strongly corpus-based and contains 570 word families. The list has been widely used for guiding decisions around learning, teaching, and curriculum and materials design (Coxhead, 2016). Another useful academic vocabulary list is Gardner and Davies's (2014) *Academic Vocabulary List*.

2.1.2.3 Technical words

Nation and Kyongho (1995) define technical words as the words that "occur with very high or moderate frequency within a very limited range of texts or just within one text" (p.36). The technical words are usually related to the subject area of the text and are not so common elsewhere. Typically, they provide up to 5% coverage of the running words in the text. It is worth noting that, in some texts, words from the AWL and the GSL might be technical words, making up between 20% and 30% of the running words in a text (Nation, 2013). Consequently, the technical words are sometimes taught along with words from the AWL to satisfy the needs of some EFL university students (Matsuoka, 2012). To identify technical words, technical

dictionaries might be used. According to these dictionaries, the technical words for any particular subject usually consist of about 1,000 words (Coxhead & Nation, 2001). Recently, the corpus approach has been adopted as a more reliable and accurate method of making word lists of technical words (as discussed in Chapter 1). Some technical word lists contain less than 300 word families (e.g. Ward's (2009) *A Basic Engineering Word List*) which make it more feasible for teaching and learning.

2.1.2.4 Mid-frequency words

The words between high-frequency words and low-frequency words are labeled as mid-frequency words. Formerly, the mid-frequency words, low-frequency words, and technical words were classified together in the same category (Nation & Kyongho, 1995). However, Nation's (2006) findings reveal the usefulness of the mid-frequency words in that knowing around 6,000-9,000 words plus proper nouns is needed in order to reach 98% coverage of the text, which is necessary for unassisted comprehension of unsimplified texts. Schmitt and Schmitt (2012) suggest that the mid-frequency words should include the words between the 3,000 word-family level (high-frequency) and the 9,000 word-family level – consisting of around 6,000 word families. As the mid-frequency words are a large group of words that occur rather infrequently, the goal for this group of words is to encourage deliberate learning, not teaching (Nation, 2013). In other words, the teacher's goal is to train learners in the use of strategies to deal with such vocabulary. The learners should begin training in the strategies for dealing with such vocabulary and spend time learning new words.

2.1.2.5 Low-frequency words

The low-frequency words of the language are words that occur very infrequently and cover only a small proportion of any text. They are those beyond the most frequent 9,000 words. It is difficult to estimate how many low-frequency words there are in English because they might be encountered once or twice and then will not be met again for a long time (Coxhead & Nation, 2001; Nation & Anthony, 2013). There are around 100,000 low-frequency word families. They might be proper names, old-fashioned words, very formal words, words belonging to a particular dialect, or foreign words (Nation, 2013).

Table 2.2
Five levels of vocabulary: high-frequency words, academic words, technical words, mid-frequency words, and low-frequency words

Vocabulary Level	Word Family Level (in total)	Nature of the Vocabulary
High-frequency*	1st 1000-3rd 1000 (3,000)	Wide range, very high frequency, essential, general purpose vocabulary
Academic	-	Wide range, high frequency, general academic purpose vocabulary
Technical	-	Narrow range, high frequency (within a particular discipline), specialized academic vocabulary unique to that discipline
Mid-frequency*	4th 1000-9th 1000 (6,000)	Wide range, moderate frequency, general purpose vocabulary
Low-frequency*	10th 1000 on (100,000)	Narrow range, low frequency, some technical vocabulary

** Note: The descriptions of these three levels are adapted from Nation and Anthony (2013).*

2.1.3 Text coverage

Text coverage is useful information about the vocabulary demand of certain texts on the EFL learners (Matsuoka, 2012). For example, according to Nation and Anthony (2013), an English learner needs to understand around 98% (including proper nouns) of the running words in a text for unassisted comprehension. This means the learner should know most words in the text and the density of unknown words is one in every fifty words (or one unknown word per five lines). To ensure that guessing meaning from context and comprehension occur, 2-5% of the running words in the text should be unknown words. In other words, knowledge of 95% of the running words of the text is essential.

Table 2.3
Coverage of the British National Corpus (BNC) by high, mid- and low-frequency word families (Nation & Anthony, 2013, p. 9)

Type of vocabulary	% Coverage
High-frequency words (3,000 word families)	90%
Mid-frequency words (6,000 word families)	5%
Low-frequency words (10th 1000 word-family level on)	1-2%
Other words (Proper nouns, exclamations, transparent compounds, abbreviations)	3-4%
Total	**100%**

Table 2.3 shows the coverage of BNC by three frequency levels. The table implies that knowing high-frequency words (3,000 word families) reaches 90% of text coverage and is not enough for being able to infer the meaning of the unknown words. Schmitt and Schmitt (2012) argue that to reach 95-98% coverage levels, the EFL leaner needs to have the knowledge of high-frequency words, academic words, technical words, and mid-frequency words.

2.2 Word lists and English language learning

Since the early 20th century, word lists have prominently emerged and played an important role in ELT. As computer technology changes over time, new word lists have been published with most of them based on corpus studies. Such corpus-based word lists provide huge benefits to learners, teachers, researchers, material developers, and test writers. In this section, a brief history of word lists and the making and using of some successful word lists are discussed.

2.2.1 History of word lists

Matsuoka (2012) profoundly reviews previous studies generating word lists. She reveals that the oldest word lists of English language can be traced back to the 16th century as glossaries of words for foreign language learning to explain either useful but difficult words or simple and common words such as names of animals, body parts, and professions. However, those word lists were composed without any frequency component.

During the 19th century, the idea that there is a basic core vocabulary for the language emerged. Several word lists made for use by stenographers were published during that period. Fries (1940) revealed that the most significant word list of the 19th century is Friedrich Wilhelm Kaeding's (1898) *Häufigkeitswörterbuch der deutschen Sprache* (Frequency List of the Words of the German Language). The list was compiled from a wide range of sources and frequency was used to determine the usefulness of a word. Furthermore, the *Häufigkeitswörterbuch* established a standard procedure of compiling a word list which has been followed with little variation ever since. Although Kaeding's intention was not for language learning, the *Häufigkeitswörterbuch* established pedagogical implication for many years (R. Jones & Tschirner, 2015).

Like the *Häufigkeitswörterbuch*, several word lists created during the early 20th century were used for particular users such as for stenographers and for the blind (Matsuoka, 2012). However, a series of word lists compiled by Edward L. Thorndike seems to be the word list widely used for language education, although it was intended for use of native speakers of English at school (Nation, 2016). Thorndike's word list series includes *The Teacher's Word Book* (Thorndike, 1921), *A Teacher's Word Book of 20,000 Words* (Thorndike, 1931), and *The Teacher's Word Book of 30,000 Words* (Thorndike & Lorge, 1944). Particularly in *The Teacher Book of 30,000 Words* (Thorndike & Lorge, 1944), the semantic count technique was proposed as a part of word selection criteria. His work made a great contribution to the techniques of compiling a word list because, apart from strict frequency counts, he decided to include the parameter of range and cross-checked against frequency lists compiled by others (Gilner, 2011; Matsuoka, 2012).

Thorndike's work had a great impact on the description of language use by helping teachers decide if a word should be taught at a certain level and how it should be introduced. Nevertheless, as the words in the lists are categorized by word types, Thorndike's word lists have been criticized. According to Matsuoka (2012), listing entries with word types contributes to two main problems. First, the different meanings of a word form cannot be distinguished. Second, it does not recognize that a word may be part of an idiom or multiword expression. Consequently, Palmer (1931) introduced an innovative approach in his word list. Based on Thorndike's data, Palmer

developed innovative principles to identify a core vocabulary by using subjective criteria; i.e. a system of headwords to categorize the entries in his word list (Gilner, 2011; Matsuoka, 2012). Obviously, this approach "distinguished word families (monologemes) from word types (monologs) and multiword units (pliologs)" (Nation, 2016, p. 9). It is worth noting that Palmer's word list was developed particularly for English language teaching. Later, Palmer's principles and Thorndike and Lorge's semantic count were employed by Michael West to create his *General Service List of English Words* (GSL).

West's (1953) GSL is still influential, unrivaled, and more widely used than many recently compiled word lists (Matsuoka, 2012; Nation, 2016; Watson Todd, 2017). The GSL was based on a corpus of five million running words and employed both objective criteria (word frequency and range) and subjective criteria (ease or difficulty of learning, necessity, overlap with words already in the list, style, and intensive and emotional words). Accordingly, the GSL contains approximately 2,000 word families with their frequencies and the relative frequencies of their different senses which are the most useful to English language learners. However, there has been criticism for its age and its size (Nation & Kyongho, 1995). Because the GSL is dated, the lack of up-to-date lexis such as *CD, DVD, television, video,* and *website* has become its major criticism (Matsuoka, 2012). While other scholars have tried to improve the GSL, Gilner and Morales' (2008 cited in Gilner, 2011, p. 72) study showed that "the GSL is neither dated nor lacking sufficient range-possibility". In other words, the GSL works in a manner similar to recent word lists based on modern corpora and adding new word families to the GSL without violating the original objective/subjective criteria is impossible.

During the 1960s-1970s, attention turned toward the field of academic language. Several word lists of general academic words were produced (Gardner & Davies, 2014). A more robust and widely used word list was Xue and Nation's (1984) *University Word List* (UWL). The list was a combination of four academic word lists created in the 1970s. The UWL gained considerable traction in language education and research for about 15 years despite the fact that the four lists were based on small corpora of academic materials and used different selection principles (Gardner & Davies, 2014; Nation, 2016). As Coxhead (2000) states that the

UWL contains many weaknesses of the prior work, there was a need for a more representative academic list. Hence, Coxhead's (2000) *Academic Word List* (AWL) was published and quickly gained popularity in the field of English language teaching.

The AWL (Coxhead, 2000) contains 570 word families and it is considered a strong corpus-based word list (Nation, 2016). It became the new standard of English language education because, in contrast to the UWL, Coxhead's AWL was created using a more robust methodology behind the list (Gardner & Davies, 2014). Coxhead (2016, p. 181) reveals that her AWL was created using "a written academic corpus, validated on a second academic corpus, and contrasted with general English in another corpus". However, some weaknesses in the AWL have been perceived such as its bias towards fields of law and economics (Matsuoka, 2012) and its focus on receptive skills (Paquot, 2007). Due to these weaknesses, Coxhead's AWL has been challenged by new lists of academic vocabulary such as Gardner and Davies's (2014) *A New Academic Vocabulary List* and Dang et al.'s (2017) *The Academic Spoken Word List*. As the present study also aims to develop a specialized word list, the strong methodology and further criticisms of the AWL will be discussed separately in the following section.

After Coxhead's (2000) AWL becomes a prominent, durable word list in the field of English language teaching, many specialized word lists for broad faculty areas (e.g. science, commerce, or humanities) and word lists for a specific discipline (e.g. medicine, engineering, or applied linguistics) have been published since the late-2000s. These word lists are usually based on their specific corpora and employ the methodology of developing a word lists from previous well-known work such as West's (1953) GSL or Coxhead's (2000) AWL. Nation (2016) states that such specialized word lists are created based of two reasons: to examine the size and the nature of subject-specific vocabularies, and to make a list that could be used in preparing learners for study in the area. Although they employ the methodology from previous work, there are significant differences in the process. As a result, there can be more than a single word list in each specific field. For example, Nation (2016) reveals that there are at least three word lists in medicine (Hsu, 2013; Quero, 2015; Wang, Liang, & Ge, 2008) and two word lists of engineering (Hsu, 2014; Ward, 2009). Section

2.3 will provide more in-depth discussion of the development and validation of these subject-specific word lists.

In summary, the review of the word list development since the 16th century helps us understand more about principles and methodology behind word list formation. The important criteria having been used are both frequency and range. Subjective criteria such as the ease or difficulty of learning should also be included in the selection principle. Although there are variations in the process, representativeness of corpora seems to be the most important factor when judging the quality and contribution of the word list.

2.2.2 Word list of academic vocabulary

As academic vocabulary knowledge is crucial for academic success, educators and language experts are calling for explicit instructions of academic vocabulary including lists of academic vocabulary (Brezina & Gablasova, 2015). The development of academic vocabulary lists can be traced back to the most influential and widely used word list – West's *General Service List* (GSL) in 1953. In response to the GSL, pioneering scholars attempted to explore academic texts to see words which are not in the list but occur frequently across academic disciplines (Coxhead & Nation, 2001). During the 1970s, according to Gardner and Davies (2014), several word lists of general academic vocabulary were developed based on small corpora of academic materials thanks to the technology at that time. Some of academic-text based word lists were made by Campion (1971), Jean Praninskas (1972), Lynn (1973), and Ghadessy (1979). A more robust academic vocabulary list called the *University Word List* (UWL) was published by Xue and Nation (1984). The developers built the UWL on the previously mentioned word lists and excluded words in the GSL. As a result, the UWL contains over 800 word families and has 8.5% coverage in a specialized corpus of academic texts (Nation & Kyongho, 1995). However, the UWL lacked consistent selection principles because it was made from different word lists. There was the need for a more representative academic vocabulary list. Since 2000, the UWL has been replaced by the new standard word list – Coxhead's (2000) *Academic Word List* (AWL).

Coxhead's AWL consists of 570 words based on a 3.5-million-word corpus of academic English texts across four disciplines: Arts, Science, Law, and Commerce. Each group consists of seven sub-disciplines. The 570 words were chosen based on the criteria that they occurred in all four disciplines, in 22 of the 28 sub-disciplines, and at least a 100 times in total. The words were then compared with a 3.5-million-word of corpus of novels to distinguish the words that were truly academic and were not in the GSL. As a result, Coxhead (2000) claims that the new list provides 10% coverage of the tokens in academic corpus, which is higher than that of the UWL. Until now, Coxhead's AWL has served as an important guideline for vocabulary learning in English language education.

Even though Coxhead's AWL is influential and widely used, the list has been criticized for several issues. Gardner and Davies (2014), for example, point out that there are two most problematic issues, which are the use of word families for initial word counts and the relationship of AWL to GSL. The use of word families has been criticized because members of some word families might not share the same core meaning. Instead, lemmas should be preferred. Besides, the AWL was built on the GSL, which is an old list and contains more general high frequency words. Yet, it is found that 79% of the AWL word families are still in the high frequency words. In other words, the good coverage of the AWL in academic texts is the directly related to the inclusion of high frequency words in the list instead of its academic representativeness. As a result, Gardner and Davies introduced a new *Academic Vocabulary List* (AVL) in 2014. One of the key characteristics of the AVL is that it represents contemporary English. The text coverage of the AVL is reported to be twice as high as the coverage of the AWL, but Nation (2013, p. 297) found that 40% of the top 500 words of the AVL are also in the GSL. This means the AVL includes high frequency words easily known to students (e.g. 'study', 'use', 'group', 'level', 'however'). S. Webb and Nation (2017) suggest that, as the AVL contains about 3,000 academic words, it is too big to be used in a language course. The AVL might be a good resource for researchers rather than for learners or teachers.

The GSL (West, 1953) has also been criticized for a number of problems. In response to them, Brezina and Gablasova's (2015) *New General Service List* (new-GSL) and Browne's (2013) *New General Service List (NGSL)* were

developed and published. However, these word lists are still under debate. For example, the new-GSL (Brezina & Gablasova, 2015) was made from a purely quantitative approach and lemmas were used as a counting unit. Browne (2013) argued that the new-GSL may not be helpful for second language learners because the developing prpocess did not involve decisions from learners and teachers. As a result, the words that are useful for the learners but not frequently occur in the corpora might be ignored. Bronwne's (2013) NGSL has been developed from both quantitative and qualitative menthods. The list has been reviewd by a number of teachers and the words in the NGSL have been adjusted according to the suggestions from those teachers. However, the NGSL uses lemmas as counting units. Accoring to Nation and Webb (2011), word lists using lemmas are more useful for supporting students' productive skills (speaking and writing). For receptive skills, which is in the focus of the present study, the word lists using words families as a counting unit are more helpful. Several specialized word lists published recently (e.g. Coxhea & Demecheleer, 2018) are still built based on the GSL (West, 1953). It is, therefore, employed as a part of the present study.

2.2.3 Specialized word lists

A specialized word list (also called *technical word list*, *field-specific academic vocabulary list*, *discipline-specific academic word list* and *discipline-based lexical repertoires*) refers to the list of academic words that are closely related to particular disciplines (Liu & Han, 2015). Experts have drawn considerable attention to the word lists of specific disciplines because several studies have proved that not all words in the interdisciplinary academic word lists (e.g. Coxhead's AWL) are equally important to learners with highly specific needs. In other words, the usefulness of the AWL varies significantly across disciplines in terms of range, frequency, collocation, and meaning (Lei & Liu, 2016). Yang (2015) also suggests that each specific discipline has its own writing conventions, including the way of explaining experience. It is, therefore, necessary to develop academic vocabulary lists for specific disciplines, which have beneficial effects on language instruction and academic vocabulary research (Liu & Han, 2015; Valipouri & Nassaji, 2013). Nation (2016, pp. 145-146) suggests the value of making a specialized word list as follows:

- help us understand the size of vocabulary of a technical area;
- suggest paths towards dealing with such vocabulary from a curriculum perspective;
- guide in the development of appropriate vocabulary learning strategies;
- help in developing subject matter material for English for Academic Purposes courses;
- help in examining the role of technical vocabulary in specialized texts and its possible effects on comprehension; and,
- help in developing tests of previous topic knowledge.

As the present study aims to investigate the academic vocabulary in scientific disciplines, previous studies of specialized word lists related to science are reviewed.

2.2.3.1 Science Word List (SWL)

Coxhead and Hirsh (2007) employed a corpus linguistics approach to conclude that there is a science-specific vocabulary consisting of words outside the GSL and AWL. They developed the Pilot Science Corpus, which includes 14 subject areas (agricultural science, biology, chemistry, computer science, ecology, engineering and technology, geography, geology, horticultural science, mathematics, nursing and midwifery, physics, sport and health science, and veterinary and animal science). The corpus contains 1,761,380 running words or token. The developers used the Range Program (Heatley, Nation, & Coxhead, 2002) to develop the SWL by excluding the words which appear in the GSL and AWL. Word families were constructed using Paul Nation's 21,000 word families of the British National Corpus (BNC) (Nation & Heatley, 2004). After that, range (at least seven subject areas), frequency (at least 50 times in the corpus), and dispersion (a dispersion factor of at least 35) criteria were applied. Finally, proper nouns, symbols, and abbreviations were removed. The SWL, eventually, consists of 318 word families and covers 3.79% over the running words of the pilot science corpus. The list, then, has been tested against other corpora to validate its usefulness as a true word list of science. Similar to the AWL, the words in the SWL are also divided into six sublists based on frequency. Coxhead and Hirsh (2007) suggest that a teacher working with students preparing to

study science at university should start with the first sublist of the AWL and then go to the first sublist of the SWL because these words have higher coverage percentage than other sublists.

However, the SWL has been developed based on the notion that the learners should acquire these specialized words after they know the first 2,000 frequency words and academic words in the AWL. Learning vocabulary in this manner sounds systematic and sensible, but it might not fit in other institutions where learners need to acquire words related to their field of study from both AWL and SWL in a limited period. Furthermore, the fact that the SWL was drawn from 14 subject areas related to science disciplines makes this word list too broad for natural science students. It is highly likely that many words that are not in their field will be introduced to them. In other words, the words in the list might not be very useful for some science students. A specialized vocabulary-learning shortcut is needed.

2.2.3.2 Chemistry Academic Word List (CAWL)

Valipouri and Nassaji (2013) developed the CAWL for EFL chemistry graduate students. CAWL is the collection of 1,400 academic word families used with high frequency in the field. The word list has been developed based on the Chemistry Research Article Corpus (CRAC), which comprises 1,185 research articles including four million words from four sub-disciplines. To be included in the CAWL, the word must occur 10 times in each sub-discipline and 114 time in all the four sub-disciplines. The words that are too technical, judged by three field experts, were excluded from the list. The researchers used Chung and Nation's (2004) rating scale approach. The list does not exclude words in the GSL and the AWL because the students are suggested to learn general service words, general academic words, and technical words respectively. The analysis revealed that the list shares 327 AWL word families, 683 GSL word families, and 390 non-AWL/non-GSL word families. The 1,400 words in the list cover 81.18% of the CRAC. The researchers claim that, by learning a list of words families in the CAWL, the chemistry graduate students can have sufficient vocabulary knowledge to read academic research articles in the field. However, despite the high coverage of AWL in the corpus, this chemistry word list did not include the validity test against other corpora in the developing process.

2.2.3.3 Microbiology Academic Word List (µAWL)

Boonyos (2014) established the µAWL, which consists of 146 word families based on the microbiology research article corpus (µRAC). The corpus contains 2,620,909 running words from 33 journals in the field. The running words consist of 2,249 word families: 1,691 words in the GSL and 558 words in the AWL. Then, the researcher simply employed the Coxhead's (2000) criteria to include the word into the list, in which the selected words had to occur in at least 17 journals and not less than 422 times over the corpus. The words in the GLS were excluded from the list, but the words in the AWL were still included. However, the researcher has failed to give accurate text coverage percentage of 146 words of the µAWL and the list has not been validated against another corpus. Hence, it is quite difficult for teachers or learners to evaluate the usefulness of the list.

2.2.3.4 Environmental Academic Word List (EAWL)

Liu and Han (2015) established the EAWL as the first academic word list for environmental science. According to their study, the AWL provides narrow coverage of some word families and it is not wholly useful for EFL environmental science learners. The EAWL includes 458 word families, in which the list has been developed based on the environmental science corpus. The corpus contains 862,242 running words from 200 research articles of 10 sub-disciplines. The research articles were published between 2010 and 2013. Before compiling the EAWL, the words that are identified in the first 2,000 most frequent word families in the GSL were eliminated. To include the word in the list, members of a word family had to occur at least 30 times and occur in at least eight of the 10 sub-disciplines. As a result, 458 word families totaling 2,332 words were chosen. The list shares 318 words with the AWL. The EAWL was also compared with the British National Corpus (BNC) as a reference corpus. The analysis revealed that the EAWL, which accounts for 15.4% of running words in the environmental corpus, covers only 1%, 3.2%, and 2.8% of the BNC fiction, magazines, and newspapers respectively. The results confirm that the word families in EAWL occur much more in academic environmental texts than they do in general texts. To test the validity of the word list, two validating corpora were created from two different sources of academic articles in the environmental disciplines. The EAWL and the AWL were used to test their coverage in each corpus. The result of the test shows

that the EAWL covered the two validating corpus better than the AWL. Hence, the EAWL appears to be more preferable than the AWL in environmental science academic learning.

2.2.3.5 Engineering English Word List (EEWL)

Hsu (2014) created the EEWL to cover the lexical shortage of EFL engineering for undergraduate students. Although the EEWL is not directly related to the natural science disciplines, it was included in the review of the present study due to the remarkable development process. First, Hsu measured the vocabulary load of English-medium engineering textbooks and discovered that the enrolling engineering students need to be trained with special vocabulary that will help them with their specialized textbooks. Then the researcher developed the 729-word EEWL for all fields of engineering students to learn within a semester or a year. The vocabulary list has been developed based on a corpus of 100 engineering textbooks across 20 engineering sub-disciplines. The Engineering Textbook Corpus contains approximately 4.57 million running words. Exclamations and spoken interjection have been removed from the corpus. Engineering abbreviations, which occur so frequently in the corpus, were added to the list. The selection criteria include (1) the word families are outside the first 2,000 most frequent words created by Taiwan's Ministry of Education for high school graduates; (2) the members of a word family must appear in at least 95 out of the 100 textbooks across 20 engineering subjects; and, (3) the members of a word family must occur at least 288 times in the corpus. The words in the BNC/COCA 1,000-word-family lists after the first 2,000 were excluded. Finally, 729 words were ultimately chosen to form the EEWL. The EEWL covered 14.3% in the corpus. Among them, there was an overlap of 304 word families within the AWL, which contributed to 7.5% of lexical coverage. It is worth noting that the EEWL included a validity test. This process aimed to confirm whether the word families in the list were truly specific to core engineering textbooks. To do so, the EEWL was tested on a corpus of general English textbooks for EFL freshmen. The analysis revealed that the EEWL only covered 2.22%, which showed that the EEWL was specific to engineering textbooks.

2.2.3.6 Nursing Academic Word List (NAWL)

Yang (2015) developed the NAWL for EFL nursing students to strengthen their academic reading and writing proficiency. Like the EEWL, the NAWL is not directly related to the target of the present study, but it was included in this review because the NAWL was built in order to solve the deficiency of the former specialized word lists of the similar field. The NAWL comprises 676 word families. To create the NAWL, the Nursing Research Articles Corpus (NRAC) was constructed. It contains 1,006,934 running words from 252 nursing research articles published between 1995 and 2011. These research articles were selected from 21 sub-disciplines in nursing. Range, frequency, and family are three criteria of creating the NAWL. Only the words which occurred in at least 11 of the 21 sub-disciplines and occurred at least 33 times in the entire NRAC, were included. The word families included had to be outside the first 2,000 most frequently occurring word families of English in the GSL. Finally, the list was compared with the AWL and the Wang, Liang, and Ge's (2008) Medical Academic Word List (MAWL). The analysis revealed that 378 AWL word families (55.92%) were identified in the NAWL and the overlap between the NAWL and MAWL is 63.46%, with 429 word families. Yang concluded that the nursing students could learn the word families in the MAWL (the old word list), but they might not encounter some words in nursing research articles. The NAWL contributed 13.64% lexical coverage of the whole NARC, which is higher than the 10.64% of the MAWL and the 8.93% of the AWL. Although the test of the word list validity is not clearly mentioned, the comparison of the NAWL with the related corpus, the MAWL, can be considered as the way of validating the nursing word list.

2.2.3.7 Medical Academic Vocabulary List (MAVL)

Lei and Liu (2016) developed the MAVL to better serve the needs of EFL medical learners. The MAVL is remarkable in the sense that it was based on more robust word lists of general service vocabulary and academic vocabulary. Moreover, the concept of *lemma* was used as a unit of counting instead of a word family. Previously, Wang et al. (2008) created the *Medical Academic Word List* (MAWL) based on the Coxhead's (2000) AWL. In Coxhead's method, general high-frequency words (West's GSL or the BNC most frequent words) are excluded. Recently, Gardner and Davies (2014) introduced the new *Academic Vocabulary List* (AVL), in which it

includes general high-frequency words and the word selection criteria are considered more rigorous. Lei and Liu (2016), as a result, claim that using the new method proposed by Gardner and Davies (2014) to develop the specialized academic vocabulary list for medical science would be more representative and valid than the MAWL. The new MAVL was created from two corpora, including a 2.7 million-word corpus of medical academic English (MAEC) and a 3.5 million-word corpus of medical English textbooks (MTEC). The MAEC contains 760 research and review articles from 21 sub-disciplines. The MTEC consists of three volumes of a comprehensive foundation textbook in medicine. A lemma-based word list was developed. It is important to note that Lei and Liu (2016) decided to use lemma, rather than word families, because the lemma-based list shows parts of speech and enables learners to focus only on the items in a word family that are used frequently in the field of study. The total lemmas from the MAEC and MTEC were extracted. Then, the lemmas in the list that were also in the new GSL (Brezina & Gablasova, 2015) were identified and considered whether they had special medical meanings by using medical dictionaries. The final list contains 819 lemmas (444 nouns, 133 verbs, 219 adjectives, and 23 adverbs). To validate and check the representativeness of the MAVL, the coverage of the word list across general, academic, and medical corpora was calculated and compared. The analysis suggested that the MAVL is a list of words that are used in academic English, far more frequently in medical English than in general academic English.

2.2.3.8 ICT Word List

Pugsee, Limgomolvilas, Wudthayagorn, and Janpugdee (2017) proposed a framework for developing a *word list for Information and Communication Technology* (ICT Word List). The study explicitly spells out six steps of generating the specialized word list: electing appropriate articles, preprocessing, counting the frequency, stemming words and grouping words, filtering function words out, and listing only content words. The study, as a result, generated the word list from a corpus of 200 research articles from eight journals, containing 1,838,518 running words. After the corpus was prepared, the selection criteria were established. The words, which occur less than 50 times and in less than 50 articles, were eliminated. Finally, the words have been grouped and function words have been filtered. The final

ICT word list contains 254 word families. The ICT word list has been compared with the AWL (40% of sharing words) and the AVL (60% of sharing words). The researchers find the ICT word list was suitable enough for ICT academic staff, teachers, and students. However, the ICT word list has not been validated against another corpus to prove its specialization and the text coverage has not been reported.

In summary, Table 2.4 presents the overall features of the specialized word list development in comparison to Coxhead's (2000) AWL. It is worth noting that the specialized word lists mentioned earlier are based on corpus studies, except Valipouri and Nassaji's (2013) CAWL in which the rating scale approach was used to complement the corpus approach. In fact, making specialized word lists does not necessarily use the corpus approach. Furthermore, it is only Coxhead and Hirsh's (2007) SWL that excluded words in the GSL and the AWL from the list to reflect the notion that the technical words should be learnt after the words in GSL and AWL are known. Most of the specialized word lists excluded only the GSL words because the words are too general. The words in AWL are kept in the lists for pedagogical purposes. There are additional ways to make word lists, which will be discussed in the following section.

Table 2.4
Comparison of specialized word lists

Features	AWL	SWL	CAWL
Word family	570	318	1,400
Purpose	university students	undergraduate students	chemistry graduate students
Corpus data	journal articles, textbooks, texts	journal articles, practical manuals, textbook chapters, lecture notes, study guides	research articles
Sub disciplines	4 (Arts, Commerce, Law, Science) 28 subject areas	14 subject areas	4 subject arears
Corpus size	3.5 million	1.7 million	4 million
Specialized occurrence	outside the GSL	outside the GSL / AWL	including the GSL/AWL
Range	15 of the 28 subject areas	7 of the 14 subject areas	2 of the 4 subject areas
Uniformity of frequency	10 times in each discipline	A dispersion factor of >35	10 times in each subject areas
Frequency	100 times in the corpus	50 times in the corpus	114 times in the corpus
Corpus Coverage	10.0%	4%	81.18%
Evaluation	Test the coverage on a fiction corpus (1.4%)	Test the coverage on the arts, law, commerce corpora (<1%)	n/a
Validation	n/a	n/a	n/a

Table 2.4 (Cont.)
Comparison of specialized word lists

Features	μAWL	EAWL	EEWL
Word family	146	458	729
Purpose	The first Microbiology word list	AWL is not entirely useful for the learners in the field.	engineering undergraduate students
Corpus data	research articles from 33 journals	research articles	engineering textbooks
Sub disciplines	n/a	10 subject areas	20 subject areas
Corpus size	2.6 million	862,242	4.57 million
Specialized occurrence	outside the GSL	outside the GSL	outside the BNC/COCA 2000
Range	17 of the 33 journals	8 of the 10 subject areas	95 of 100 textbooks
Uniformity of frequency	n/a	n/a	14 in each subject areas
Frequency	422 times in the corpus	30 times in the corpus	288 times in the corpus
Corpus Coverage	n/a	15.43%	14.3%
Evaluation	n/a	Test the coverage on 3 BNC corpora (1-3.8%)	Test the coverage on the General English book corpus (2.22%)
Validation	n/a	Compare the coverage of AWL and EAWL in 2 validating corpora	n/a

Table 2.4 (Cont.)
Comparison of specialized word lists

Features	NAWL	MAVL	ICT-WL
Word family	676	819 (lemmas)	254
Purpose	nursing graduate students	medical students or prospective medical students	ICT teachers and students
Corpus data	research article	journal articles, textbooks	200 journal articles
Sub disciplines	21 subject areas	21 subject areas	n/a
Corpus size	1 million	6.2 million	1,838,518
Specialized occurrence	outside the GSL	excluding the high-frequency words without medical meaning / checking special meaning in 2 medical English dictionaries	outside the GSL / AWL
Range	11 of 21 subject areas	12 of 21 of subject areas	50 of 200 articles
Uniformity of frequency	n/a	Juilland's D value: 0.5.	n/a
Frequency	33 times in the corpus	28.57 times per million words / 50% higher in the academic corpus than in the non-academic corpus (1.5 frequency ratio)	50 times in the corpus
Corpus Coverage	13.64%	80%	n/a
Evaluation	Compare the coverage with other word lists	Compare the coverage across general, academic, and medical corpora	n/a
Validation	n/a	n/a	n/a

2.2.4 Identifying technical words

According to Chung and Nation (2004), there are four approaches to identify technical words in a specialized text for making a word list: using a rating scale, using clues, using a technical dictionary, and using corpora. The reliability degree of these approaches has been compared and Table 2.5 summarizes the findings from the study.

Table 2.5
Summary of Chung and Nation (2004)

Approach	Average accuracy rate	Practicality
1. Rating scale	100%	Laborious to apply
2. Corpus-based	82.7%	Easy to apply
3. Dictionary-based	79.8%	Depending on the dictionaries
4. Clue-based	75.5%	Laborious to apply

The findings show that the corpus-based approach provides a very high degree of accuracy and it is not difficult to apply. This approach seems to be the most practical way to make a specialized word list. The result is also in line with the popularity of this approach as it has been used by aforementioned specialized word lists (e.g. Boonyos, 2014; Coxhead & Hirsh, 2007). The present study will also employ the corpus-based approach as the main method of making the specialized word list.

The rating scale approach and the dictionary-based approach also give high accuracy degrees, but they are not widely used. However, the two approaches have been adapted to supplement the quality of some specialized word lists. The rating scale approach (or expert-judged approach), in which a panel of experts is given four-point Likert scale to judge whether to include a word in the list, was applied in Valipouri and Nassaji's (2013) *Chemistry Academic Word List* (CAWL) and Ackermann and Chen's (2013) *Academic Collocation List* (ACL). In the CAWL, three chemistry professors were consulted to identify the words in the list that were too technical. The judged words were excluded from the final CAWL. In the ACL, six experts from different professional backgrounds were consulted to judge whether to include the words in the final ACL. The experts had to decide if the words should be excluded from or included in the list from a pedagogical point of view. The rating scale approach will also be applied to the current study.

The dictionary-based approach was also employed by Lei and Liu (2016) in the *Medical Academic Vocabulary List* (MAVL). The special meaning criterion was used to make the decision whether or not the general high frequency words should be included in the MAVL. To check the word meaning, two well-known medical English dictionaries were consulted. The present study will also adapt the dictionary-based approach in order to assign the CEFR levels (namely A1, A2, B1, B2, C1, and C2) to each word in the list. Instead of using paper dictionaries, two online vocabulary profile databases will be used as references. More details about these databases will be discussed in section 3 of this chapter.

2.2.5 Making a word list

As the corpus-based approach will be used as the main procedure of the present study, the common word list making method is deliberately discussed in this section. Nation and Webb (2011, pp. 135-144) propose six steps involved in constructing a corpus-based word list.

2.2.5.1. Deciding on the reason for making the list

The word list should have a reason for making it or a research question that the list will be used to answer. For example, in Coxhead's (2000) AWL, the aim was to "make a list of words that are beyond the first 2,000 words of English that would be particularly important for all students wishing to proceed with academic study through English" (p.150). This objective makes the list become a widely focused list of academic words and paves the way for later researchers to create narrower, discipline-focused lists.

2.2.5.2. Deciding in the unit of counting

The unit of counting should be considered. There are two ways of counting words: a word family and a lemma. According to Nation and Webb (2011), a word family (i.e. the inflected and derived forms are closely related in form and meaning to each other) is suitable for receptive use, as in listening and reading. For productive use, as in speaking and writing, a lemma (a base word and its inflected forms that are all the same part of speech) is more accurate. Additionally, Dang and Webb (2016) argue that a lemma should be the unit of counting in a word list aimed at beginners because it is most commonly used both inside and outside the classroom. In

Coxhead's (2000) AWL, as reading academic texts was the primary interest, the word family was chosen as the unit of counting.

2.2.5.3 Choosing or creating a suitable corpus

A major task of creating a word list is to choose or to create a suitable corpus. The needs of people who will benefit from the word list are the main concern. Nowadays, there are many corpora available and building new corpora is easier because many materials are in electronic form. However, building a corpus is not necessarily easy. In Coxhead's (2000) AWL, as there was no well-prepared academic corpus available, a new corpus was needed. As a result, the AWL was based on the four equal-sized, well-balanced faculty divisions – science, humanities, law, and commerce. The corpus comprises 3.6 million tokens, which Nation and Webb (2011) claim to be a good representation of written academic texts students would need to read.

2.2.5.4 Making decisions about what will be counted as words

The forth step of creating a word list is to make decisions about what will be counted as words. The word list will be more closely related to the objectives and support the learning burden principle. Nation and Webb (2011) provide criteria of word selection, in which *transparent compounds*, *proper names*, *non-words and marginal words*, *foreign words*, *abbreviations*, *homonyms and homographs* should be consistently considered. Instead of including them in the same list, it might be more practical to treat some types of words separately. For example, proper nouns were excluded from Coxhead's (2000) AWL.

2.2.5.5 Deciding on the criteria that will be used to order the words in the lists

This step is to decide on the criteria that will be used to order the words in the list. The criteria usually refer to the range and frequency of a word in the corpus. Although the range (the occurrence of the word in almost every text in the corpus) is more important that the frequency, they need to be combined in a sensible way. The criteria employed in Coxhead's (2000) AWL were: (1) the word families should not be in West's (1953) General Service List; (2) they must occur in all four faculty divisions and over half of the 28 discipline divisions; and, (3) they must occur at least ten times in each of the four faculty divisions with a frequency at least 100 in the whole 3,600,000 token corpus.

2.2.5.6 Checking the lists against another corpus

The word list requires another corpus against which to be cross-checked. Running the word list on the independent corpus could show any omissions or unusual inclusions in the list. Coxhead's (2000) AWL was cross-checked against two different corpora – a corpus of fiction texts and a corpus of academic texts. The performance of the AWL on the fiction texts was poor (2% coverage), but it worked well on the academic corpus (10% coverage). Nation and Webb (2011) also state that a good word list should perform well on other corpora of the same nature.

2.3 The Common European Framework of Reference for Languages

The Common European Framework of Reference for Languages (CEFR) has become a global standard framework of English language teaching, learning and assessment. It has unquestionable influence on foreign language learning worldwide (Figueras, 2012). As suggested in Chapter 1, foreign language policies and English education curricula of several countries around the world, including Thailand, have referred to the CEFR. This section reviews literature related to the development of the CEFR, the roles of CEFR in English education, and two CEFR-graded lexical inventories: the English Vocabulary Profile and the Global Scale of English Vocabulary.

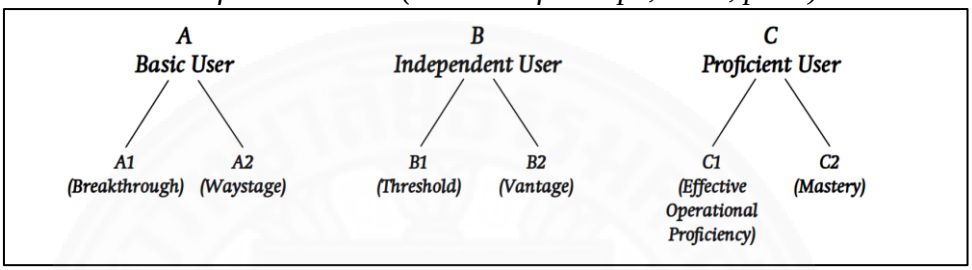
2.3.1 The introduction to the CEFR

The *Common European Framework of Reference for Languages: learning, teaching, assessment* was created by the Council of Europe in 2001. The CEFR aims to provide "a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc. across Europe" (Council of Europe, 2001, p. 1). Originally, the CEFR was developed as a part of the Council of Europe's language policy to promote the learning of several European languages. Nowadays, it is widely used in English language learning and teaching.

The language learners, according to the CEFR, are described and classified into six common reference levels (Figure 2.1). This classification shows learning progression through the levels in relation to cognitive abilities (North, 2014).

The CEFR, furthermore, describes the context of language use, communicative language activities and strategies for language learners, and communicative language competences. Its clear definitions and information have made the CEF successful as a framework for foreign language education in many countries outside European Union countries.

Figure 2.1
The CEFR common reference levels (Council of Europe, 2001, p. 23).



2.3.2 The Roles of CEFR in English Language Education

The CEFR coincides with the rising approach of communicative language teaching (Jones, 2014). It has become a benchmark of English language learning in many countries. Globally, the CEFR has been referred to in English language leaning, teaching, and testing. The CEFR is used in English education for many different purposes (Cambridge University Press, 2013, p. 3):

- developing syllabuses;
- creating tests/exams;
- marking exams;
- evaluating learning needs;
- designing courses;
- developing materials;
- describing language policies;
- continuous/self-assessment; and,
- teacher training programs.

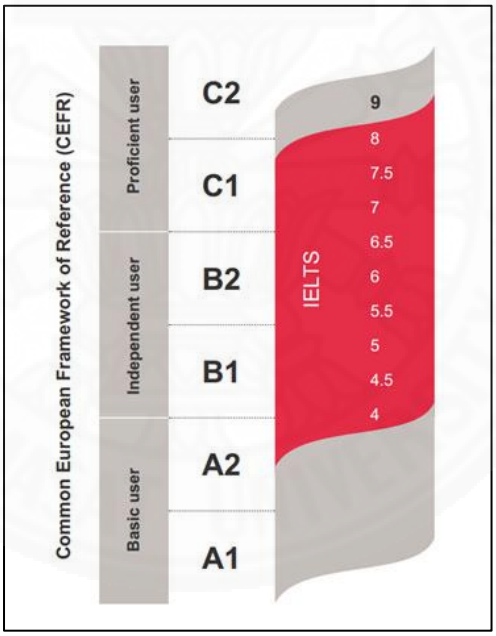
Many research studies towards the applicability and effectiveness of the CEFR have been conducted. For example, to ease the burden on English language

learners, the Association of Language Testers (ALTE) proposes the guided learning hours to progress between levels (Table 2.6).

Table 2.6
Guided learning hours to progress between CEFR levels (Cambridge University Press, 2013, p. 4)

CEFR Level	Guided Learning Hours
C2	Approximately 1,000-1,200
C1	Approximately 700-800
B2	Approximately 500-600
B1	Approximately 350-400
A2	Approximately 180-200
A1	Approximately 90-100

Figure 2.2
The relationship between IELTS bands and the six CEFR levels (IELTS, 2017)



Several studies reflect the significance of the CEFR towards the development of English learning materials and assessment. In material development, coursebooks and language learning materials have been graded with the CEFR levels. Hasselgreen (2011), for instance, reports the process of developing materials for the reading and writing level of young learners and linking the materials to the CEFR. Valldepérez Castillo (2017) found a wide gap between the speaking activities in language coursebooks and the CEFR descriptors. In language assessment, famous

standardized English examinations have aligned their test scores with the CEFR levels. *TOEFL iBT*®, for example, reports that the total minimum score corresponding to B1 is 42 scaled points; the total minimum score to enter B2 is 72 scaled points; and it is 95 scaled points for C1 (Papageorgiou, Tannenbaum, Bridgeman, & Cho, 2015). IELTS also links the IELTS 9 bands to the CEFR based on many studies conducted by Cambridge Assessment English (IELTS, 2017). Figure 2.2 shows the relationship between IELTS bands and the six CEFR levels.

Although the CEFR has become an influential framework to understand performance levels of foreign language learners, it was developed as language-neutral descriptions of levels of language proficiency (Council of Europe, 2001). In other words, the CEFR does not specify any language in particular. As a result, it is difficult to link the CEFR to the knowledge of specific language, i.e. grammar and vocabulary knowledge. The Council of Europe has encouraged the state members to create language specific guidance for users of the CEFR. These inventories of linguistic forms are known as *Reference Level Descriptions* (RLDs). According to Council of Europe (2014), there have been RLDs of ten specific languages finalized and currently being developed: Croatian, Czech, English, French, German, Georgian, Greek, Italian, Portuguese, and Spanish. Especially for the English language, three inventories have been produced in response to the Council of Europe recommendation.

2.3.2.1 The Core Inventory for General English

Based on the original CEFR, *The Core Inventory for General English* (North, Ortega, & Sheehan, 2010) was developed by the British Council and EAQUALS (European Association for Quality Language Services). The project is aimed to embed the CEFR in English language teaching for international use. One of the remarkable outputs of this project is to provide validated, scientifically calibrated descriptors of three additional CEFR levels: A2+, B1+, and B2+. This concept has been adopted by many coursebook publishers and language assessment organizations. Table 2.7 shows that salient characteristics of CEFR levels.

Table 2.7

Salient characteristics of CEFR levels (North et al., 2010, pp. 23-25)

PROFICIENT USER	C2	**Level C2**, labelled, '**Mastery'** is intended to characterise the degree of precision, appropriateness and ease with the language which typifies the speech of those who have been highly successful learners. Descriptors calibrated here include: *convey finer shades of meaning precisely by using, with reasonable accuracy, a wide range of modification devices; has a good command of idiomatic expressions and colloquialisms with awareness of connotative level of meaning; backtrack and restructure around a difficulty so smoothly the interlocutor is hardly aware of it.*
	C1	**Level C1** was labelled **Effective Operational Proficiency**. What seems to characterize this level is good access to a broad range of language, which allows fluent, spontaneous communication, as illustrated by the following examples: *Can express him/herself fluently and spontaneously, almost effortlessly. Has a good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions. There is little obvious searching for expressions or avoidance strategies; only a conceptually difficult subject can hinder a natural, smooth flow of language.* The discourse skills characterising the previous band continue to be evident at Level C1, with an emphasis on more fluency, for example: *select a suitable phrase from a fluent repertoire of discourse functions to preface his remarks in order to get the floor, or to gain time and keep it whilst thinking; produce clear, smoothly-flowing, well-structured speech, showing controlled use of organisational patterns, connectors and cohesive devices*
INDEPENDENT USER	B2+	This band represents a strong Vantage performance. The focus on argument, effective social discourse and on language awareness which appears at B2 continues. However, the focus on argument and social discourse can also be interpreted as a new focus on discourse skills. This new degree of discourse competence shows itself in conversational management (co-operating strategies): *give feedback on and follow up statements and inferences by other speakers and so help the development of the discussion; relate own contribution skilfully to those of other speakers.* It is also apparent in relation to coherence/cohesion: *use a variety of linking words efficiently to mark clearly the relationships between ideas; develop an argument systematically with appropriate highlighting of significant points, and relevant supporting detail*. Finally, it is at this band that there is a concentration of items on negotiating.
	B2	Descriptors calibrated at **Level B2** represent quite a break with the content so far. For example at the lower end of the band there is a focus on effective argument: *account for and sustain his opinions in discussion by providing relevant explanations, arguments and comments; explain a viewpoint on a topical issue giving the advantages and disadvantages of various options; develop an argument giving reasons in support of or against a particular point of view; take an active part in informal discussion in familiar contexts, commenting, putting point of view clearly, evaluating alternative proposals and making and responding to hypotheses.* Secondly, running right through the level there are two new focuses. The first is being able to more than hold your own in social discourse: e.g. *understand in detail what is said to him/her in the standard spoken language even in a noisy environment; initiate discourse, take his turn when appropriate and end conversation when he/she needs to, though he/she may not always do this elegantly; interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without imposing strain on either party.* The second new focus is a new degree of language awareness: *correct mistakes if they have led to misunderstandings; make a note of "favourite mistakes" and consciously monitor speech for it/them; generally correct slips and errors if he/she becomes conscious of them*.

Table 2.7 (Cont.)
Salient characteristics of CEFR levels (North et al., 2010, pp. 23-25)

INDEPENDENT USER (Cont.)	B1+	This band seems to be a strong Threshold performance. The same two main features at B1 continue to be present, with the addition of a number of descriptors which focus on the exchange of quantities of information, for example: *provide concrete information required in an interview/consultation (e.g. describe symptoms to a doctor) but does so with limited precision; explain why something is a problem; summarise and give his or her opinion about a short story, article, talk, discussion interview, or documentary and answer further questions of detail; carry out a prepared interview, checking and confirming information, though he/she may occasionally has to ask for repetition if the other person's response is rapid or extended; describe how to do something, giving detailed instructions; exchange accumulated factual information on familiar routine and non-routine matters within his field with some confidence*.
	B1	**Level B1** reflects the **Threshold Level** specification and is perhaps most categorised by two features. The first feature is the ability to maintain interaction and get across what you want to, in a range of contexts, for example: *generally follow the main points of extended discussion around him/her, provided speech is clearly articulated in standard dialect; express the main point he/she wants to make comprehensibly; keep going comprehensibly, even though pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production*. The second feature is the ability to cope flexibly with problems in everyday life, for example *cope with less routine situations on public transport; deal with most situations likely to arise when making travel arrangements through an agent or when actually travelling; enter unprepared into conversations on familiar topics.*
BASIC USER	A2+	This band represents a **strong Waystage** (A2+) performance. What is noticeable here is more active participation in conversation given some assistance and certain limitations, for example: *understand enough to manage simple, routine exchanges without undue effort; make him/herself understood and exchange ideas and information on familiar topics in predictable everyday situations, provided the other person helps if necessary; deal with everyday situations with predictable content, though he/she will generally have to compromise the message and search for words;* plus significantly more ability to sustain monologues, for example: *express how he/she feels in simple terms; give an extended description of everyday aspects of his environment e.g. people, places, a job or study experience; describe past activities and personal experiences; describe habits and routines; describe plans and arrangements; explain what he/she likes or dislikes about something.*
	A2	**Level A2** appears to reflect the level referred to by the **Waystage** specification. It is at this level that the majority of descriptors stating social functions are to be found, like *use simple everyday polite forms of greeting and address; greet people, ask how they are and react to news; handle very short social exchanges; ask and answer questions about what they do at work and in free time; make and respond to invitations; discuss what to do, where to go and make arrangements to meet; make and accept offers.* Here too are to be found descriptors on getting out and about: the simplified cut-down version of the full set of transactional specifications in "The Threshold Level" for adults living abroad, like: *make simple transactions in shops, post offices or banks; get simple information about travel; use public transport: buses, trains, and taxis, ask for basic information, ask and give directions, and buy tickets; ask for and provide everyday goods and services*.
	A1	Level A1 is the lowest level of generative language use - the point at which the learner can *interact in a simple way, ask and answer simple questions about themselves, where they live, people they know, and things they have, initiate and respond to simple statements in areas of immediate need or on very familiar topics*, rather than relying purely on a very finite rehearsed, lexically organised repertoire of situation-specific phrases.

2.3.2.2 English Profile Programme (EP)

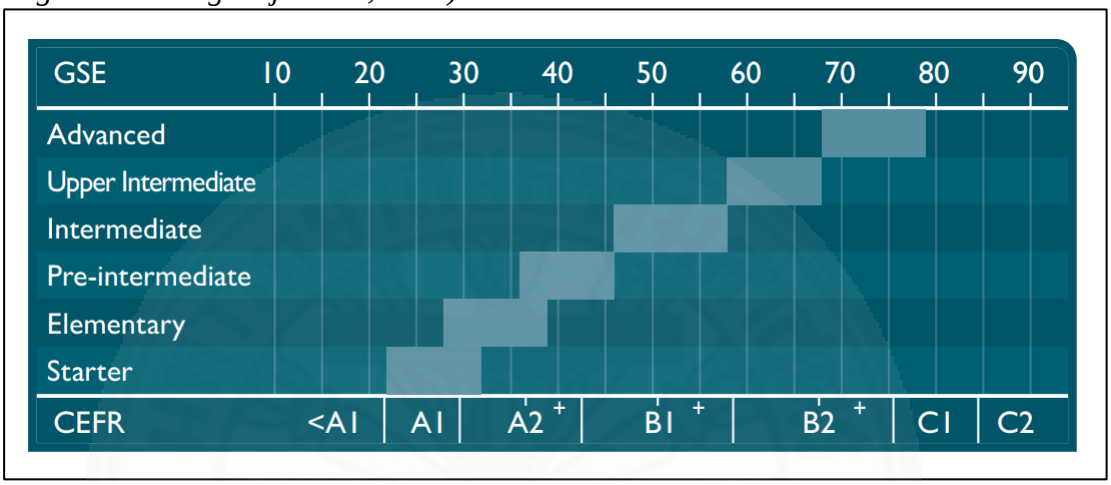
The second English RLD is the *English Profile Programme* (EP), developed by Cambridge ESOL and Cambridge University Press. The main objective of the EP, as stated in its introduction booklet, is "to analyse language produced by learners of English in order to throw light on what they can and can't do with the language at each of the Common European Framework of Reference levels, for example, in using the grammar and lexis at their disposal" (UCLES/CUP, 2011, p. 2). The EP clearly focuses on providing validated descriptions of English across all six levels of the CEFR, in which the grammatical and lexical progression is central to the learning of the language. At present, the EP provides two online tools: English Vocabulary Profile (EVP) and English Grammar Profile (EGP). These online tools are the searchable databases that provide the research findings on the suitable grammar and vocabulary for learning and teaching at each CEFR level. The EP can be reached through http://www.englishprofile.org. As this study will make use of the EVP, this resource tool will be elaborated again in section 3.3.

2.3.2.3 Global Scale of English (GSE)

The third English RLD is the *Global Scale of English* (GSE). The GSE is a psychometric tool developed by Pearson Education. It is the scale of English proficiency from 10-90 aligned with the CEFR (see Figure 2.3) and identifies what a learner can do at each point on the scale across four skills of language: listening, speaking, reading, and writing. The term "global" in the title reflects that this framework is not specific to any region in particular. Instead, it can be a standard for language courses and assessment worldwide (*Introducing the Global Scale of English Learning Objectives*, 2015). The GSE has been widely known and accepted as an innovative framework for national English curricula. In Turkey, for example, several limitations of the CEFR were found and it did not significantly help English language teaching. The GSE has been adopted in the curricula instead (Aydın, Unver, Alan, & Saglam, 2017). Similar to the EP, the GSE also provides *GSE Teacher Toolkit* – an interactive webpage providing research-based information about learning objectives, grammar, and vocabulary for each GSE and CEFR level. It is worth noticing that, for GSE, the CEFR contains nine levels – A1, A2, A2+, B1, B1+, B2, B2+, C1, and C2. The concept proposed by Brian North et al. (2010) in *The Core Inventory for General*

English has been adopted. See https://www.english.com/gse for more information. As the present study will also make use of the GSE Teacher Toolkit, further details will be discussed in Section 3.4.

Figure 2.3
The Global Scale of English aligned to the CEFR (Introducing the Global Scale of English Learning Objectives, 2015)



2.3.3 The English Vocabulary Profile (EVP)

As discussed in the previous section, the EVP is a branch of the English Profile (EP) Programme. The EVP is defined as "an interactive online resource which describes the vocabulary which the learners of English know at each level of the CEFR" (UCLES/CUP, 2011, p. 53). The EVP has aimed to identify which words, phrases, phrasal verbs, and idioms the English learners around the world generally know. It presents the level of each word meaning in CEFR order, which suggests learning priorities. The EVP can be reached through http://www.englishprofile.org/wordlists.

2.3.3.1 The development of the EVP

The EVP has been developed based on corpus-informed research, which are divided into two phrases: A1-B2 vocabulary (Capel, 2010) and C1-C2 vocabulary (Capel, 2012). The first phrase of the research focused on words and phrases known by the learners at CEFR levels A (Basic User) and B (Independent User). Most words and phrases at these levels are derived from lexicographic research, making use of the *Cambridge International Corpus*. This corpus contains a billion

words of written and spoken English. The top 5,000-6,000 words were selected and lexicographers manually assigned one of the three frequency levels to each word: E, I, and A (Essential, Improver, and Advanced), where E represents the highest frequency of occurrence. The words tagged 'E' or 'I' were primarily assumed to fall within the four CEFR levels A1-B2. Then, to capture words that are useful to learners, the *Cambridge Learner Corpus* was analyzed together with word lists from coursebooks and other materials for learners. At this phase, the *Cambridge English Lexicon* (Hindmarsh, 1980) was also consulted because the book has proved a very helpful graded word list for English language teaching. Finally, the pilot version of the A1-B2 word list contains around 4,700 headwords. The pilot version was validated by sending an online questionnaire to the known user groups focusing on the aim and usability.

The second phase dealt with completion of the CEFR levels C1 and C2. At the C levels, General English and Academic English were covered. The analysis focused on 'core vocabulary' and it was based on actual learner evidence, frequency information from first language corpora, and additional sources for Academic English. The remaining less frequency words from the *Cambridge English Lexicon* (Hindmarsh, 1980) were considered for inclusion. Next, all words tagged 'I' and 'A' from the *Cambridge International Corpus* in the previous phase were added to the database. The learner corpora from IELTS data and Cambridge English examinations were analyzed. It is worth noting that the words from the *Academic Word List* (Coxhead, 2000) were also employed to scrutinize against words from the learner corpus evidence before inclusion. As a result, around 2,000 new headwords were added, which make the EVP contain about 7,000 headwords from A1-C2 levels.

2.3.3.2 Key features of the EVP

As mentioned in the earlier sections, the EVP is more than a list of words graded by the CEFR levels. In fact, the EVP contains many features, which can be useful for teachers, leaners, material writers, course designers, and researchers. According to *English Profile: Introducing the CEFR for English* (UCLES/CUP, 2011), the EVP contains the following features:

- containing words, phrases, phrasal verbs, and idioms,
- presenting the CEFR level of each meaning of a word,
- providing detailed dictionary-style entry,
- including audio and written pronunciations,
- containing many real examples,
- having different filters for searching; and,
- containing both American English and British English versions.

Take for example the word *foot*. The following figures show the core results for *foot* in British English and extensive information about some of the results.

Figure 2.4
The core results for "foot" in British English

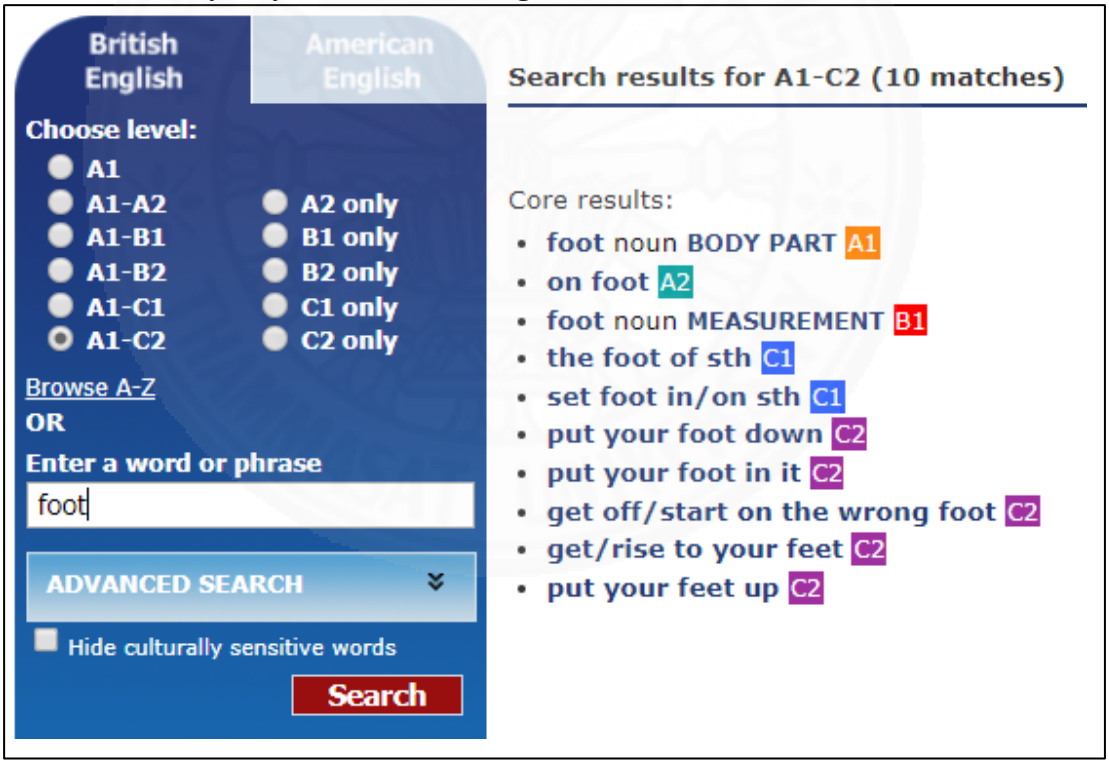
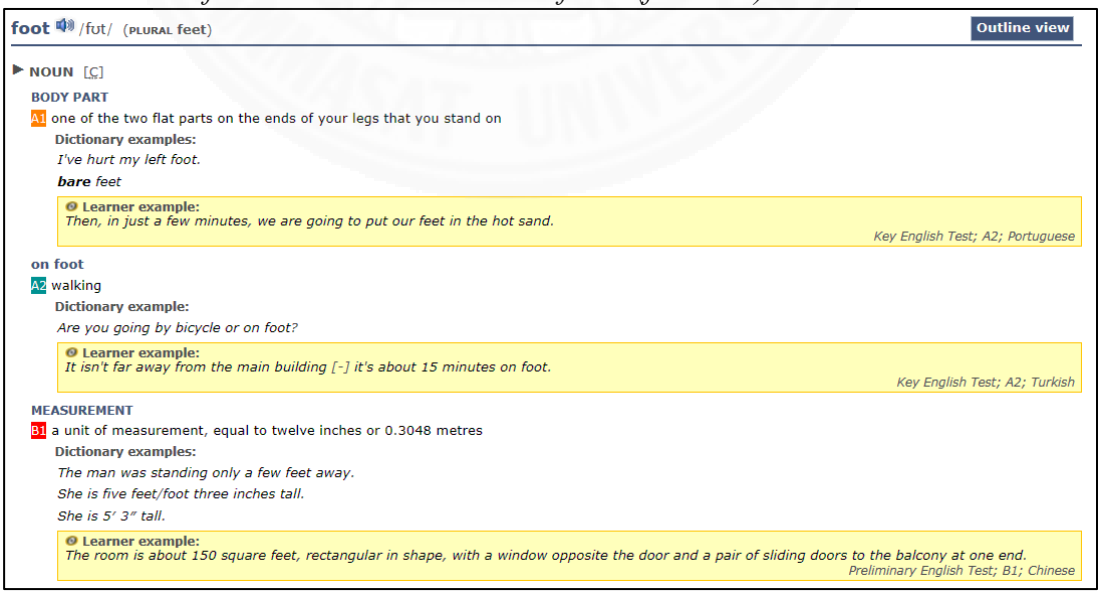


Figure 2.5
The extensive information about the word "foot" (outline view)

Figure 2.6
The extensive information about the word "foot" (full view)



2.3.4 The Global Scale of English (GSE) Vocabulary

The GSE Vocabulary is an ongoing project to develop a graded lexical inventory that aligned the GSE and the CEFR. Similar to the EP Programme, the GSE is also created as Reference Level Descriptions (RLDs) of English. According to Benigno and de Jong (2017), the GSE Vocabulary is "a graded inventory of general English for adult learners which indicates which and how many words meanings learners should be able to understand at different levels to successfully communicate in English" (p.4). The GSE Vocabulary is freely available online as a section in the GSE Teacher Toolkit at: https://www.english.com/gse/teacher-toolkit/user/vocabulary.

2.3.4.1 The development of the GSE Vocabulary

The GSE Vocabulary has been established based on the theoretical framework that vocabulary knowledge and proficiency levels should be related from both quantitative process and qualitative point of view. It is believed that purely frequency-based word lists have some limitations, such as biases of the corpus and the omission of low-frequency words. The GSE Vocabulary was created using frequency data and teacher judgments to identify level-appropriate vocabulary.

Developing the GSE Vocabulary inventory comprised four main steps. First, three corpora were analyzed to produce a word frequency list. The corpora included *Longman Corpus Network*, *UKWAC*, and *A Corpus of Contemporary American English (COCA).* The analysis also included most entries from the *Longman Active Study Dictionary of English*. The frequency word list contained more than 20,000 lemmas. The second step was to identify the meanings of each lemma. The resulting list of around 37,000 word meanings was created. A team of lexicographers manually assigned a topic and sub-topic to each word meaning. The third step involved rating the 37,000 word meanings by 10 raters out of 19 English teachers using an overlapping design. The raters were asked to rate the word meaning using a pre-defined scale on the principle of usefulness (see Table 2.8). The fourth step was to combine information on the frequency words and on the usefulness of words. The data were analyzed using descriptive statistics, linking the word meanings to the GSE and the CEFR. Finally, the GSE Vocabulary contains more than 36,000 word meanings organized by topic, 80,000 collocations, and 70,000 phrases.

Table 2.8
Pre-described rating scale

	Rating Scale
Essential	"Essential" items are the words/phrases that learners would want to acquire first. They are essential for basic communication.
Important	"Important" items are words/phrases that become necessary at a next stage; they are still very common. They are perhaps a little more detailed or a little more specific in their meaning. "Useful" items are words/phrases that expand the user's vocabulary enabling more detailed and specific language use.
Useful	"Useful" items are words/phrases that expand the user's vocabulary enabling more detailed and specific language use.
Nice to have	"Nice to have" items are for users to express themselves accurately and precisely.
Extra	The "extra" category is for items that some language users will use occasionally, but they are not needed for everyday communication.

2.3.4.2 Key features of the GSE Vocabulary

The GSE Vocabulary is a graded lexical inventory which is available online for free at: https://www.english.com/gse/teacher-toolkit/user/vocabulary. The GSE Vocabulary, according to Benigno and de Jong (2017), has five key features. First, the inventory was developed using a mixed methodology of corpus frequency analysis and teacher judgement on communicative usefulness. Second, the inventory presents each lexical entry as a word meaning, not a lemma or a word family. Third, it includes more than 37,000 word meanings (about 20,000 lemmas), 80,000 collocations and 7,000 phrases. Fourth, it can be searched by keyword, part of speech, topic, subtopic, and proficiency level (on the GSE and the CEFR). Finally, the GSE Vocabulary helps learners, teachers, and material designers prioritize vocabulary.

To compare the search results with the EVP, the word *foot* was taken as an example. Figure 2.7 shows the online search results for *foot* to teach adult learners. The database allows users to download the search results with more extensive information, as presented in Figure 2.8.

Figure 2.7
The search results of the word "foot"

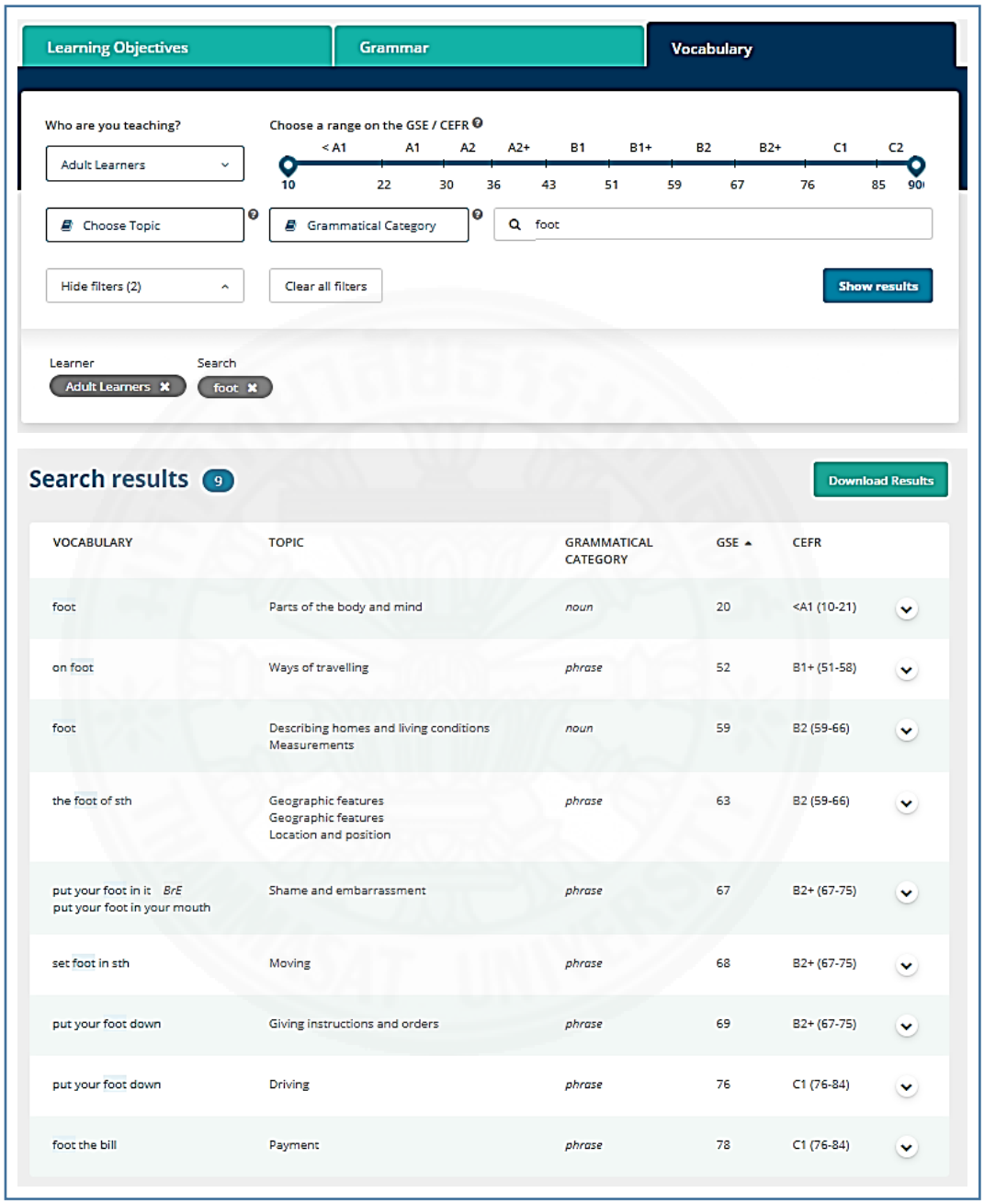
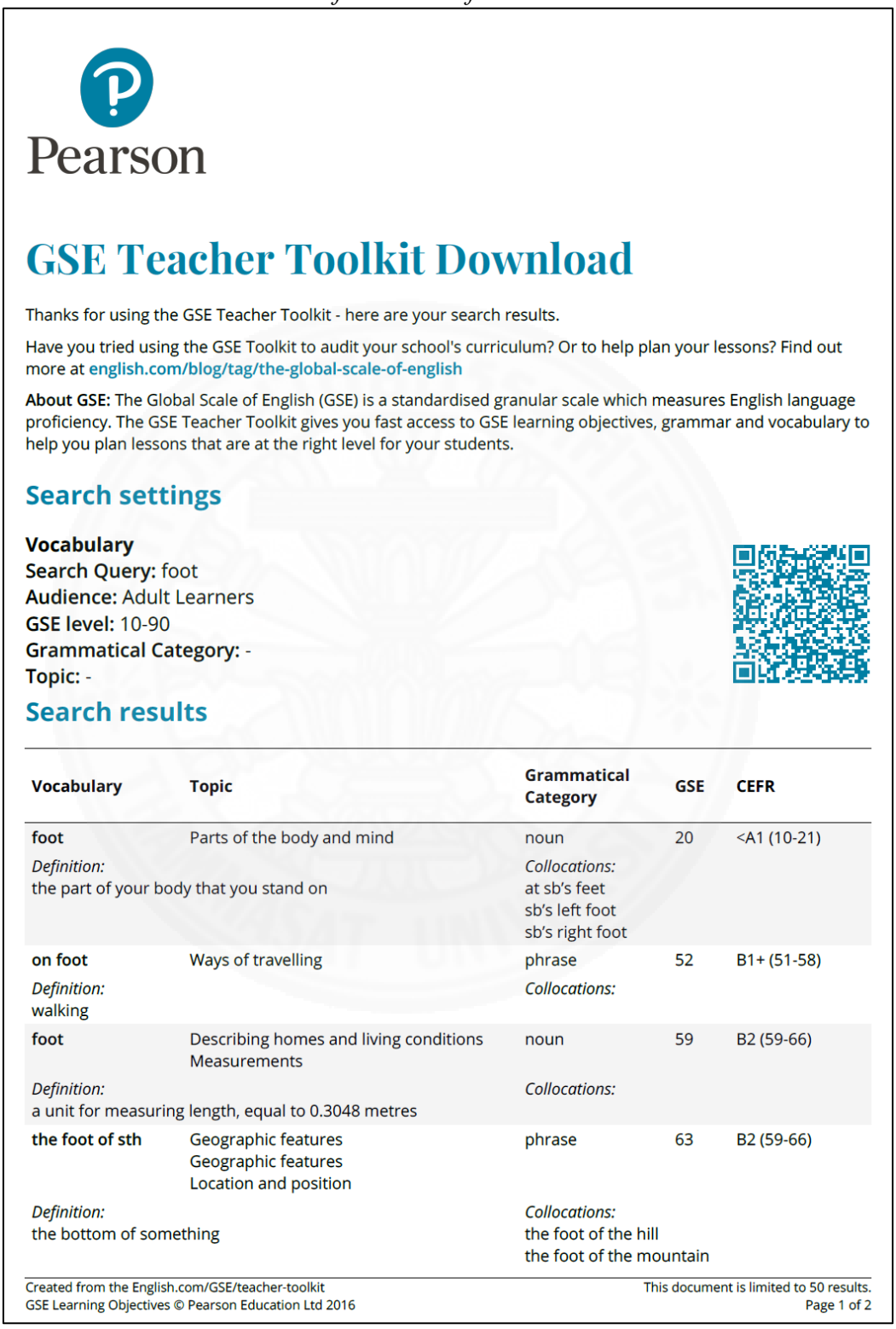


Figure 2.8
The downloaded search results of the word "foot"



Pearson

GSE Teacher Toolkit Download

Thanks for using the GSE Teacher Toolkit - here are your search results.

Have you tried using the GSE Toolkit to audit your school's curriculum? Or to help plan your lessons? Find out more at english.com/blog/tag/the-global-scale-of-english

About GSE: The Global Scale of English (GSE) is a standardised granular scale which measures English language proficiency. The GSE Teacher Toolkit gives you fast access to GSE learning objectives, grammar and vocabulary to help you plan lessons that are at the right level for your students.

Search settings

Vocabulary
Search Query: foot
Audience: Adult Learners
GSE level: 10-90
Grammatical Category: -
Topic: -

Search results

Vocabulary	Topic	Grammatical Category	GSE	CEFR
foot	Parts of the body and mind	noun	20	<A1 (10-21)
Definition: the part of your body that you stand on		*Collocations:* at sb's feet sb's left foot sb's right foot		
on foot	Ways of travelling	phrase	52	B1+ (51-58)
Definition: walking		*Collocations:*		
foot	Describing homes and living conditions Measurements	noun	59	B2 (59-66)
Definition: a unit for measuring length, equal to 0.3048 metres		*Collocations:*		
the foot of sth	Geographic features Geographic features Location and position	phrase	63	B2 (59-66)
Definition: the bottom of something		*Collocations:* the foot of the hill the foot of the mountain		

Created from the English.com/GSE/teacher-toolkit
GSE Learning Objectives © Pearson Education Ltd 2016

This document is limited to 50 results.
Page 1 of 2

2.4 English as a language of science

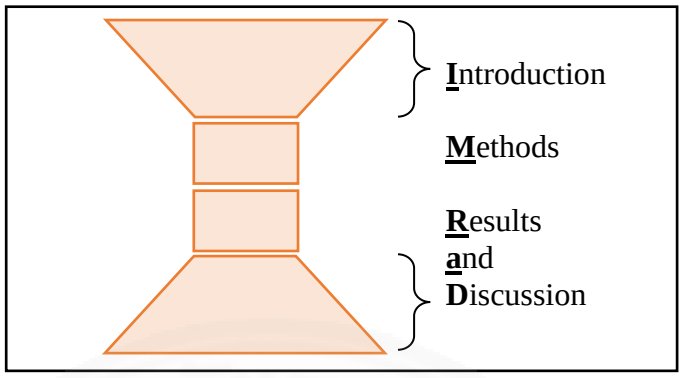
As discussed in Chapter 1, science has great impact on the country's social and economic development. Thai government aims to have well-equipped graduates in science disciplines to serve the development of the country. Those graduates must have a strong knowledge of their disciplines and communicate fluently in English because English has been accepted as *the international language of science* (Wood, 2001). According to Sapir-Whorf hypothesis which explains the relation between language, thought, and culture (Kramsch, 2004), being able to use English language makes a considerable impact on being successful in the scientific community. This section will discuss the characteristics of scientific English, the aspects of science vocabulary, and the existing pedagogical lists of science vocabulary.

2.4.1 The characteristics of scientific English

A major task of scientists is to tell someone about their findings in clear, complete, and concise manners. Therefore, they must have a good command of scientific language (Davis, Davis, & Dunagan, 2012). Reeves (2005) describes that scientists must use the language that is free from connotations. They try to avoid emotional or persuasive language and try to employ the most efficient and objective one. Ahmad (2012) also shows that scientific English has stylistic features and the special use of passive voice, nominalization, questions, and hedges in scientific texts make it different from English used in other disciplines.

From a genre analysis point of view, scientific English also has special characteristics. Although there is an increasing number of scientific research articles written by non-native speakers of English, those research articles have followed similar rhetoric structures. According to Wood (2001), the IMRAD (Introduction-Methods-Results-and-Discussion) format has become the conventional way of organizing scientific communication whether it is for journal articles, laboratory report, seminar presentations, theses, or dissertations. Davis et al. (2012) explain that the IMRAD is an influential communication device because it is what the readers understand and expect in scientific texts.

Figure 2.9
The IMRAD organization of scientific research articles



Furthermore, genre analysis also sheds light on the rhetorical structures in the research articles. Wood (2001) demonstrates that each section of scientific articles written in English has a definite rhetoric structure such as introduction sections (Swales, 1990), results sections (Thompson, 1993), and discussion sections (Hopkins & Dudley-Evans, 1988). These move structures are used by the authors to advance their argument and to have their claims accepted by the wider scientific community. Apart from scientific rhetoric, science vocabulary also needs special attention in scientific English.

2.4.2 The aspects of science vocabulary

Scientific English is a part of becoming a scientist and it is the property of scientists of any language background (Wood, 2001). Student scientists need to integrate the language into their developing content knowledge. In learning science, the English language plays a large and important role. As English learners, the students also need to acquire science vocabulary in order to meet the academic success.

Employing the contemporary perspectives on types of vocabulary (Templeton, Bear, & Johnston, 2010), the vocabulary which the students encounter can be divided into three classifications: conversational vocabulary (high frequency words), core academic vocabulary, and content-specific vocabulary. As the nature of these types of vocabulary has been discussed earlier in this chapter, this section will focus on learning science vocabulary as content-specific vocabulary.

Table 2.9
The challenges and assets of learning science vocabulary (Short, Vogt, & Echevarria, 2010)

Challenges	Assets
1. Dense	1. Clear delineation of vocabulary
2. Conceptually difficult	2. Build-up of ideas
3. Central to text	3. Concepts can be taught thematically
4. General academic vocabulary	4. Many clear Spanish cognates
5. Not much time for instruction	5. Potential for high levels of engagement

Short et al. (2010) suggest the challenges and assets that exist in learning science vocabulary. According to Table 2.9, the learners still need to learn science vocabulary because it can be found to be very frequent in scientific texts and it often carries important concepts of the texts. General academic words are still important for learning science. Therefore, vocabulary instruction for English learners needs to include general academic vocabulary. Coxhead (2000)'s Academic Word List is still recommended. Polysemous words, or words with multiple meanings, also need specific attention because they are often used and the learners might know only one meaning, but not the others – especially the scientific usage of the terms. Reeves (2005) suggests that even English-speaking scientists still have to take care when using terminology because the polysemous terms can be confusing, misleading, and vague to different people. This suggestion implies that, apart from the corpus-based data, subjective judgment from experts is necessary for making a word list for scientific disciplines.

2.5 Summary of the related studies

As the present study aims to develop an academic word list for undergraduate students in science disciplines, this chapter has reviewed related studies in terms of English vocabulary knowledge, vocabulary list development, the Common European Framework of Reference for Languages, and the nature of scientific English. In this section, the previous reviewed studies will be summarized before explaining the methodology in the following chapter. The aspects to be covered in this section are the word selection criteria and the decision on words with multiple meanings.

2.5.1 The word selection criteria

The research questions or the reasons for making a word list are very crucial as they influence the later steps of word list development. According to the word lists reviewed in section 2 of this chapter, most of them were made for pedagogical purposes. For example, Valipouri and Nassaji's (2013) CAWL and Yang's (2015) NAWL targeted at the graduate students in the fields. Most word lists are also aimed at reflecting the vocabulary use in real contexts. Hence, a corpus-based approach was used. The corpus has some impact on words included in the list.

Another factor that influences the word list is the word selection criteria. When the criteria are loose, there will be too many words in the list. The CAWL is a good example. The list contains 1,400 headwords because the words from GSL and AWL were included into the list. Although the words in the final list were classified into four categories, the total 1,400 headwords are still too large for the graduate students to learn. On the other hand, it does not mean that the word list should contain a small number of words. The Boonyos' (2014) μAWL contains only 146 words. The words in the list are, for instance, *analysis*, *isolate*, *significant*, *sequence*, …, *domain*, *energy*, and *community*. The list does not really reflect the specialized words in the microbiology discipline. This word list has two problems. First, the words inside the AWL were included. As the AWL contains the very high frequency words from general academic disciplines, the high frequency words in the μAWL corpus are mostly the same as the AWL words. Second, only words that occurred 422 times over the corpus were selected. This number is quite tough for the 2.6-million-word corpus. When compared to other word lists, the standard is less fastened.

As discussed in section 2 of this chapter, there are several academic word lists for science disciplines and sub-disciplines. These word lists were developed to serve their own specific purposes. For example, Coxhead and Hirsh's (2007) *Science Word List* (SWL) was targeted at first year undergraduate students in English-speaking countries. Hence, their word list consists of words from broad subject areas of science, ranging from biology to sport and health science. The SWL is also built on the notion that vocabulary should be learned in order. In other words, the learners should be familiar with high frequency words before learning academic words and be familiar with academic words before learning technical words. As a result, their word list was

aimed to identify science-specific words outside the GSL and AWL. This notion has not been used in the several following word lists.

Valipouri and Nassaji (2013), for example, developed the *Chemistry Academic Word List* (CAWL). The CAWL was targeted at Iranian chemistry graduate students. The researchers made the decision to include words with high frequency in their corpus. The words inside the GSL and AWL were also counted. As a result, the word list consists of 1,400 academic word families. Although the words were classified into four categories, the CAWL still contains a large amount of vocabulary. A more compromising method of selecting the words for the scientific word list might be to exclude GSL words and to include AWL words, as used in Boonyos (2014) and Liu and Han (2015). This way, the word list will be concise enough for a learner to study in a short period of time and useful for science students across sub-disciplines. For example, the words in Coxhead's (2000) AWL, based on 3.5-million-word corpus, occurred 100 times in the corpus. The words in Coxhead and Hirsh's (2007) SWL occurred 50 times in the 1.7-million-word corpus. The decision to include or exclude the GSL and AWL and the decision on the frequency rate are the example factors affecting the total words in the word list.

However, the present study tries to challenge this convention by setting the expected total number of words in the list in the first place. Taking into account that the manageable learning rate for EFL learners is around 1,000 word families a year (Nation, 2014, 2015), the present study aims to have around 300-400 word families in the word list. This number of words is considered manageable for preparing the ESL undergraduate students in science disciplines. To reach that number, the GSL words must be excluded. Yet, the AWL should be included because the undergraduate students are rarely familiar with the AWL words. The frequency rate will be adjusted later to meet the expected results. Chapter 3 will present the detailed information about the procedure of the study.

2.5.2 Polysemes in scientific English

A polysemous word is defined as a word that has more than one meaning. The classic example is the word *bank* as in a financial *bank* and in the *bank* beside a river. The word *bank* in each situation has a different meaning. The argument is whether these two different meanings should be counted as one word or two words in a word list.

In scientific language, Reeves (2005) remarked that scientists need to be very careful when dealing with the words that may have other meanings and those different technical meanings are not obviously noticed, such as the word *bank*. Scientists from the different sub-disciplines may define the same word in different ways. For example, the word *homology* in the fields of evolutionary biology and biochemistry have different technical meanings. In evolutionary biology, the word *homology* means similarity between organisms based on genetics, while similarity based on similar adaptation to a common function is called *analogy*. However, biochemists use the word *homology* for both types of similarity. The present study considers Reeves's remarks. Distinguishing general words and specialized words cannot be done solely through a corpus approach, although it is very objective. The polysemous words that have specialized meanings can be differentiated by subjective judgement, i.e. using expert's judgement.

Chung and Nation (2004) claim that using expert's judgement is the most thorough way of identifying specialized words. In spite of its high degree of reliability, this approach is very laborious. The expert's judgement approach is commonly used to fulfill the limitation of the main approach. Ackermann and Chen (2013), for example, used expert's judgement to select English collocations in the corpus-driven *Academic Collocation List* (ACL). The GSE Vocabulary (Benigno & de Jong, 2017) also invited many English teachers to judge the usefulness of the words in the corpus-based word list. The results were used to connect the words to the CEFR and GSE levels. The present study also requires the expert-judged approach to distinguish specialized words because they might contain different general and scientific meanings. This approach, therefore, will be used in the study to complement the corpus approach.

After the polysemous words are detected, there are several ways to deal with them. If necessary, the polysemous words might be pooled together in another list of polysemous words. This will be very useful for low proficiency learners. However, Nation (2016) recommends that polysemous words should not be counted as separate words because they are not exactly separate words. They are difficult to differentiate reliably. Moreover, English learners should be able to see the connection between each polysemous words because "this is the essence of dealing with unfamiliar uses of known words" (p.53). Therefore, the present study will not separate the polysemous words nor make a list of polysemous words.

This section has summarized the important, controversial aspects of word list development. In the following chapter, the procedure of making the *Science Academic Word List* (SAWL) in the present study will be demonstrated.

CHAPTER 3
RESEARCH METHODOLOGY

This study focuses on the development of the Science Academic Word List (SAWL). This word list has three main features: (1) the list contains academic words reflecting 11 sub-disciplines of science; (2) the list share some words from the AWL (Nation, 2000); and, (3) the words in the list are aligned with the CEFR levels. Also, this study aims to compare the new science academic word list with the previous Science Word List (Nation & Hirsh, 2007). To achieve these goals, the methodology was designed to contribute to the fulfillment to research questions and objectives mentioned in Chapter 1. The chapter presents the research methodology used in this study which comprises of three sections: (1) materials, (2) research instrument, and (3) data collection. The pilot study of this research project was conducted prior to the full-scale research. To avoid confusion, only significant suggestions from the pilot study were mentioned in the related sections without revealing its results.

3.1 Materials

The SAWL was designed for the EFL undergraduate students majoring in natural science disciplines at Burapha University. In their programs, the students are assigned to read several academic journal articles written in English. Therefore, the articles from scientific journals were used as the materials for this study. The following steps were applied in order to gain the data.

3.1.1 The lecturers in the 11 scientist majors were asked to recommend six journal titles, the articles of which are frequently assigned to the students. There should be 66 journal titles in total.

3.1.2 Five journal titles from each major were selected based on their impact factor index and the availability of the journals in the university online databases. Each of the 11 majors contains five journal titles or 55 journal titles in total. Table 3.1 shows the selected journal titles.

Figure 3.1
The process of selecting journal articles for creating a corpus

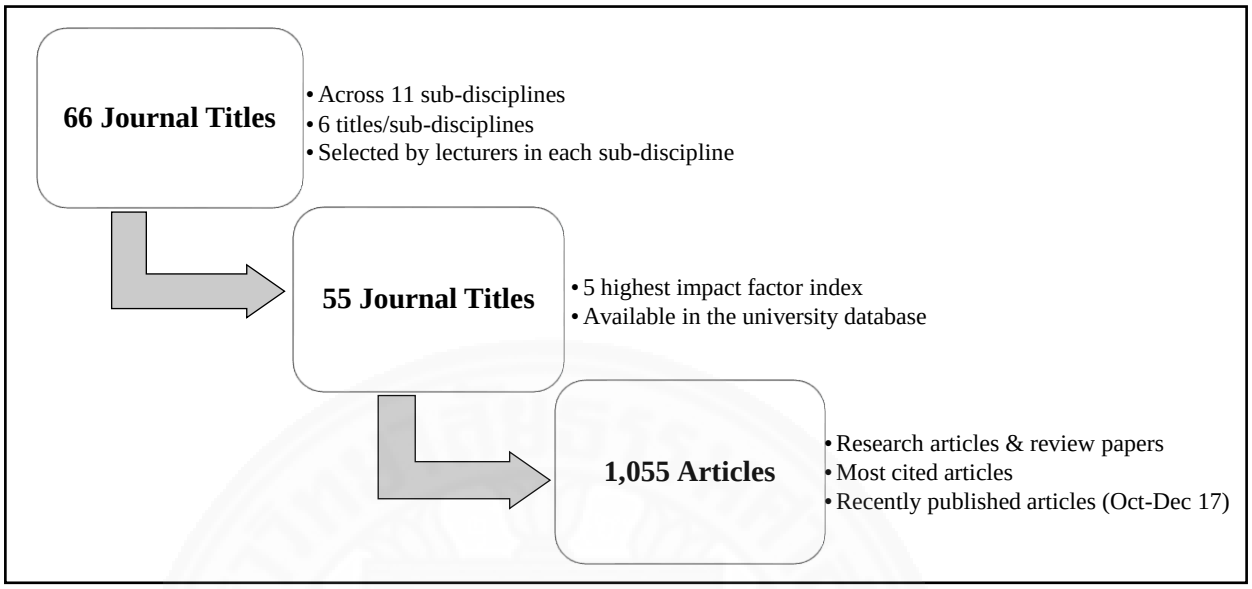


Table 3.1
The selected journal titles for compiling the SAJ corpus

Disciplines	Journals
1. Applied Physics	1.1 Applied Surface Science Journal
	1.2 Journal of Alloys and Compounds
	1.3 Surface and Coatings Technology Journal
	1.4 Thin Solid Films Journal
	1.5 Wear Journal
2. Aquatic science	2.1 Aquaculture Journal
	2.2 Coral Reefs Journal
	2.3 Hydrobiologia Journal
	2.4 Marine Biology Journal
	2.5 Zoological Studies Journal
3. Biochemistry	3.1 Biochemical Journal
	3.2 Biochemistry Journal
	2.3 Journal of Biochemistry
	3.4 Journal of Biological Chemistry
	3.5 PLOS One Journal
4. Biology	4.1 Cell Stem Cell Journal
	4.2 Nature Protocols Journal
	4.3 Nature Reviews Microbiology Journal
	4.4 The FEBS Journal
	4.5 Translational Research Journal

Table 3.1 (Cont.)
The selected journal titles for compiling the SAJ corpus

Disciplines	Journals
5. Biotechnology	5.1 Applied Microbiology and Biotechnology Journal 5.2 Bioresource Technology Journal 5.3 Biotechnology and Bioengineering Journal 5.4 Current Opinion in Biotechnology Journal 5.5 Plant Biotechnology Journal
6. Chemistry	6.1 Analytica Chimica Acta Journal 6.2 Analytical Chemistry Journal 6.3 Journal of Chromatography A 6.4 Talanta Journal 6.5 The Analyst Journal
7. Food Chemistry	7.1 Food Chemistry Journal 7.2 Food Microbiology Journal 7.3 Journal of Food Science 7.4 Journal of the Science of Food and Agriculture 7.5 Meat Science Journal
8. Mathematics	8.1 International Journal of Mathematical Education in Science and Technology 8.2 Mathematics Magazine 8.3 Mathematics Teacher Journal 8.4 The American Mathematical Monthly 8.5 The College Mathematics Journal
9. Microbiology	9.1 Biocontrol Journal 9.2 Biological Control Journal 9.3 Mycologia Journal 9.4 Phytopathology Journal 9.5 Plant Disease Journal
10. Physics	10.1 ACS Nano Journal 10.2 Journal of Computational Physics 10.3 Nature Physics Journal 10.4 Physics of Life Reviews 10.5 Physics Reports
11. Statistics	11.1 Computational Statistics & Data Analysis Journal 11.2 Journal of Statistical Planning and Inference 11.3 Open Journal of Statistics 11.4 Statistics and Computing Journal 11.5 Statistics and Probability Letters Journal

3.1.3 The most cited articles and recently published articles were selected equally to maintain balance between influential work and current work. Several journal websites provide a list of the most cited articles, as shown in Figure 3.1. These articles were usually published before 2017. The recent articles were chosen from those

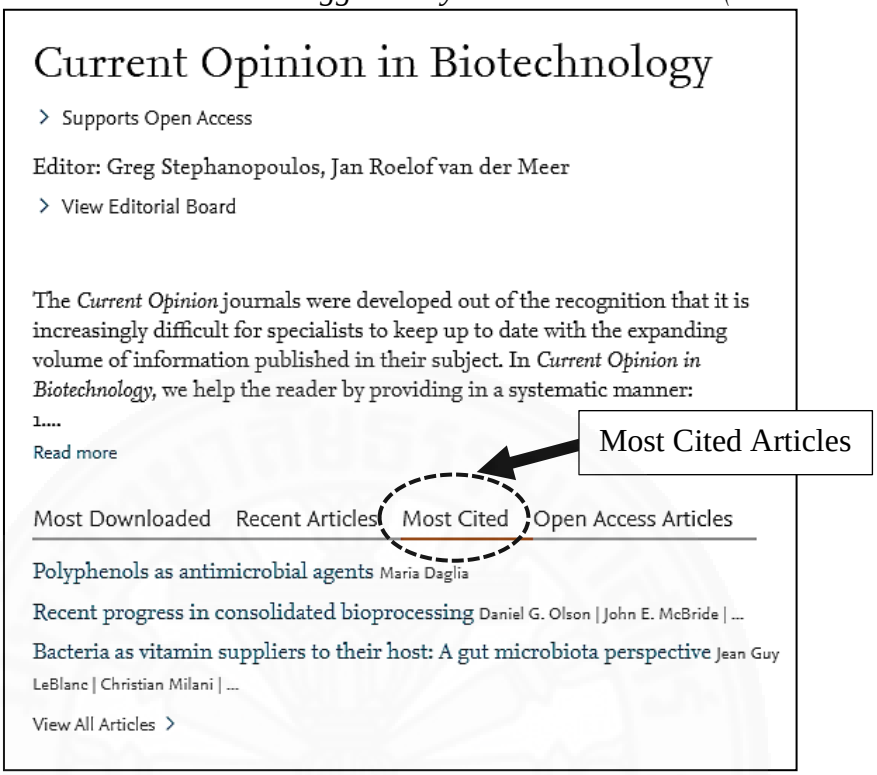
published in October - December 2017, at the time this study was undertaking. As shown in Table 3.2, the target corpus was expected to contain around 5.5 million running words, i.e. 500,000 running words from each of the 11 disciplines. However, the number of articles in each major might be slightly different because each discipline has different writing convention. For example, the average length of articles in aquatic science is around 9,000 running words per article, while the article in applied physics contains only 3,500 running words per article on average. It is worth noting that the review papers and research articles included in the SAWL was derived from the Scientific Academic Journal Corpus (hereafter SAJ corpus) because both of them are important for the study in science disciplines. The accurate figures of journal articles and running words will be presented in Chapter 4, Table 4.1.

3.1.4 All articles were saved into TXT files by using Notepad program. The irrelevant sections in the articles such as acknowledgements, references, appendices, and biography were excluded from the analysis.

Table 3.2
The target size of the Scientific Academic Journal (SAJ) Corpus

Majors/Sub-disciplines	**Journals**	**Articles (Approximate Number)**	**Running Words (Approximate Number)**
1. Applied Physics	5	100	500,000
2. Aquatic science	5	100	500,000
3. Biochemistry	5	100	500,000
4. Biology	5	100	500,000
5. Biotechnology	5	100	500,000
6. Chemistry	5	100	500,000
7. Food Chemistry	5	100	500,000
8. Mathematics	5	100	500,000
9. Microbiology	5	100	500,000
10. Physics	5	100	500,000
11. Statistics	5	100	500,000
Total	55	1,100	5,500,000

Figure 3.2
Examples of most cited articles as suggested by the online database (Science Direct)



3.2 Research instrument

The present study will use five research tools, which are divided into three phases: analyzing the corpus, judging by the experts, and aligning with the CEFR.

3.2.1 Corpus analysis

To analyze the corpus, three computer software programs were used: AntWordProfiler (Anthony, 2014), AntConc (Anthony, 2018), and TagAnt (Anthony, 2015). These programs are comprehensive and freely available software programs for corpus linguistics research. The programs have been developed by Dr. Laurence Anthony, a Professor in the Faculty of Science and Engineering at Waseda University, Japan. In addition, they are recommended by Nation (2016) and are widely used for making many word lists (e.g. Chanasattru & Tangkiengsirisin, 2016; Pugsee, Limgomolvilas, Wudthayagorn, & Janpugdee, 2017). Figures 3.3, 3.4, and 3.5 illustrate the interface of the three software programs.

Figure 3.3
AntWordProfiler (Version 1.4.1)

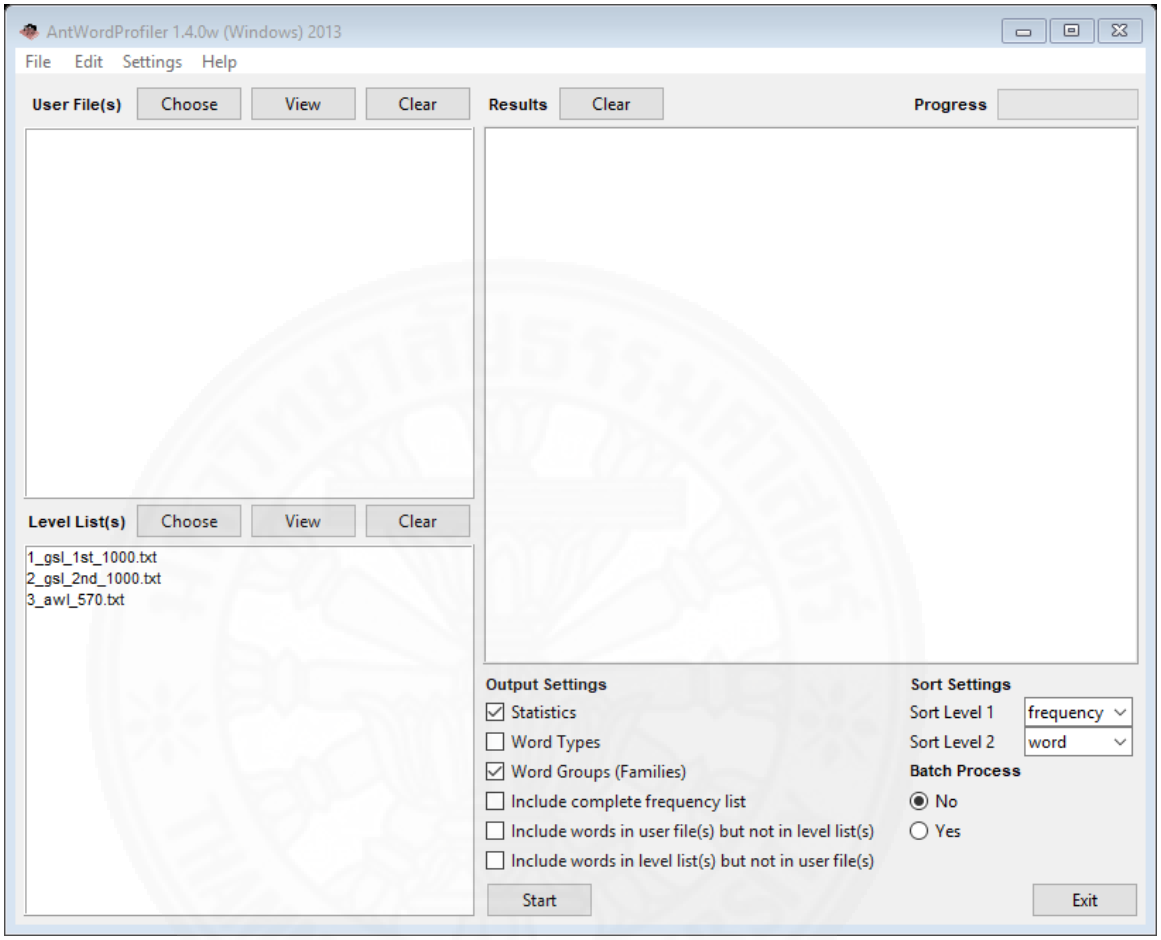


Figure 3.4
AntConc (Version 3.5.7)

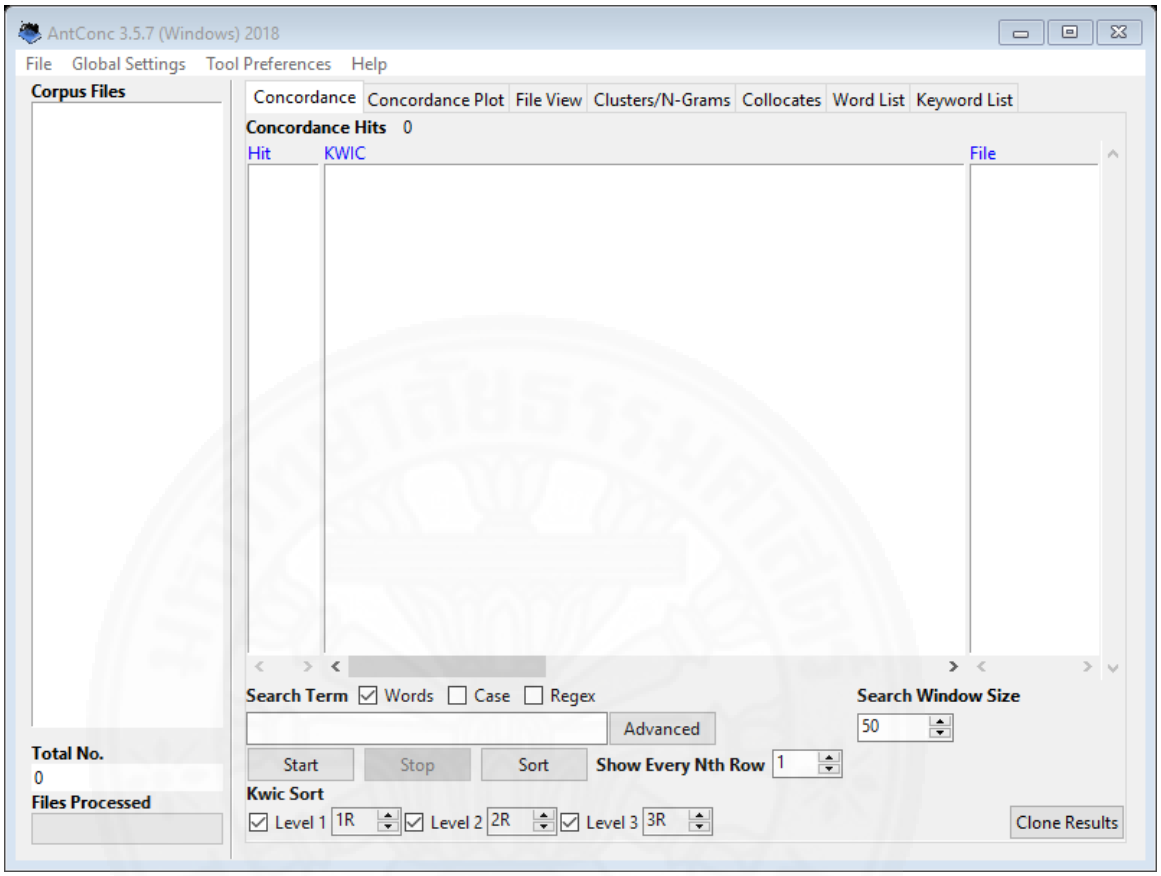
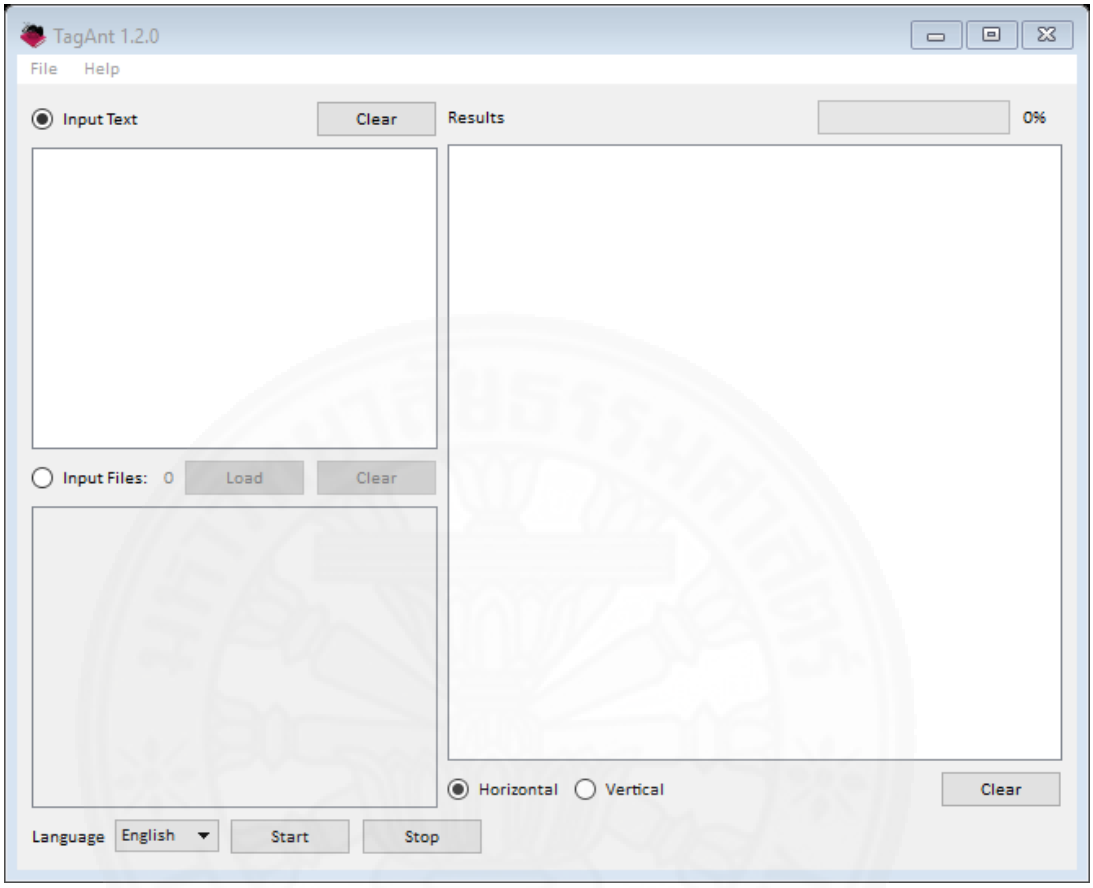


Figure 3.5
TagAnt (Version 1.2.0)



In this study, *AntWordProfiler* was used to generate word lists from the corpus and to compare the lists against reference word lists which are embedded in the program, i.e. West's (1953) GSL 1st 1K, GSL 2nd 1K, and Coxhead's (2000) AWL. In addition, the AntWordProfiler was used to evaluate the SAWL by analyzing its text coverage rate on other corpora. *AntConc* was used to supplement the prior program. AntConc is a concordance program, which can build the corpus and create a frequency word list from the corpus. Its concordance function was used to investigate the use of the key words in context (known as KWIC format). The results from this function were given to the experts for receiving their qualitative judgement and used to align the SAWL words with the CEFR. Its collocation function can be used to find co-occurring words of the target item based on frequency and statistics. In this study, the results from this function were used to establish the list of SAWL collocations. *TagAnt* is a freeware Part-Of-Speech (POS) tagger. It was used to tag POS of all words in the SAJ corpus.

The results were then analyzed with *AntConc* to support the CEFR alignment of the SAWL words.

3.2.2 Expert judgment

According to Chung and Nation (2004), the expert-judge approach is the most reliable method of identifying technical words. As this approach needs a lot of time and effort, some specialized word lists (e.g. Ackermann & Chen, 2013; Valipouri & Nassaji, 2013) employed this approach to enhance the corpus approach. The main tool for the expert-judged approach is a rating scale. The scale used in the present study is the three-point scale adapted from Chung and Nation (2004), in which it is used to measure the strength of the relationship of a word for a particular specialized field. Originally, the scale consists of four levels with words classified at levels 3 and 4 being considered technical words and added into their list (see Figure 3.6). Valipouri and Nassaji (2013) employed the similar scale. They included words rated at levels 1-3 in their Chemistry Academic Word List (CAWL) because the aim was to make a word list for undergraduate students who also need to know general words used in the chemistry discipline. They did not include words at level 4 because these words were considered too technical.

However, the scale has been criticized for being difficult to distinguish between levels (Coxhead, 2018; Schmitt, 2010). The scale has been reduced from four to three levels in recent studies (i.e. Coxhead & Demecheleer, 2018; Tongpoon-Patanasorn, 2018). In the pilot study of this research project, the raters also reflected that, although the level descriptions were clear enough, they were uncertain about levels 1 and 2. Therefore, in the full-scale study, the original 4-level scale was adjusted by removing the first level because all the words in the potential SAWL were considered to have a relationship with the natural science disciplines and the three-level scale is easier for the raters to judge, shown in Figure 3.7. All words classified at levels 1 and 2 were included in the final SAWL. However, words at level 3 were not included because their meanings are too specific.

Figure 3.6
The original 4-point rating scale from Chung and Nation (2004), p.254

Level 1 Words such as function words whose meaning has no particular relationship with the field of anatomy, that is, words independent of the subject matter. Examples are: *the, is, between, it, by, 12, adjacent, amounts, common, commonly, directly, constantly, early,* and *especially*.
Level 2 Words whose meaning is minimally related to the fields of anatomy in that they describe the positions, movements, or features of the body. Examples are: *superior, part, forms, pairs, structures, surrounds, supports, associated, lodges,* and *protects*.
Level 3 Words whose meaning is **closely related** to the field of anatomy. They refer to parts, structure or functions of the body, such as the regions of the body and systems of the body. Such words are also used in general language. The words may have some restrictions of usage depending on the subject field. Examples are: *chest, trunk, neck, abdomen, ribs, breast, cage, cavity, shoulder, girdle, skin, muscles, wall, heart, lungs, organs, liver, bony, abdominal,* and *breathing*. Words in this category may be technical terms in a specific field like anatomy and yet may occur with the same meaning in other fields and not be technical terms in those fields.
Level 4 Words that have a meaning specific to the field of anatomy and are not likely to be known in general language. They refer to structures and functions of the body. These words have clear restrictions of usage depending on the subject fields. Examples are: *thorax, sternum, costal, vertebrae, pectoral, fascia, traches, mammary, periosteum, hematopoietic, pectoralis, viscera, intervertebral, demifacets,* and *pedicle.*

Figure 3.7
The 3-point rating scale for the present study modified from Chung and Nation (2004)

Level 1 Words that have a meaning that is **minimally related** to the 11 disciplines of natural science
Level 2 Words that have a meaning that is **closely related** to the 11 disciplines of natural science. The words are also used in general language, but they may have some restrictions of usage depending on the subject fields.
Level 3 Words that have a meaning specific to the 11 disciplines of natural science and are not likely to be known in general language. The words have clear restrictions of usage depending on the subject fields.

3.2.3 The CEFR alignment

The words in the SAWL were aligned with the CEFR levels. To identify their CEFR levels, the words were measured through two online databases: the English Vocabulary Profile (EVP) and the Global Scale of English (GSE) Vocabulary. As discussed in Chapter 2, these two online databases are well-established online tools for researchers and teachers to measure the CEFR level of English words. The EVP is available at http://www.englishprofile.org/wordlists and the GSE Vocabulary can be found at https://www.english.com/gse/teacher-toolkit/user/vocabulary.

Both EVP and GSE Vocabulary were used to identify the CEFR levels of the SAWL words. This technique is adapted from technical dictionary technique, in which technical dictionaries were used to identify technical words (Chung & Nation, 2004; Coxhead, 2018). The most frequent family members in the SAWL were checked individually against both databases.

Prior to finding the CEFR levels in the databases, the parts of speech of the SAWL words were detected. This process aimed to provide further lexical information to support the CEFR alignment. As mentioned in the previous section, *TagAnt* (Anthony, 2015) was used. With Nation's (2016) concept of learning polysemous words (discussed in Chapter 2) taken into account, the words in the SAWL were tagged with their lower CEFR levels.

3.3 Data collection

The present study aims to develop the SAWL by employing three main approaches: corpus analysis, expert judgement, and CEFR alignment. The first phase of this study used six steps involved in making a corpus-based word list as suggested in Nation and Webb (2011). The second phase applied the expert-judged approach as used in Chung and Nation (2004), Valipouri and Nassaji (2013), Tongpoon-Patanasorn (2018), and Coxhead and Demecheleer (2018). The third phase adapted the technical dictionary method and used the two CEFR online databases.

3.3.1 Phase one: Corpus-based word list development

Nation and Webb's (2011) steps involved in making a word list have been used as a framework of the first phase of this study.

Step 1: Decide on the research question the list will be used to answer, or the reason for making the list

This study aims to develop the SAWL, which is the list of academic words frequently used in journal articles of multi-disciplines in science. The list was also designed to facilitate the EFL undergraduate science students preparing for reading scientific journal articles within one year of study. The list was expected to comprise 400-500 headwords. However, the final SAWL eventually contains more words than expected, but they are still manageable. This point will be discussed later in the implications.

Step 2: Decide on the unit of counting

The SAWL will be used to develop the student's reading skill. According to Nation and Webb (2011), the most suitable counting unit for receptive knowledge is the word family. That is, words with the same base and inflections are considered as members of the same word family (Bauer & Nation, 1993). The word family for *access* is illustrated below:

ACCESS
- ACCESSED
- ACCESSES
- ACCESSING
- ACCESSIBLE
- INACCESSIBLE
- ACCSSIBILITY
- INACCSSIBILITY

Step 3: Choose or create a suitable corpus

The SAWL was derived from the SAJ corpus, which comprises over 5.5 million words of academic written texts from 55 journals across 11 disciplines of natural science. The process of article selections has been explained in the previous section. As a result, the SAJ corpus consists of 11 sub-corpora and each sub-corpus contains approximately 500,000 running words.

Step 4: Make decisions about what will be counted as words

Words, which do not require previous knowledge or can figure out from previous knowledge, were not used as a headword in the SAWL. Based on this notion, the SAWL did not include transparent compounds, proper names, non-words, foreign words, and abbreviations.

Step 5: Decide on the criteria that will be used to order the words in the list

The present study adopted the word selection criteria from the creation of AWL (Coxhead, 2000). According to the AWL, words were selected based on three criteria: specialized occurrence, range, and frequency.

Special occurrence refers to the occurrence of the words in specialized manners. Coxhead (2000) did not include words in West's (1953) GSL. Many specialized academic word lists developed after the AWL also follow this rule and the word list developers insist that the specialized words should not occur in the AWL either. For example, Coxhead and Hirsh's (2007) SWL focuses on words outside the GSL and AWL. However, as mentioned in Chapter 2, this is debatable. Some specialized word lists allow words in the GSL and AWL (e.g. Valipouri & Nassaji, 2013), while other word lists may exclude words in the AWL (e.g. Boonyos, 2014; Liu & Han, 2015; Yang, 2015). For the creation of the SAWL of this study, the words occurring in the GSL were removed, but the words from the AWL were included.

The range of a word refers to the occurrence of the word in each of the sections (or sub-corpora) of the corpus (Nation & Webb, 2011). The AWL was developed from a large corpus divided into four faculty divisions. Each division comprises eight disciplines, or 28 discipline divisions in total (4 divisions x 8 disciplines = 28 discipline divisions). To be included in the AWL, the words have to occur at least 10 times in each of the four faculty divisions and in at least 15 of the 28 discipline divisions (53.6%). However, the corpus for the present study does not have faculty divisions. Instead, it contains 11 discipline divisions. By applying Coxhead's (2000) principle to the present study, the words to be included in the SAWL have to occur in at least six of the 11 discipline divisions (54.5%).

The last condition is that the words should occur with a high frequency rate in the corpus. According to the AWL, each word in the list had to occur with a frequency of at least 100 times in the whole corpus of 3.5 million running words. That means approximately 28.6 times in every one million running words of the corpus. This principle has been adopted in many specialized word lists. For example, Coxhead and Hirsh's (2007) SWL was derived from a 1.7 million-word corpus. The acceptable frequency rate was 50 times in the corpus (28.6 x 1.7 = 48.6). Valipouri and Nassaji's (2013) CAWL was based on a 4 million-word corpus. The words in the list must occur at least 114 times in the corpus (28.6 x 4 = 114.4). Liu and Han's (2015) EAWL was developed from an 0.86 million-word corpus. The frequency rate for EAWL was 30 times in the corpus (28.6 x 0.86 = 24.6). In the present study, the corpus will contain around 5.5 million running words. Hence, the appropriate frequency rate for the SAWL is 155 times in the whole corpus (28.6 x 5.5 = 157.3).

Table 3.3
Summary of Word Selection Criteria of SAWL and other word lists

Criteria	**SAWL**
Corpus size	5.5 million-word corpus
Special occurrence	Excluding West's (1953) GSL Including Coxhead's (2000) AWL
Range	6 of 11 sub-disciplines
Frequency	155 times on the whole corpus

After deciding the word selection criteria, the SAJ corpus was loaded into the AntWordProfiler program. The SAJ corpus comprises 11 text files. Each file contains around 500,000 running words derived from research articles and review papers published in selected scientific academic journals. The list of words was created. Then, the words in the list were refined and compared with West's (1953) GSL and Coxhead's (2000) AWL. In the first stage, the words at Levels 1 and 2 (GSL 1st 1K and 2nd 1K) were removed because they are inside the GSL. Next, the Level 3 and 4 words (AWL and others) were explored. Words like transparent compounds, proper names, non-words, foreign words, and abbreviations were removed from the result. Finally, the words that met all selection criteria were kept and the rest were removed. The potential SAWL, which categorized individual words into word families, was generated based

on this result. At this stage, AntConc program was employed to investigate the KWIC of some words to make a decision whether they should be counted as a word or not.

Step 6: Check the list against another corpus

In the pilot study of this research project, the pilot-SAWL was tested against another corpus before having the experts rate the words in Phase 2. Unfortunately, several words were dismissed from the pilot list at this stage and resulted in inaccurate coverage rate of the pilot list on the pilot corpus. In the full-scale study, the potential SAWL was reviewed by the panel of experts in Phase 2 and adjusted according to the rating results prior to running this cross-checking test (the last step of Phase 1).

The final SAWL was tested against other corpora to verify its good performance. There are three stages at this step. First, the SAWL was checked on the SAJ corpus to see its performance on the corpus the list is made from. As the SAWL combines general academic words and technical words together, it was hypothesized that the SAWL should have an equal or higher text coverage rate than the AWL alone in the SAJ corpus. Next, the SAWL was cross-checked against other corpora to see if there were any unusual inclusions or notable omissions. This process is based on Nation and Webb (2011) that "word lists will usually perform well on the corpus from which they are made, but the real test comes from an independent corpus (p.144)". The SAWL was tested against the English News Corpus (ENC) (Goldhahn, Eckart, & Quasthoff, 2012) available online at http://wortschatz.uni-leipzig.de/en/download/. As mentioned in Chapter 1, this study employed five latest sub-corpora of the ENC. The ENC contains around five million running words in total, which is similar to the size of the SAJ corpus. As the SAWL was created to benefit students who read scientific journal articles, it was hypothesized to have a low text coverage rate in the comparative corpus. Finally, the final list was tested against the pilot-SAJ corpus, an equivalent SAJ corpus with smaller size made for the pilot study, to investigate its usefulness in real journal articles. As this step, research questions 1 and 2 were answered.

To answer research question 4, the final SAWL list was compared against two different word lists of science. First, the SAWL was compared with the SWL (Nation & Hirsh, 2007) to figure out their performances in the natural science disciplines. The two lists were cross-checked against the SAJ corpus to explore

their coverage rate. Next, the SAWL was compared with the English for Science and Technology Coursebook Word List (ESTCWL). This process was added after undertaking the pilot study and it was discovered that the results could not adequately reflect pedagogical implications. The ESTCWL was compiled from five target vocabulary lists of five commercially available English for Science and Technology (EST) coursebooks printed by a global publisher. The SAWL and ESTCWL were cross-checked against the SAJ corpus to explore their pedagogical usefulness.

3.3.2 Phase two: Expert judgement

The purpose of the expert judgement is to review whether the words should be included in the final SAWL from a scientific point of view. As mentioned in Chapter 2, scientists from different sub-disciplines might define the same term in different ways. Some words might be too vague or inappropriate for some subfields.

In the present study, the panel of three experts judged the potential SAWL words to ensure accuracy of the list entries. The panel consisted of three experienced lecturers from the Faculty of Science who volunteered to participate in the study. They hold doctorate degrees in science and have been teaching science subjects for at least five years. A detailed written summary of the scope and objectives of this study were sent to all the panel experts. They also received the questions and rating scale (Figure 3.4), which was modified from Chung and Nation (2004). Each of the experts was asked to make an independent judgement based on the question whether the word was specific to any discipline of natural science. The words were excluded in the SAWL if they were rated too specific by two of the three raters. The inter-rater reliability test (the Kappa statistic) was applied to the analysis.

3.3.3 Phase three: The CEFR alignment

After having the panel of experts review the SAWL, the words in the final SAWL were aligned with the six levels of the CEFR in order to answer research question 3. To identify and validate the levels, the most frequent family members in the SAWL were measured through two online databases: the EVP and the GSE Vocabulary. Both of these databases are freely available, but subscription is required. Meanwhile, the information regarding parts of speech (POS) of these representative words based on the POS-tagged SAJ corpus was used to support the decision. Figures 3.8 and 3.9 show the searching results of the word “abstract”.

This CEFR alignment process has been measured through comparing two different resources. They might yield different results. Hence, the following criteria were applied to this process. It should be mentioned that the criteria had been used in the pilot study and some conditions were adjusted to improve their accuracy.

(1) This study classified words into six CEFR levels: A1, A2, B1, B2, C1, and C2. These six levels are used in both reference databases. Nonetheless, the GSE Vocabulary supplies four more levels (<A1, A2+, B1+, and B2+). This study considers <A1 as A1, A2+ as A2, B1+ as B1, and B2+ as B2. For example, the word "abstract" is classified as B2 in the EVP and as B2+ in the GSE Vocabulary. It will be classified as B2.

(2) If the word has many senses of meaning, it will be tagged with the one frequently occurring in the SAJ corpus. For example, the word "cell" has many senses of meaning such as an organism, a room, and an electrical device. In the SAJ corpus, this word was mostly used as a noun referring to an organism. The CEFR levels of the word "cell" as an organism will be used.

(3) If the word exists in neither resources, it will be tagged as "unlisted". For example, both EVP and GSE Vocabulary do not contain the word "absorption", the word will be tagged as "UL" in the SAWL.

(4) As the EVP and the GSE Vocabulary were built on different resources, they may provide different CEFR levels of the same words. For example, the word "core" is classified at the C2 level in the EVP, but the GSE Vocabulary shows this word at the B2 level. In such a case, the word will be tagged at the lower level (i.e. B2). The purpose of this condition is to maintain the optimum distribution of CEFR levels in the SAWL. If the higher levels are chosen, chances are that too many words in the SAWL will fall into the high CEFR levels or be classified as "unlisted".

To verify the accuracy of CEFR levels of SAWL words, the CEFR-mapped SAWL was divided by level into four sublists, i.e. A, B, C, and UL. They were then tested against two corpora. First, each of the four sublists was tested against the SAJ corpus in order to investigate its coverage rate and then compared with the three other sublists. The purpose of this process is to prove that the words with lower CEFR levels occur more frequently than the higher-level words, which is adopted from Capel (2010).

Figure 3.8

The searching interface of the word "abstract" from the English Vocabulary Profile

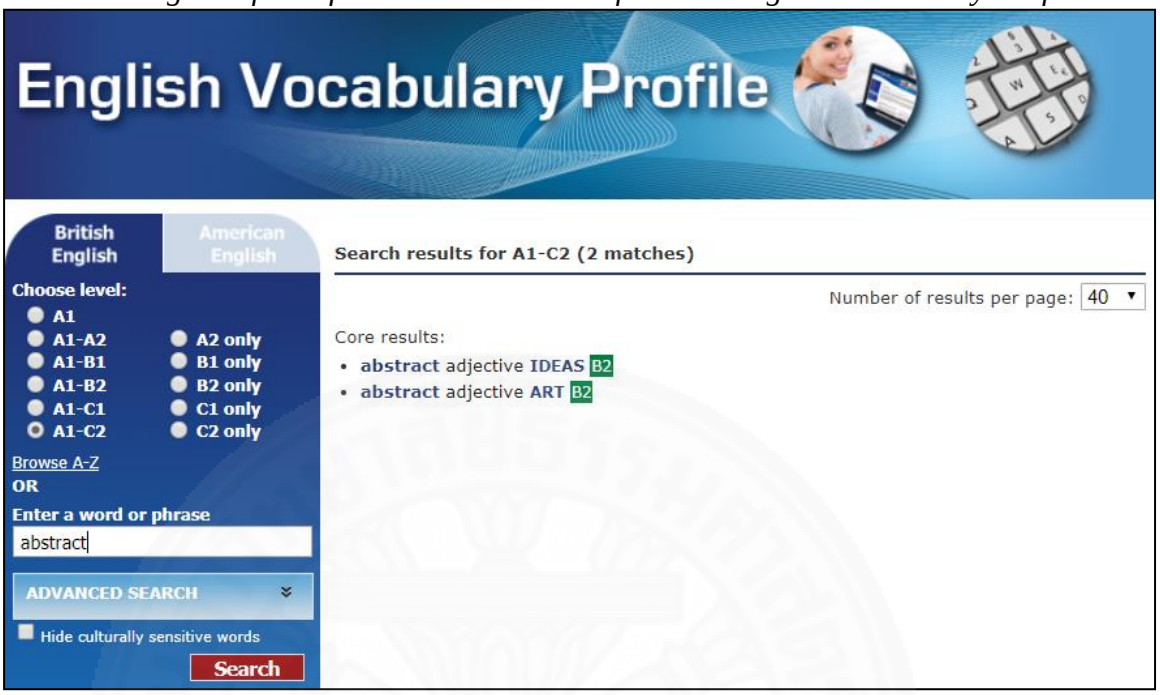
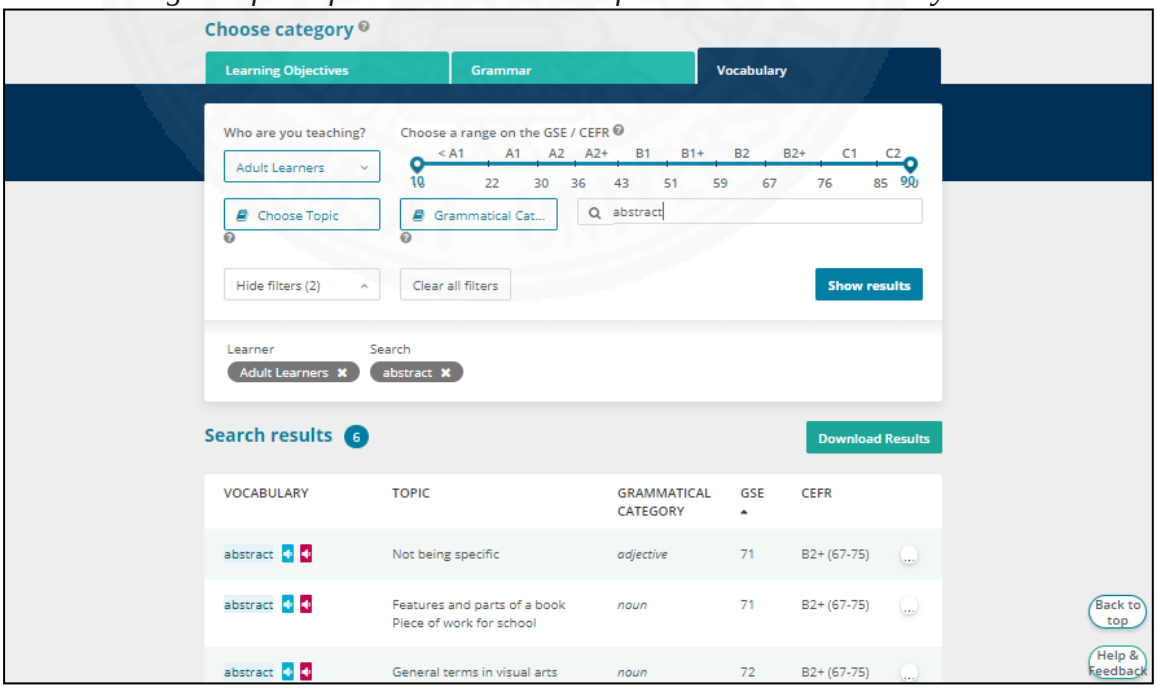


Figure 3.9

The searching interface of the word "abstract" from the GSE Vocabulary



Second, the CEFR-mapped SAWL sublists were tested against the corpus of essays written by learners at different CEFR levels. This process has been added after the pilot study was carried out and yielded that there should be another test to confirm the results. Therefore, in the full-scale study, the CEFR-mapped SAWL sublists were tested against the Written Essay Corpus (WEC) compiled by the International Corpus Network of Asian Learners of English (ICNALE) (Ishikawa, 2013). The WEC contains academic essays written by learners of English at different CEFR levels. This corpus was employed in the present study because it represents English language used by EFL university students, similar to the purpose of making the SAWL. However, the WEC comprises only three levels: A2, B1, and B2 (and above). In this study, the WEC was restructured and divided into two sub-corpora, i.e. CEFR-A and CEFR-B. The SAWL A1 and A2 sublists were tested against the CEFR-A sub-corpus and the SAWL B1 and B2 sublists were checked against the CEFR-B sub-corpus. The purpose of this test is to prove whether the SAWL words are used by the learners at the corresponding CEFR levels. On the other hand, the SAWL C1 and C2 sublists were tested against the CEFR-B sub-corpus with the opposite objective. It is hypothesized that C1 and C2 words are too difficult and should not frequently occur in the CEFR-B corpus. Likewise, the UL sublist was tested against both WEC sub-corpora with the similar objective and hypothesis. The results from this cross-checking with the WEC would support the validity of the CEFR-mapped SAWL.

3.4 Summary

The chapter presents the research methodology used in the present study. The first section shows the process of compiling the SAJ corpus, which is the corpus of journal articles in natural science disciplines. The second section presents research instruments used in each of the three phases of the study: the corpus-based approach, the expert-judged approach, and the CEFR alignment. The last section illustrates the process of analyzing the data and verifying the results. The results together with the discussion of related issues will be presented in Chapter 4.

CHAPTER 4
RESULTS AND DISCUSSION

This chapter reports the results of the development of the corpus-based Science Academic Word List (SAWL) and the alignment of the SAWL words with the Common European Framework of Reference (CEFR) levels. In order to answer the research questions, this chapter is divided into three parts:

1. Development of the SAWL and its validation
2. Alignment of the SAWL with the CEFR levels
3. Comparison of the SAWL, the SWL (Coxhead & Hirsh, 2007), and the EST coursebook word list.

4.1 The development and validation of the SAWL

This section aimed to answer the first two research questions: (1) Which specialized academic words are frequently found the in Science Academic Journal (SAJ) corpus?, and (2) Which academic words are frequently found in the natural science disciplines and on the AWL (Coxhead, 2000)? In this study, the specialized academic words refer to the words that are outside the GSL (West, 1953) but may coincide with the AWL (Coxhead, 2000). The frequently used specialized academic words in natural science journal articles were identified and systematically gathered in order to construct the SAWL. To do so, a corpus-based approach and an expert-judged approach were employed. This section is divided into four parts. The first two parts report the results of the two approaches, while the third part shows the validation process. The last part initiates the discussion on the issues involved.

4.1.1 Corpus-based approach

This study followed the corpus-based word list making process suggested by Nation and Webb (2011). As presented in Chapter 3, the corpus compiled for this study is called the SAJ corpus. This corpus consists of 5.5 million running words representing 11 disciplines of natural science. Table 4.1 illustrates the components in

SAJ corpus and Table 4.2 shows the distribution of GSL, AWL, and other words in the corpus. Eleven disciplines of the SAJ corpus contain approximately 500,000 running words. They were derived from 1,062 research articles and review papers from 55 journals suggested by 11 lecturers from different subfields of science in a Thai university. It should be noted that these natural science disciplines have been normally taught in several universities nationwide. In the SAJ corpus, the GSL contributes 63.36% coverage of the running words. The AWL and other word types cover 36.64% of the corpus. As this study aimed to make a list of specialized academic words that are out of the GSL, the AWL words and other word types found in the SAJ corpus were classified in the following steps.

Table 4.1
The components in SAJ corpus

Subject areas	Articles	Running words
1. Aquatic science	88	508,337
2. Applied Physics	98	507,044
3. Biology	66	508,004
4. Biochemistry	89	501,808
5. Biotechnology	99	505,450
6. Chemistry	89	505,922
7. Food Chemistry	96	503,609
8. Mathematics	146	502,157
9. Microbiology	91	507,532
10. Physics	76	505,310
11. Statistics	124	507,772
Total	**1,062**	**5,562,996**

Table 4.2
The distribution in the SAJ corpus

Levels of words		Running words		Groups	
		No. of running word	Percent	No. of Groups	Percent
1	GSL K-1	3,239,363	58.23%	994	0.95%
2	GSL K-2	285,525	5.13%	898	0.86%
3	AWL	561,119	10.09%	568	0.54%
4	Others	1,476,989	26.55%	102,259	97.65%
	Total	**5,562,996**	**100.00%**	**104,719**	**100.00%**

After compiling the SAJ corpus, the established word selection criteria were applied. All word families will be included in the SAWL if they meet three conditions. First, the words must be outside the GSL. In compliance with the first criterion, from 5.5 million running words in the SAJ corpus, 3.2 million words were removed from the study. Second, the words must occur in at least six of the 11 disciplines. As shown in Table 4.3, the SAJ corpus 102,827 word families (groups) that are outside the GSL and 5,615 word families meet this occurrence condition. Finally, they must occur with a high frequency rate in the SAJ corpus. The appropriate rate for SAWL is 155 times in the corpus. It was found that 899 word families satisfy this criterion. Therefore, the potential SAWL contains 899 word families – 421 AWL words and 478 non-AWL words. Figure 4.1 presents the 100 highest frequency words of potential SAWL together with their range and frequency (see Appendix G for the full potential list). The words coinciding with the AWL are marked with asterisk (*). However, it should be noted that these are the prospective words. Some of them were removed from the final list in the following step.

Table 4.3
Rage levels of the word families outside the GSL in the SAJ corpus

Range Levels (Disciplines)	Word Families
11	873
10	617
9	787
8	871
7	1,050
6	1,417
5	1,912
4	2,874
3	4,893
2	11,848
1	75,685
Total	102,827

Figure 4.1
The potential Science Academic Word List (SAWL) – 100 highest frequency words

Note: *Numbers in parenthesis = (Range / Frequency); * = appearing in AWL (Coxhead, 2000)*

001.cell (11/16056)
002.analyse (11/10162)*
003.protein (11/9058)
004.method (11/8798)*
005.species (11/8798)
006.image (11/8356)*
007.function (11/7973)*
008.data (11/7711)*
009.vary (11/7367)*
010.process (11/7045)*
011.structure (11/5954)*
012.gene (11/5784)
013.significant (11/5486)*
014.obtain (11/5328)*
015.distribute (11/4953)*
016.indicate (11/4784)*
017.specific (11/4715)*
018.similar (11/4666)*
019.sequence (11/4572)*
020.detect (11/4492)*
021.identify (11/4441)*
022.factor (11/4200)*
023.estimate (11/4011)*
024.define (11/3944)*
025.mathematics (11/3928)
026.react (11/3790)*
027.interact (11/3762)*
028.phase (11/3749)*
029.region (11/3710)*
030.concentrate (11/3697)*
031.approach (11/3622)*
032.range (11/3616)*
033.parameter (11/3590)*
034.require (11/3579)*
035.select (11/3564)*
036.potential (11/3503)*
037.extract (11/3493)*
038.molecule (11/3431)
039.respond (11/3429)*
040.matrix (11/3396)
041.site (11/3350)*
042.area (11/3347)*
043.section (11/3321)*
044.energy (11/3284)*
045.previous (11/3281)*
046.target (11/3156)*
047.culture (11/3088)*
048.induce (11/3027)*
049.stable (11/2952)*
050.dense (11/2910)
051.occur (11/2867)*
052.activate (11/2816)
053.regulate (11/2776)*
054.complex (11/2762)*
055.linear (11/2739)
056.correspond (11/2705)*
057.chemical (11/2694)*
058.consist (11/2694)*
059.infect (11/2686)
060.resistance (11/2679)
061.derive (11/2630)*
062.environment (11/2622)*
063.dimension (11/2615)*
064.volume (11/2586)*
065.mechanism (11/2583)*
066.investigate (11/2543)*
067.layer (11/2541)*
068.demonstrate (11/2525)*
069.component (11/2512)*
070.cluster (11/2507)
071.role (11/2504)*
072.tissue (11/2492)
073.isolate (11/2475)*
074.research (11/2450)*
075.domain (11/2444)*
076.assume (11/2441)*
077.ratio (11/2434)*
078.evaluate (11/2423)*
079.approximate (11/2406)*
080.equate (11/2396)*
081.generate (11/2394)*
082.design (11/2389)*
083.involve (11/2370)*
084.positive (11/2365)*
085.individual (11/2357)*
086.compound (11/2344)*
087.strategy (11/2330)*
088.modify (11/2323)*
089.normal (11/2245)*
090.available (11/2239)*
091.pathway (11/2227)
092.affect (11/2226)*
093.cycle (11/2225)*
094.cellular (11/2189)
095.construct (11/2179)*
096.final (11/2178)*
097.peak (11/2175)
098.maximise (11/2143)*
099.vector (11/2134)
100.technique (11/2098)*

4.1.2 Expert-judged approach

The expert-judged approach was used in this study in order to complement the corpus-based approach. The purpose of this process is to use personal judgement of experts to identify words that might be too specific and are not frequently used across disciplines. In this study, a panel of raters was set up and invited to judge the strength of relationship of each prospective SAWL words based on the 3-point rating scale given. The panel consists of three experienced lecturers of natural science disciplines who were invited to participate in the rating process. They received a prospective word-rating sheet with description of the three levels. The raters had about one week to rate the prospective words. The words rated at Level 3 by at least two out of three raters were excluded from the final SAWL. The results from expert-judged approach revealed that, out of 899 prospective word families, 46 word families met the criterion and were not included in the final SAWL. Figure 4.2 shows the rating score of these words..

Figure 4.2
The words excluded from the final SAWL

Words	Level of Word		
	Rater 1	**Rater 2**	**Rater 3**
1.algorithm	2	3	3
2.antibiotics	3	3	3
3.antibody	3	2	3
4.axis	2	3	3
5.beta	2	3	3
6.binary	2	3	3
7.bootstrap	3	3	1
8.cells	3	2	3
9.cluster	2	3	1
10.conjugated	2	3	3
11.convex	2	3	3
12.diagnosis	3	3	2
13.diagnostic	3	3	2
14.diameter	3	3	2
15.dose	3	3	3
16.dotted	2	3	3
17.equilibrium	2	3	3
18.gamma	2	3	3

Figure 4.2 (Cont.)
The words excluded from the final SAWL

Words	Level of Word		
	Rater 1	**Rater 2**	**Rater 3**
19.gradient	2	3	3
20.heterogeneity	3	3	2
21.homogeneous	3	3	2
22.horizontal	2	3	3
23.humidity	2	3	3
24.orthogonal	2	3	3
25.pairwise	2	3	3
26.plasmid	3	2	3
27.quantitative	3	3	2
28.radiation	2	3	3
29.radius	2	3	1
30.recombination	3	3	3
31.resistance	3	3	3
32.sigma	2	3	3
33.situ	3	3	3
34.spectrometer	3	3	3
35.spectrometry	3	3	3
36.spectroscopy	2	3	3
37.spectrum	2	3	3
38.stationary	2	3	1
39.stimulated	3	3	3
40.substrate	3	3	3
41.subunits	3	2	3
42.topology	3	2	3
43.vacuum	2	3	3
44.vector	3	3	3
45.vitro	3	3	3
46.vivo	3	3	3

Apart from technical words, some general science words such as *antibody, axis, cells,* and *horizontal* were also removed from the final SAWL. The rating results illustrated that 41 out of the 46 removed words were suggested by the second rater, while the first and third raters selected 24 and 36 words respectively. It should be noted that the second rater rated five words (i.e. *antibody, cells, plasmid, subunits, topology*) at Level 2, but the other raters chose Level 3. This disagreement illustrates the effect of rater's length of experience in teaching and working. The second rater is a science lecturer who has been teaching for more than 20 years and frequently

publishes research articles in English, while the first and third raters have only 5 years of teaching experience. These different characteristics have impact on the raters' perspective towards technicality of words. That is, the rater with considerable experience might feel that the words related to his or her field are common words (Levels 1-2) and categorize other words as too technical (Level 3). The raters with shorter length of experience, in contrast, might primarily focus on their discipline-specific words and classify them as technical words (Level 3). As a result, some general science words were incidentally removed from the final SAWL. However, the agreement among the three experts was tested and still showed a high rate of agreement among the experts: 82.37%. For future research, it is suggested that there should be an opportunity for the invited experts to discuss together before rating the prospective words independently.

Table 4.4
The coverage of 15 SAWL Sublists on the SAJ corpus

Levels of words	Running words		Groups	
	No. of running word	Percent	No. of Groups	Percent
SAWL Sublist 01	264,462	4.75%	60	0.06%
SAWL Sublist 02	134,023	2.41%	60	0.06%
SAWL Sublist 03	94,791	1.70%	60	0.06%
SAWL Sublist 04	72,923	1.31%	60	0.06%
SAWL Sublist 05	58,703	1.06%	60	0.06%
SAWL Sublist 06	49,101	0.88%	60	0.06%
SAWL Sublist 07	40,288	0.72%	60	0.06%
SAWL Sublist 08	33,433	0.60%	60	0.06%
SAWL Sublist 09	27,926	0.50%	60	0.06%
SAWL Sublist 10	25,864	0.46%	60	0.06%
SAWL Sublist 11	19,247	0.35%	60	0.06%
SAWL Sublist 12	15,983	0.29%	60	0.06%
SAWL Sublist 13	13,277	0.24%	60	0.06%
SAWL Sublist 14	10,660	0.19%	60	0.06%
SAWL Sublist 15	1,322	0.02%	13	0.01%
Total	862,003	15.48%	853	0.85%

The final SAWL consists of 853 word families – 421 AWL words and 432 non-AWL words. The headwords of SAWL is presented in Appendix A and the complete word family SAWL is shown in Appendix B. The family members are based on Nation (2012). The whole SAWL is also divided by frequency into 15 sublists

as shown in Appendix C. Each sublist comprises 60 word families, except for the final sublist, i.e. Sublist 15, which contains the remaining 13 word families. The coverage of the sublists begins at 4.75% of the first sublist and drops to 0.01% by the 15th sublist. These sublists are useful for the users who focus on learning specialized academic words by their frequency.

4.1.3 Validity test

According to Liu and Han (2015) and Nation and Webb (2011), the development of a specialized academic word list should be complemented by a validity test. The purpose of this process is to prove that the SAWL is appropriate for academic learning of the natural science disciplines. This validity test evaluates the SAWL in two aspects: its performance in specialized context and in general context.

First, the SAWL was checked against the SAJ corpus, which is the corpus it was made from, to see that SAWL contains academic words of natural science disciplines. It was also hypothesized that the SAWL should have equal or higher text coverage rate than the AWL. As shown in Table 4.5, the AWL provides 10% coverage of the corpus. However, the SAWL brought up to 15.5% coverage of the corpus. This means the SAWL contains useful academic words and it works better than the AWL in natural science disciplines. The SAWL was also tested against the Pilot-SAJ corpus, which is the parallel corpus of the SAJ corpus. As a result, the SAWL covers 14.65% of a running word. This implies that the SAWL also works well in other journal articles outside the SAJ corpus.

Second, to demonstrate that general words are not included in the SAWL, the list was cross-checked against another independent corpus. In the present study, the English News Corpus (ENC) (Goldhahn, Eckart, & Quasthoff, 2012) was used. The ENC is a corpus of online English news and available online at http://wortschatz.uni-leipzig.de/en/download/. As presented in Table 4.5, although 805 SAWL word families occur in the ENC corpus, the coverage rate is very low (4%). The results confirm that the SAWL contains specialized academic words in natural science disciplines.

In conclusion, the coverage rates of the SAWL in different corpora support the hypothesis that the SAWL is the list of specialized word list frequently

found in natural science disciplines. All 853 word families of SAWL are commonly used in journal articles in these disciplines. The list has been proven beneficial for EFL science students. Further discussion on the development and validation of the SAWL will be presented in the following section.

Table 4.5
The distribution of SAWL in three corpora

Corpora	Token	SAWL Token		SAWL Type		SAWL Group	
		Token	Token%	Type	Type%	Group	Group%
SAJ	5,562,996	862,003	15.50	3,473	3.13	853	0.79
Pilot-SAJ	1,111,298	162,875	14.65	2,900	7.60	853	2.37
ENC	5,994,530	267,387	4.46	3,206	2.35	805	0.60

4.1.4 Discussion

4.1.4.1 Academic language of science

As mentioned earlier in Chapter 2, scientific English is important for being successful in the scientific community as it has several special characteristics. One of the findings of this study supports Templeton, Bear, and Johnston (2010) that science students encounter three levels of vocabulary: general words, academic words, and content-specific words. Evidence from this corpus-based study elaborates more on this point.

The SAJ corpus shows the proportion of the three types of vocabulary in science fields (see Table 4.2). General high frequency words, i.e. the GSL (West, 1953), generally accounts for 75-80% in academic texts (Gilner, 2011; Yang, 2015). Recent studies in the financial corpus (Tongpoon-Patanasorn, 2018) and social sciences corpus (Chanasattru & Tangkiengsirisin, 2016) also yield similar results. However, the portion of general high frequency words in science disciplines is smaller. In the present study, the GSL accounts for approximately 63% coverage of the SAJ corpus. The figure is in line with other corpus-based studies of scientific English. For example, Coxhead and Hirsh (2007) found that the GSL covers 65% of their corpus of science disciplines. The GSL also reaches 65% coverage in the chemistry corpus (Valipouri & Nassaji, 2013) and 60% coverage in the microbiology corpus (Boonyos, 2014).

While the portion of general words is smaller, academic words gain larger portions in science disciplines. As a word list of general academic words, the AWL (Coxhead, 2000) normally covers 10% of the corpus. However, in the SAJ corpus of natural science disciplines, the AWL together with specialized academic words account for 15.5% coverage. Similarly, Valipouri and Nassaji (2013) reports that the chemistry academic words covers 16.83% of the corpus and Liu and Han (2015) discover that the environmental science academic words covers 15.43% of the corpus.

According to Coxhead (2014), learners need to know about 98% of words in a text in order to read and comprehend without support. In natural science disciplines, general high frequency words (2,000 words) with science academic words (853 words) would account for 75-80% coverage. That means 2,853 words are still not enough for the science student to comprehend the text. Therefore, the science students need to deal with low-frequency words and content-specific words accounting for 20-25% coverage. Although it is a large proportion, linguistic knowledge does not solely play a role in acquiring these words. Low-frequency words normally take up 3-5% of the texts and students can use vocabulary-learning strategies, e.g. guessing the meaning or using a dictionary, to understand their meaning. The content-specific words may include proper names, abbreviations, and foreign words, in which science students can figure out the meaning of these words from their previous knowledge.

4.1.4.2 Mixed approach of word list making

The present study used two approaches to make the SAWL: the corpus-based approach and the expert-judged approach. While the first approach is a purely quantitative method (e.g. Brezina & Gablasova, 2015; Coxhead, 2000), the later approach is a mixture of qualitative and quantitative methods (e.g. Ackermann & Chen, 2013; Browne, 2013).

A four-point rating scale (Chung & Nation, 2004) is frequently adopted in the qualitative part of several studies (e.g. Coxhead & Demecheleer, 2018; Tongpoon-Patanasorn, 2018; Valipouri & Nassaji, 2013) to find if the words that might be either too general or too specific. However, the rating scale used in the present study was modified by reducing from four levels to three levels because of feedback from the raters during the pilot study stage. This problem is also mentioned in Ha and Hyland (2017). The use of the modified three-level scale to judge

too technical words in this study corresponds with Coxhead and Demecheleer (2018) and Tongpoon-Patanasorn (2018).

Recommendations for using this tool in the future are that, instead of solely rating by experts in the disciplines, an ESP teacher should be added to the panel of raters in order to voice opinions from a pedagogical perspective. In addition, the rating scale has been used to dismiss too specific words from the list. There should be an opportunity for the raters as an expert to suggest some important words to be added in the list. This is similar to the pedagogical principal behind the GSL (West, 1953) and the NGSL (Browne, 2013), in which useful words for EFL learners are intentionally inserted into the list even if they do not frequently occur in their corpora. In the present study, one rater made a comment that *"limit"* has specialized meaning in science disciplines and science students should know it. She personally suggested that this word should be on the list. However, as the term already exists in the GSL, including the GSL word in the SAWL would violate the established criteria. This reflects that the criteria could be improved in future research.

4.2 The alignment of the SAWL with the CEFR

This section aimed to answer the third research question: To what extent do the lexical items in the SAWL align with the CEFR? The purpose of this question is to propose an alternative method of selecting specialized vocabulary to teach or to learn. To assign 6 levels of CEFR to each of 853 SAWL word families, two research-based CEFR Level Descriptions (RLD) were employed, namely *English Vocabulary Profile (EVP)* and *Global Scale of English Vocabulary (GSE Vocabulary)*. The first part of this section reports the results of the CEFR alignment process together with the list of SAWL words with their CEFR levels. Then, the second part shows how the results were verified. The discussion of issues involved are presented in the last part.

4.2.1 The CEFR alignment approach

The process of mapping the SAWL words with the CEFR consists of four main steps. First, the representatives of each word family were selected based on their frequency. In Capel (2012), the most frequent family members of the AWL

(Coxhead, 2000) were selected as word family representatives and were aligned with the CEFR levels based on their frequency in the corpus. The present study adopted this process. The most frequent family members that occur in the SAJ corpus have been included in the list of 853 SAWL most frequent family members. For example, the word family *"acquire"* comprises six word members, i.e. *acquire, acquired, acquires, acquiring, acquisition,* and *acquisitions*. The frequency rates of each word in the SAJ corpus are 84, 326, 11, 34, 358, and 10 times respectively. *Acquisition* occurs most often in the corpus. Hence, it was chosen as the representative of the family *"acquire"*. Figure 4.3 presents the list of 853 most frequent family members (also marked with ^ with Appendix B).

Figure 4.3

The list of 853 SAWL most frequent family members

001.absorption
002.abstract
003.abundance
004.academic
005.access
006.accommodate
007.accompanied
008.accuracy
009.acetate
010.acetic
011.acetone
012.achieved
013.acid
014.acquisition
015.activation
016.acute
017.adapted
018.additives
019.adequate
020.adhesion
021.adjacent
022.adjusted
023.administration
024.adult
025.adverse
026.affect
027.affinity
028.agar
029.agarose
030.aggregation
031.aid
032.albumin
033.alcohol
034.alignment
035.alkaline
036.allele
037.allocation
038.altered
039.alternative
040.ambient
041.amino
042.ammonia
043.amplification
044.amplitude
045.anaerobic
046.analogous
047.analysis
048.annealing
049.annual
050.antimicrobial
051.antioxidant
052.apoptosis
053.apparent
054.appendix
055.approach
056.appropriate
057.approximately
058.aqueous
059.arbitrary
060.architecture
061.area
062.aromatic
063.array
064.aspects
065.assay
066.assembly
067.assessed
068.assigned
069.assisted
070.assume
071.atmosphere
072.atoms
073.attached
074.attained
075.attitude
076.attributed
077.authors
078.automated
079.available
080.aware
081.bacillus
082.bacteria
083.barrier
084.basal
085.baseline
086.batch
087.beads
088.beneficial
089.bias
090.bioactive
091.biochemical
092.biology
093.biomass
094.biomedical
095.biosynthesis
096.biotechnology

097.bonds
098.bovine
099.breakdown
100.breast
101.briefly
102.buffer
103.bulk
104.calcium
105.calibrate
106.cancer
107.candidate
108.capable
109.capacity
110.capillary
111.capture
112.carbohydrate
113.carbon
114.cardiac
115.cardiovascular
116.cascade
117.catalytic
118.categories
119.cavity
120.cellular
121.cellulose
122.centrifugation
123.cerevisiae
124.challenge
125.chamber
126.channel
127.chemical
128.chemistry
129.chip
130.chloride
131.cholesterol
132.chromatography
133.chromosome
134.chronic
135.clarify
136.classical
137.climate
138.clinical
139.clones
140.coating
141.code
142.coefficient
143.coincide
144.colon
145.column
146.comments
147.communication
148.community
149.compatible
150.compensate
151.complementary
152.complex
153.components
154.composition
155.compounds
156.comprehensive
157.comprised
158.computational
159.concept
160.conclusion
161.concurrent
162.conducted
163.conference
164.configuration
165.confirmed
166.confocal
167.conformation
168.consequently
169.conserved
170.considerable
171.consistent
172.constant
173.constituents
174.constraints
175.construct
176.consumption
177.contact
178.contamination
179.context
180.contrast
181.contribute
182.conventional
183.conversely
184.conversion
185.cooperation
186.coordinates
187.core
188.correlation
189.corresponding
190.coupled
191.covalent
192.created
193.criteria
194.crucial
195.crude
196.crystal
197.culture
198.cumulative
199.cycle
200.cysteine
201.dashed
202.data
203.database
204.dataset
205.decades
206.decline
207.decomposition
208.defects
209.deficient
210.defined
211.definite
212.degradation
213.dehydrogenase
214.demonstrated
215.denote
216.density
217.dependence
218.depicted
219.depletion
220.deposited
221.deposition
222.depression
223.design
224.despite
225.detection
226.developmental
227.device
228.devoted
229.diagram
230.diet
231.differential
232.differentiation
233.diffraction
234.diffusion
235.digestion
236.digital
237.diluted
238.dimensional
239.discrepancy
240.discrete
241.discrimination
242.dispersion
243.displacement
244.display
245.disruption
246.dissolution
247.dissolved
248.distilled
249.distinct
250.distribution
251.divergence
252.diversity

253.documented
254.domain
255.dominant
256.donor
257.downstream
258.dramatically
259.droplet
260.drug
261.dual
262.duration
263.dye
264.dynamics
265.ecological
266.economic
267.ecosystems
268.edited
269.efficiently
270.electrode
271.electron
272.electrophoresis
273.electrostatic
274.elemental
275.elements
276.elevated
277.elimination
278.elongation
279.embedded
280.emergence
281.emission
282.emphasis
283.empirical
284.enable
285.encoding
286.encountered
287.endogenous
288.energy
289.engineered
290.enrichment
291.ensure
292.entities
293.environmental
294.enzymatic
295.enzyme
296.epithelial
297.equation
298.equipped
299.equivalent
300.error
301.established
302.esters
303.estimated
304.ethanol
305.evaluated
306.evaporation
307.eventually
308.evidence
309.evolution
310.exceed
311.excitation
312.excluded
313.exhibited
314.exogenous
315.expanded
316.experts
317.explicit
318.exploited
319.exponential
320.exposure
321.external
322.extraction
323.fabricated
324.facilitate
325.factors
326.favorable
327.favorably
328.feasible
329.features
330.feedback
331.feeding
332.fermentation
333.fiber
334.file
335.filter
336.finally
337.finite
338.flexible
339.fluctuations
340.fluid
341.fluorescence
342.flux
343.focus
344.formula
345.fraction
346.fragments
347.framework
348.frequencies
349.function
350.functionalized
351.fundamental
352.fungal
353.furthermore
354.fusion
355.gel
356.gene
357.generated
358.generation
359.genetic
360.genome
361.genotypes
362.genus
363.geographic
364.geometry
365.germ
366.global
367.glucose
368.glycerol
369.goal
370.grade
371.graph
372.grid
373.guarantee
374.gut
375.height
376.hence
377.hierarchical
378.highlighted
379.hybrid
380.hydrogen
381.hydrolysis
382.hydrophilic
383.hydrophobic
384.hydroxyl
385.hypothesis
386.identical
387.identified
388.illustrated
389.image
390.imaging
391.immobilized
392.immune
393.impact
394.implemented
395.implications
396.implies
397.imposed
398.incidence
399.incorporated
400.incubated
401.index
402.indicated
403.individual
404.induced
405.infection
406.inference
407.inflammatory
408.infrared

409.inherent
410.inhibition
411.initial
412.initiation
413.injection
414.injury
415.innovation
416.input
417.insertion
418.inset
419.insights
420.instance
421.instructions
422.intact
423.intake
424.integral
425.integrated
426.intensity
427.interactions
428.interestingly
429.interface
430.interior
431.intermediate
432.internal
433.interpretation
434.interval
435.intervention
436.intestinal
437.intrinsic
438.invasive
439.inverse
440.investigated
441.involved
442.ion
443.isolated
444.items
445.kernel
446.kidney
447.kinase
448.kinetic
449.kit
450.labeled
451.laboratory
452.lactic
453.laser
454.latent
455.lateral
456.lattice
457.layer
458.linear
459.linked
460.lipid
461.liver
462.localization
463.location
464.locus
465.longitudinal
466.loop
467.lysine
468.magnesium
469.magnetic
470.magnification
471.magnitude
472.maintained
473.major
474.manipulation
475.manually
476.mapping
477.marginal
478.marine
479.mathematics
480.matrix
481.mature
482.maximal
483.maximum
484.mechanism
485.media
486.median
487.mediated
488.medical
489.medium
490.membrane
491.mesh
492.metabolic
493.metabolism
494.metabolites
495.methanol
496.methionine
497.method
498.micro
499.microbial
500.microorganisms
501.microscopy
502.migration
503.minimal
504.minimize
505.minimum
506.minor
507.mitochondrial
508.mobile
509.mode
510.modified
511.molar
512.molecular
513.monitoring
514.morphology
515.mortality
516.motivated
517.mounted
518.muscle
519.mutant
520.mutations
521.nanoparticles
522.negative
523.negligible
524.network
525.neural
526.neutral
527.nevertheless
528.nitrogen
529.nodes
530.norm
531.normal
532.notion
533.novel
534.nuclear
535.nucleotide
536.nucleus
537.null
538.nutrient
539.nutritional
540.objective
541.obtained
542.obvious
543.occupied
544.occur
545.odd
546.online
547.onset
548.optical
549.optimal
550.optimization
551.optimum
552.option
553.oral
554.organic
555.organisms
556.orientation
557.outcomes
558.output
559.oven
560.overall
561.overlap
562.overnight
563.overview
564.oxide

565.oxidized
566.oxygen
567.panel
568.parallel
569.parameters
570.participants
571.partners
572.passive
573.pathogen
574.pathogenic
575.pathway
576.patients
577.peak
578.penetration
579.peptide
580.percentage
581.perception
582.period
583.peripheral
584.peroxide
585.persistence
586.perspective
587.pharmaceutical
588.phase
589.phenomenon
590.phenotypic
591.phosphate
592.phylogenetic
593.physical
594.physiological
595.plasma
596.plastic
597.platform
598.plot
599.plus
600.polar
601.poly
602.polymer
603.polymerase
604.polynomial
605.pooled
606.pore
607.portion
608.pose
609.positive
610.posterior
611.potassium
612.potent
613.potential
614.preceding
615.precipitation
616.precision
617.precursor
618.predicted
619.predominantly
620.preliminary
621.presumably
622.prevalence
623.previous
624.primary
625.prime
626.primers
627.principal
628.principle
629.prior
630.probe
631.procedure
632.process
633.profile
634.progression
635.projection
636.proliferation
637.proline
638.promote
639.propagation
640.proportion
641.protease
642.protein
643.protocol
644.publication
645.published
646.pulse
647.purchased
648.purified
649.purity
650.pursued
651.putative
652.quantification
653.radical
654.random
655.range
656.ratio
657.rational
658.reaction
659.reagents
660.receptor
661.recovery
662.redox
663.reference
664.refinement
665.regime
666.region
667.regression
668.regulation
669.rejection
670.relaxation
671.release
672.relevant
673.reliable
674.removed
675.replication
676.required
677.research
678.reservoirs
679.residence
680.residues
681.resolution
682.resonance
683.resources
684.response
685.restoration
686.restricted
687.resuspended
688.retention
689.revealed
690.reverse
691.revision
692.robust
693.role
694.rotation
695.route
696.routine
697.salinity
698.saturation
699.scaling
700.scan
701.scenario
702.scheme
703.score
704.seasonal
705.secretion
706.section
707.sector
708.seek
709.segment
710.selected
711.sensing
712.sensor
713.sequence
714.series
715.serine
716.serum
717.setup
718.sex
719.shift
720.significant

721.silica
722.silicon
723.similar
724.simulation
725.simultaneously
726.site
727.skeletal
728.sodium
729.software
730.solely
731.soluble
732.solvent
733.somewhat
734.source
735.spatial
736.species
737.specific
738.specified
739.spectra
740.spectral
741.sphere
742.spontaneous
743.stability
744.static
745.statistical
746.status
747.storage
748.straightforward
749.strains
750.strand
751.strategy
752.stress
753.structure
754.subsequent
755.subset
756.substitution
757.successive
758.sufficient
759.sum
760.summary
761.superior
762.supernatant
763.supplementary
764.suppression
765.survey
766.survival
767.susceptible
768.suspension
769.sustainable
770.switch
771.symbols
772.symmetry
773.symptoms
774.synergistic
775.synthesis
776.synthetic
777.tank
778.target
779.task
780.taxonomic
781.team
782.technical
783.techniques
784.technology
785.template
786.temporal
787.terminal
788.tertiary
789.text
790.theory
791.therapeutic
792.therapy
793.thereby
794.thermal
795.threshold
796.tissue
797.tolerance
798.topic
799.toxic
800.toxin
801.trace
802.tract
803.traditional
804.transcription
805.transfer
806.transformation
807.transient
808.transition
809.transmission
810.transport
811.trend
812.triangle
813.trigger
814.triple
815.triplicate
816.tumor
817.tyrosine
818.ultimately
819.unclear
820.undergo
821.underlying
822.uniform
823.unique
824.untreated
825.uptake
826.urea
827.utilized
828.validation
829.variables
830.vectors
831.velocity
832.verify
833.version
834.vertical
835.via
836.viability
837.video
838.violation
839.virtually
840.virus
841.vitamin
842.volatile
843.voltage
844.volume
845.vs
846.wavelength
847.weighted
848.whereas
849.widespread
850.worldwide
851.yeast
852.zinc
853.zone

The second step of mapping the SAWL words with the CEFR is to find out how the words were commonly used in the SAJ corpus. In other words, the parts of speech of the 853 SAWL high frequency words were investigated. A Part-Of-

Speech (POS) tagger, namely TagAnt (Anthony, 2015), was used to convert the original SAJ corpus into the POS-tagged files. The computer program then analyzed words in the SAJ corpus files and attached the POS tags (e.g. **_JJ** for adjective, **_NN** for noun, and **_VV** for verb) at the end of every word automatically. Excerpts from the original SAJ and tagged SAJ files are shown in Figure 4.4. Then AntConc (Anthony, 2018), a concordance computer software, was employed in order to explore parts of speech of each SAWL word in the whole corpus. Figure 4.5 presents that sample report from AntConc. After the parts of speech of all SAWL words had been determined (as shown in Appendix D), the 853 words were cross-checked against two databases in the following step.

Figure 4.4

Excerpts from the original and the POS-tagged files

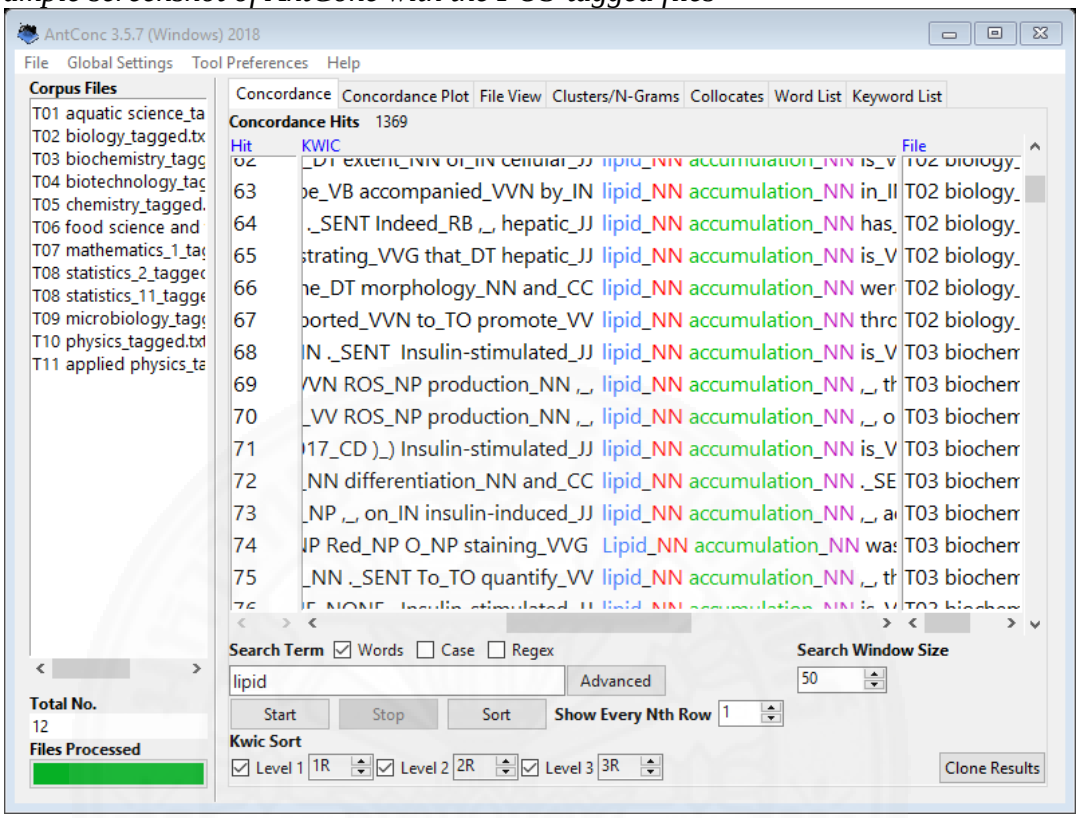
Original

The persistence of coral reef ecosystems under climate change scenarios depends on the ability of reef-building corals to survive and grow in warmer oceans. Rising seawater temperature is a major threat facing coral reefs, as thermal stress causes destabilization of the symbiosis between corals and their symbiotic algae (Symbiodinium spp.), a phenomenon known as coral bleaching. In association with climate warming, coral reef bleaching events worldwide are becoming more frequent and severe, often leading to mass coral mortality (Hoegh-Guldberg et al. 2007; Baker et al. 2008).

POS-Tagged

The_DT persistence_NN of_IN coral_JJ reef_NN ecosystems_NNS under_IN climate_NN change_NN scenarios_NNS depends_VVZ on_IN the_DT ability_NN of_IN reef-building_VVG corals_NNS to_TO survive_VV and_CC grow_VV in_IN warmer_JJR oceans_NNS ._SENT Rising_VVG seawater_NN temperature_NN is_VBZ a_DT major_JJ threat_NN facing_VVG coral_JJ reefs_NNS ,_, as_IN thermal_JJ stress_NN causes_VVZ destabilization_NN of_IN the_DT symbiosis_NN between_IN corals_NNS and_CC their_PP$ symbiotic_JJ algae_NN (_(Symbiodinium_NP spp_NP ._SENT)_) ,_, a_DT phenomenon_NN known_VVN as_IN coral_JJ bleaching_VVG ._SENT In_IN association_NN with_IN climate_NN warming_NN ,_, coral_JJ reef_NN bleaching_VVG events_NNS worldwide_RB are_VBP becoming_VVG more_RBR frequent_JJ and_CC severe_JJ ,_, often_RB leading_VVG to_TO mass_VV coral_JJ mortality_NN (_(Hoegh-Guldberg_NP et_NP al_NP ._SENT 2007_LS ;_: Baker_NP et_NP al_NP ._SENT 2008_LS)_) ._SENT

Figure 4.5
Sample screenshot of AntConc with the POS-tagged files



The third step of the process is to check the SAWL words against two CEFR-RLD databases, namely EVP and GSE Vocabulary. The CEFR levels of the same word from two databases were compared based on four established criteria: (1) the six CEFR levels used in this study include A1, A2, B1, B2, C1, and C2; (2) only the CEFR levels of a word that match meanings and the parts of speech of the one found in the SAJ corpus will be counted; (3) the lower CEFR levels will be used if both databases yield different results; and (4) the words will be tagged with UL (unlisted) if they are not in both databases.

Finally, after applying the criteria, it was revealed that the majority of SAWL words are at B1 (19%), B2 (41%), and C1 (20%) levels (see Table 4.6). The CEFR-based sublists are presented in Appendix D. Although there are six levels, it is not necessary to start learning from A1 to C2. The users may focus on the level that is suitable for learners' proficiency.

Table 4.6
Number of SAWL words in each CEFR levels

CEFR Level	No. of Words	Coverage over the SAJ corpus
A1	3	0.12%
A2	39	0.49%
B1	158	2.99%
B2	350	3.79%
C1	171	1.63%
C2	7	0.05%
UL	125	0.86%
Total	**853**	**9.93%**

4.2.2 The Validity Test of CEFR Alignment

To confirm that the CEFR levels of 853 SAWL words are valid, two methods were applied – testing on its own corpus and testing on another independent corpus. First, the CEFR-mapped SAWL sublists were tested against the SAJ corpus. It was hypothesized that A1/A2 words should have higher occurrences in the corpus than B1/B2 and C1/C2 words. The results shown in Table 4.7 support the claim. The average occurrence of A1/A2 words is 803 times per word, which is obviously higher than B1/B2 words (742 times per word) and C1/C2 words (523 times per word). The UL words, in which their CEFR levels cannot be classified, also occur with a very low frequency rate (381 times per word) in the SAJ corpus. The occurrence of the CEFR-mapped SAWL has been proven in line with the hypothesis..

Table 4.7
The average occurrence of CEFR-based SAWL words

CEFR Levels	No. of Words	Average Occurrence	S.D.
A1/A2	42	803	900
B1/B2	508	742	924
C1/C2	78	523	381
UL	125	381	332

The second validity test is to prove that the CEFR levels of SAWL words are reliable and accurate. Ishikawa's (2013) The International Corpus Network of Asian Learners of English (ICNALE), which is a corpus of written essays composed by 5,600 Asian learners of English with CEFR-linked proficiency control, was employed in this study in order to investigate the occurrence of SAWL words at

different CEFR levels of the learners. The ICNALE Written Essay corpus divides proficiency levels of learners into four levels – A2, B1 Lower, B1 Upper, and B2 or higher.

In this study, the ICNALE was divided into three subcorpora – A2, B1, and B2. Each subcorpus was examined against CEFR-mapped SAWL sublists. The SAWL-A1 and-A2 sublists were matched with the ICNALE-A2 subcorpus. The SAWL-B1 and –B2 sublists were tested against the ICNALE-B1 and –B2 subcorpora. This is to assess the good performance of the CEFR-mapped SAWL sublists of A1-B2 levels. As the ICNALE does not explicitly separate C1 and C2 levels, the hypothesis for examining sublists of C1-C2 levels is to prove that most words do not belong to the B1/B2 subcorpus. Similarly, the test of UL sublist is to prove that the list does not belong to any level of the whole ICNALE Written Essay. Table 4.8 summarizes the second validity test process. As shown in Table 4.6, the majority of the SAWL words falls into the B1 and above levels of the CEFR. This helps to illuminate the connection between the CEFR and cognitive development. North (2014) explains that learning using a language involves more cognitive abilities and the CEFR embraces this notion as seen in many descriptors for the B1 and above levels. For example, language users at the B2 level "*can understand the main ideas of complex text on both concrete and abstract topics, including* ***technical discussions*** *in his/her field of specialization*". The findings from the present study can confirm the significance of the B2 level for science disciplines because it was proven that professional scientists mostly use academic words at this level when they compose academic texts. This is also in line with linguistic relativity theories, like Sapir-Whorf hypothesis, that language is tied together with cognitive competence and culture. People of different discourses in the same language or across languages will have different cultural worldviews (Kramsch, 2004). For the EFL science students, apart from general vocabulary and content knowledge, they need to learn specialized academic words at least until B2 level in order to acquire the knowledge, understand the scientific viewpoints, and become a part of the scientific community.

In Thailand, according to the National Education Plan, students at the tertiary level are expected to achieve the B2 level of the CEFR. This is also in line with the findings from the study. When the science students read academic texts in

English, specialized academic words are proven necessary for them. However, English teachers cannot fully carry out this responsibility. Teaching specialized vocabulary, like the SAWL words, can be a joint venture between English teachers and content teachers. That means English teachers and science teachers need to work together. For example, the ESP course for science students can focus on common SAWL words and the difficult SAWL words could be integrated with the content subjects and employ new teaching approaches, like CLIL (Content and Language Integrated Learning) and EMI (English Medium Instruction).

Table 4.8
The second validity test of the SAWL

Levels	SAWL	ICNALE	Expected Results
A	Sublist 1 (A1) Sublist 2 (A2)	Subcorpus 1 (A2)	Appear
B	Sublist 3 (B1) Sublist 4 (B2)	Subcorpus 2 (B1/B2)	Appear
C	Sublist 5 (C1) Sublist 6 (C2)	Subcorpus 2 (B1/B2)	Disappear
UL	Sublist 7 (UL)	Subcorpora 1 – 2 (A2-B2)	Disappear

Table 4.9 shows the result of the second validity test. At CEFR-A and CEFR-B levels, it was found that 81-83% of the SAWL sublists occur in the ICNALE subcorpora. That implies that the learners at CEFR-A and CEFR-B levels know most words in the SAWL sublists of CEFR-A and CEFR-B levels respectively. For CEFR-C level, because the ICNALE does not spell it out from the CEFR-B level explicitly, the validity test was done by expecting that the majority of the words in SAWL sublists 5 and 6 should occur minimally in the ICNALE subcorpora 2. The investigation yielded the expected results. That is 67 out of 178 (38%) of the words in SAWL sublists 5 and 6 occur in the ICANLE subcorpora 2, which means most words in the aforementioned sublists are beyond CEFR-B level. The test of UL level also reports that only 9 out of 125 (7%) of the words in SAWL sublist 7 occur in the whole ICANLE Written Essay corpus. This means most words in this sublist are beyond any CEFR levels represented in the corpus. Again, the second validity test confirms the CEFR levels tagged on the 853 SAWL words.

Table 4.9
The results of the second validity test of the CEFR-mapped SAWL

Levels	Total SAWL Words	SAWL Words in ICANLE Subcorpora	Percent
A	42	34	81%
B	508	420	83%
C	178	67	38%
UL	125	9	7%

4.2.3 Discussion

4.2.3.1 The CEFR levels of academic words in science disciplines

As shown in Table 4.6, the majority of the SAWL words falls into the B1 and above levels of the CEFR. This helps to illuminate the connection between the CEFR and cognitive development. North (2014) explains that learning using a language involves more cognitive abilities and the CEFR embraces this notion as seen in many descriptors for the B1 and above levels. For example, language users at the B2 level "*can understand the main ideas of complex text on both concrete and abstract topics, including* ***technical discussions*** *in his/her field of specialization*". The findings from the present study can confirm the significance of the B2 level for science disciplines because it was proven that professional scientists mostly use academic words at this level when they compose academic texts. This is also in line with linguistic relativity theories, like Sapir-Whorf hypothesis, that language is tied together with cognitive competence and culture. People of different discourses in the same language or across languages will have different cultural worldviews (Kramsch, 2004). For the EFL science students, apart from general vocabulary and content knowledge, they need to learn specialized academic words at least until B2 level in order to acquire the knowledge, understand the scientific viewpoints, and become a part of the scientific community.

In Thailand, according to the National Education Plan, students at the tertiary level are expected to achieve the B2 level of the CEFR. This is also in line with the findings from the study. When the science students read academic texts in English, specialized academic words are proven necessary for them. However, English teachers cannot fully carry out this responsibility. Teaching specialized vocabulary, like the SAWL words, can be a joint venture between English teachers and content teachers.

That means English teachers and science teachers need to work together. For example, the ESP course for science students can focus on common SAWL words and the difficult SAWL words could be integrated with the content subjects and employ new teaching approaches, like CLIL (Content and Language Integrated Learning) and EMI (English Medium Instruction).

4.2.3.2 Choosing words to teach

There are several ways to choose words to teach. Previously, Watson Todd (2017) suggested that there are at least three bases. English teachers may prioritize frequency words, learnable and teachable words, or opaque words. The present study proposes an alternative method of choosing words to teach, i.e. using the CEFR levels of the words. By using the proposed criteria and databases, specialized words can be appropriately mapped with their CEFR levels and the words that suit the learners' proficiency will be brought into focus. The teachers might select words according to themes or topics.

After the target words are tagged with CEFR levels, the next question is *"Which levels should be selected to teach – at the same level of the learners' proficiency or above their level?"* To answer this question, the principle behind the mapping criteria should be explained. In line with Capel (2010), the CEFR levels of the SAWL words mean they are commonly known by the learners at that level. For example, the word '*affect*' in this study is at B1 level. That means the learners at B1 level commonly know this word. The evidence derived from the ICANLE also supports that it is frequently used by B1 learners. Furthermore, for the B1 learners, most words on the third sublist (B1) of CEFR-mapped SAWL might be known. However, it is not always necessary to make a jump onto the higher sublists because the sublists are only a rough guide for vocabulary teaching. There are other aspects of the words that can be elaborated further such as collocations or discipline-specific. Lexical information provided in Appendix D might be useful.

In summary, although the SAWL words are tagged with the level corresponding to the learners' proficiency, it does not mean that the learners could easily comprehend them or use them correctly in context. They may need assistance from the teachers. In other words, choosing the right words to teach is important, but

what is more important is vocabulary teaching strategies as mentioned in Gholami and Zeinolabedini (2018) and Nation (2013). Quality of teaching still comes into play.

4.2.3.3 SAWL collocations

Valipouri and Nassaji (2013) proposed that, apart from the meanings of words, students need to know how they collocate with other words in the context of specific disciplines because learning word combinations is useful for language production (speaking and writing skills). Although this is not the main objective of the present study, collocations of the CEFR-mapped SAWL words will be highlighted. Recently, Ackermann and Chen (2013) developed the Academic Collocation List (ACL), which contains 2,468 highly frequent and pedagogical relevant collocations. The full list is available online at https://pearsonpte.com/organizations/researchers/academic-collocation-list/. The SAWL most frequent members were checked against the ACL to figure out the collocations of SAWL words. Out of 853 SAWL words, 234 words were reported to coincide with the ACL either as the first component or as the second component of the collocations, as shown in Figure 4.6 below.

Figure 4.6

The list of 234 SAWL coinciding with the ACL (Ackermann & Chen, 2013)

I. First component of the collocations (e.g. ***abstract*** *concept*)

1. abstract	22. environmental	43. normal
2. academic	23. external	44. nuclear
3. adverse	24. facilitate	45. occur
4. annual	25. finite	46. online
5. assume	26. flexible	47. optimal
6. beneficial	27. fundamental	48. oral
7. briefly	28. genetic	49. overall
8. chemical	29. global	50. physical
9. classical	30. hierarchical	51. pose
10. comprehensive	31. initial	52. positive
11. considerable	32. integral	53. preliminary
12. consistent	33. internal	54. previous
13. contribute	34. intrinsic	55. primary
14. conventional	35. linear	56. prime
15. crucial	36. longitudinal	57. principal
16. developmental	37. major	58. prior
17. digital	38. maximum	59. promote
18. distinct	39. medical	60. published
19. dominant	40. minimum	61. radical
20. economic	41. modified	62. random
21. empirical	42. negative	63. reliable

64. seek
65. statistical
66. subsequent
67. sufficient
68. superior
69. technical
70. traditional
71. undergo
72. underlying
73. unique
74. virtually

II. Second component of the collocations (e.g. *allow* ***access***)

75. access
76. accuracy
77. acute
78. administration
79. affect
80. analysis
81. apparent
82. approach
83. area
84. array
85. aspects
86. assigned
87. attitude
88. aware
89. capacity
90. challenge
91. communication
92. community
93. concept
94. conclusion
95. conference
96. constraints
97. construct
98. contact
99. context
100. contrast
101. correlation
102. created
103. criteria
104. culture
105. database
106. decades
107. defined
108. degradation
109. demonstrated
110. design
111. discrimination
112. display
113. distribution
114. diversity
115. documented
116. domain
117. dramatically
118. duration
119. elements
120. embedded
121. emphasis
122. encountered
123. energy
124. equivalent
125. error
126. evidence
127. evolution
128. explicit
129. factors
130. features
131. feedback
132. formula
133. fraction
134. framework
135. function
136. generation
137. goal
138. identical
139. identified
140. image
141. impact
142. implications
143. incidence
144. innovation
145. instructions
146. intensity
147. interpretation
148. intervention
149. involved
150. linked
151. location
152. method
153. motivated
154. network
155. norm
156. orientation
157. overview
158. panel
159. parameters
160. peak
161. percentage
162. perception
163. period
164. perspective
165. phase
166. phenomenon
167. portion
168. principle
169. procedure
170. profile
171. proportion
172. range
173. reaction
174. reference
175. removed
176. resolution
177. resources
178. response
179. role
180. score
181. section
182. sector
183. selected
184. sex
185. shift
186. species
187. sphere
188. stability
189. status
190. straightforward
191. strategy
192. structure
193. summary
194. task
195. technology
196. text
197. theory
198. topic
199. transformation
200. transition
201. trend
202. unclear
203. version

III. First or second component of the collocations (e.g. ***alternative*** *approach / provide (an)* ***alternative***)

204. alternative	215. individual	226. significant
205. appropriate	216. integrated	227. similar
206. available	217. media	228. source
207. climate	218. minor	229. specific
208. complex	219. objective	230. stress
209. constant	220. obvious	231. survey
210. core	221. potential	232. target
211. data	222. preceding	233. transport
212. established	223. process	234. widespread
213. focus	224. relevant	
214. geographic	225. research	

These 234 words are a good starting point for learning collocations in natural science disciplines because they are frequently used and also occur in other academic disciplines. Based on the ACL, these 234 SAWL words occur in either or both positions of 1,390 collocations or approximately 56% of all collocations of the ACL. For example, the word *published* occurs in the first part of its collocations. According to the ACL, this word usually co-occurs with *literature, research, work,* and *material*. The word *process* occurs in both first and second parts of its collocations such as *assessment process*, *begin a process*, *process data*, and *process information*. The complete SAWL collocation list is shown in Appendix F.

In addition, further analysis of the 1,390 ACL collocations in the SAJ corpus illustrated that 1,032 collocations (approximately 75%) exist in the natural science discipline. For instance, the SAWL collocation list shows 15 words that co-occur with the word *resources*. Only 12 collocations were found in the SAJ corpus. *Additional resources, electronic resources,* and *learning resources* are not presented. The collocations found in the SAJ corpus are tagged with "sci" in Appendix F. The present study illustrates the distribution and confirms the usefulness of the ACL (Ackermann & Chen, 2013) in natural science disciplines.

4.3 The comparison with other word lists

This section aimed to answer the fourth research question: To what extent do the SAWL differ from the SWL (Coxhead & Hirsh, 2007) and the EST Coursebook Word List? The purpose of this research question is to demonstrate the distinctive

properties of the SAWL. There are two parts in this section. The comparison between the SAWL and the SWL (Coxhead & Hirsh, 2007) is discussed in the first part. The comparison between the SAWL and the EST Coursebook Word List is revealed in the second part. The discussion on the issues involved is generated in the final part.

4.3.1 The SAWL and the SWL (Coxhead & Hirsh, 2007)

Both SAWL and SWL (Coxhead & Hirsh, 2007) are established as a specialized word list of science disciplines. Yet, the development and the performance of these two word lists are not similar and needed to be discussed. First, the two word lists have different objectives of development. The SWL was made as a technical word list of 14 science disciplines, mainly including natural science, health science, technological science, and veterinary and animal science. The SWL was created from a corpus of various text types such as textbooks, practical manuals, lecture notes, study guides, and handouts. In addition, the 570 AWL word families were not included in the final SWL list. In contrast, the SAWL focuses on a narrower range of disciplines, i.e. 11 natural science disciplines. The corpus was compiled from only one text type, which is academic journal articles. The AWL words were also allowed in the development of SAWL. These features make the SAWL more useful for the EFL science students who aim to enhance vocabulary for reading skills and may not be familiar with the AWL words.

To demonstrate that the SAWL is more suitable for natural science disciplines the SWL, the distribution of the two lists were compared. Coxhead and Hirsh (2007) state that the SWL contains 318 word families and covers around 4% over its corpus. However, the SWL's coverage was tested again in the present study to compare it with the SAWL. The SWL was checked against the SAJ corpus to explore its coverage over the corpus. As the SWL does not include the AWL, the coverage of the AWL was calculated together with the SWL to make the result comparable with the SAWL. As shown in Table 4.10, the SWL and AWL altogether provide 14% coverage over the SAJ corpus while the SAWL covers 15.5% of the corpus. Comparing the coverage of SAWL, AWL, and SWL sublists presented in Table 4.11 shows that the first sublist of SAWL still provides highest coverage (4.75%, 3% and 1.84% respectively). Comparing the cumulative coverage of the first five sublists of SAWL,

AWL, and AWL also yields similar results (11.23%, 7.94%, and 3.89% respectively). Clearly, the coverage of the SAWL is higher even though it contains less word families than the SWL (318 word families) together with the AWL (570 word families). Again, the SAWL performs better than another word list of science disciplines.

Table 4.10
The coverage of the SWL and the SAWL over the SAJ corpus

Levels of words	Running words		Groups	
	No. of running word	Percent	No. of Groups	Percent
1 AWL	561,119	10.09%	568	0.54%
2 SWL	222,628	4.00%	318	0.30%
3 SAWL	862,003	15.50%	853	0.79%

4.3.2 The SAWL and the EST Coursebook Word List

To prove that the SAWL covers more specialized vocabulary than commercial coursebooks of English for Science and Technology (EST), the coverage of the SAWL was compared with the one of the EST coursebook word list. The EST coursebook word list (ESTCWL) was created by combining five target vocabulary lists from a series of five commercial EST coursebook printed in 2015 by a global publisher. Five books contain 1,067 targeted words. After removing 17 repeated words, the whole ESTCWL comprises 1,050 words, 649 words of which are outside the GSL (West, 1953). As the SAWL does not include the GSL, the coverage of these 649 words of EST coursebook word list were compared with the SAWL. The words were tested against the SAJ corpus to find their coverage in natural science disciplines.

Table 4.12 shows the comparison of the coverage between ESTCWL and the CEFR-mapped SAWL. The CEFR-mapped SAWL was used to compare because it is a list of the 853 most frequent SAWL family members. The findings revealed that, of 649 ESTCWL words, 587 occur in the SAJ corpus and cover approximately 3% of the corpus. Nevertheless, the CEFR-mapped SAWL provides approximately 10% coverage over the SAJ corpus. The results clearly imply that the SAWL contains more words that are useful for EFL natural science students than ESTCWL. Yet, it should be kept in mind that this comparison is merely about the coverage of vocabulary. Low coverage of the ESTCWL does not mean this series of

EST coursebook have inferior quality. Evaluating of ELT coursebooks needs to consider other features.

Table 4.11
The coverage of 16 SAWL, 10 AWL (Coxhead, 2000), 6 SWL (Coxhead & Hirsh, 2007) Sublists on the SAJ corpus

Levels of words	SAWL		AWL		SWL	
	Running word	%	Running word	%	Running word	%
Sublist 01	264,462	4.75	166,718	3.00	102,173	1.84
Sublist 02	134,023	2.41	93,331	1.68	43,244	0.78
Sublist 03	94,791	1.70	66,106	1.19	26,712	0.48
Sublist 04	72,923	1.31	60,093	1.08	25,304	0.45
Sublist 05	58,703	1.06	55,224	0.99	18,865	0.34
Sublist 06	49,101	0.88	37,758	0.68	6,327	0.11
Sublist 07	40,288	0.72	39,689	0.71	-	-
Sublist 08	33,433	0.60	33,385	0.60	-	-
Sublist 09	27,926	0.50	22,200	0.40	-	-
Sublist 10	25,864	0.46	7,765	0.14	-	-
Sublist 11	19,247	0.35	-	-	-	-
Sublist 12	15,983	0.29	-	-	-	-
Sublist 13	13,277	0.24	-	-	-	-
Sublist 14	10,660	0.19	-	-	-	-
Sublist 15	1,322	0.02	-	-	-	-
Total	**862,003**	**15.48**	**582,269**	**10.47**	**222,625**	**4.00**

Table 4.12
The coverage of the ESTCWL and the CEFR-mapped SAWL over the SAJ corpus

Levels of words	Running words		Groups	
	No. of running word	Percent	No. of Groups	Percent
1 ESTCWL	167,387	3.01%	587	0.53%
2 SAWL-CEFR	551,632	9.93%	853	0.77%

4.3.3 Discussion

4.3.3.1 EAP coursebook

Previous studies on coursebooks of English for Academic Purposes (EAP) (e.g. Coxhead, Yen Dang, & Mukai, 2017; Miller, 2011) suggested that many commercially available EAP coursebooks do not represent the authentic language in university textbooks. The coursebooks normally present large amounts of high frequency words. The present study also comes up with similar results. The ESTCWL contains 1,050 target words derived from a series of five EST coursebooks. It was found that the AWL accounts for only 33% (349 words) of the list. However, the

ESTCWL was made from five coursebooks, which implies that each book might contain only 70 AWL words. Normally, one coursebook would have 248-391 of the AWL word families (Miller, 2011). The coursebook used for making the ESTCWL might not be a good textbook for academic vocabulary learning.

4.4 Summary

This chapter has reported the findings of this study. Based on the findings, the SAWL is established as an academic word list of 11 natural science disciplines. The SAWL contains 853 word families and it accounts for 15.5% coverage of the running words in SAJ corpus. The list is also divided to 15 sublists according to frequency. To facilitate the users who are familiar with the CEFR, the most frequent family members of SAWL have been mapped with six CEFR levels. The collocations of the SAWL are also presented in Appendix F. This specialized academic word list is useful for EFL science students, teachers, material developers, test developers, etc. Chapter 5 will present the conclusion of this study, recommendations for further studies, and pedagogical implications.

CHAPTER 5
CONCLUSIONS AND RECOMMENDATIONS

The present study was aimed to identify specialized academic words frequently occur in natural science disciplines. The series of Science Academic Word List (SAWL) has been established based on the findings from this study. This chapter is divided into three sections. The first section presents the summary of the study in accordance with the four research questions. The second section discusses the pedagogical implications of this study. The series of SAWL and the SAWL website are introduced. Finally, the limitations and recommendations for further studies are presented.

5.1 Conclusions

The first two research questions attempted to search for specialized academic words in natural science disciplines that occur outside the GSL (West, 1953) but may coincide with the AWL (Coxhead, 2000). The corpus-based approach and the expert-judged approach were employed in this study and the SAWL was established as a result. The SAWL contains 853 word families of specialized academic words in natural science disciplines. All of them are outside the GSL while 421 word families correspond with the AWL. Although the SAWL contains more word families than expected in Chapter 3 (400-500 word families), the frequency-based SAWL sublists are provided to assist the users in searching for the appropriate words to teach. The components of the SAWL will be elaborated in section 5.3 of this chapter.

The next research question tried to propose another method of choosing suitable words to teach. The process of mapping the 853 SAWL words with the CEFR levels was demonstrated. The findings revealed that the majority of SAWL words are mapped with B1, B2, and C1 – 158 words, 350 words, and 171 words respectively. The proportion of SAWL words at each CEFR level is in agreement with the AWL words when they were categorized into CEFR levels (Capel, 2012). The CEFR-mapped

SAWL sublists are also presented in this study together with the useful lexical information.

The last research question sought to find the distinct differences between the SAWL and other related word lists. First, it was compared with the SWL (Coxhead & Hirsh, 2007) and the results yielded that the SAWL performs better than the SWL in journal articles related to natural science disciplines. Second, the SAWL was tested against the EST coursebook word list. Although the two word lists are not equivalent, the results imply that the SAWL is more useful. It is confirmed that the SAWL can serve as a good resource for designing syllabuses, developing materials, or planning curricula specifically for EFL natural science students.

5.2 Recommendations for further research

Future research of specialized academic word lists can be conducted to address the following issues. First, one of the limitations of this study was that the SAJ corpus was compiled from only one text type, i.e. journal articles. This limits the usefulness and implications of the SAWL. It is recommended that the future studies include other text types such as textbooks, conference papers, reports, and theses. Second, although this study provides information about collocations of SAWL words, it is not the focus of this study. All 1,388 SAWL collocations (presented in Appendix F) are based on the Academic Collocation List (Ackermann & Chen, 2013). Therefore, some of them might not occur in the SAJ corpus or in science-based written texts. Systematic investigations in this area are needed. Finally, this study has established the criteria for mapping specialized academic words with the CEFR levels. Previous studies in this area are rarely found. Future studies can be done to verify CEFR levels of the SAWL words and to verify the criteria proposed by this study.

5.3 Pedagogical implications of the study

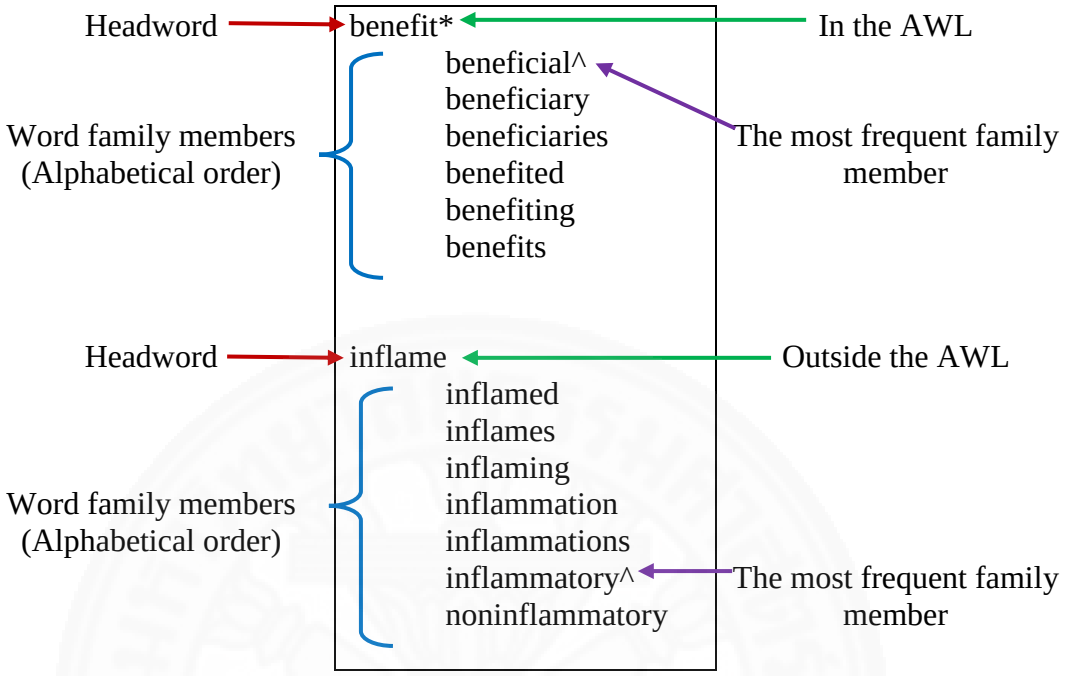
The results of this study have several pedagogical implications. As the SAWL provides a high coverage of science English in research articles, it should be a good resource for students and teachers of science English, syllabus designers, and material developers. Several learning tools have been created to serve pedagogical implications of the study. In the following parts, the structure of the complete SAWL word families will be explained. Next, the frequency-based SAWL sublists, the CEFR-based SAWL sublists, and the SAWL collocation lists will be described.

5.3.1 The complete SAWL

The complete SAWL is presented in Appendix B. The list contains 853 headwords and their family members. As discussed earlier in this chapter, 421 SAWL word families are shared with the AWL (Coxhead, 2000). As shown in Figure 5.1, the headwords of these 421 families are marked with an asterisk (*). The most frequent members in each word family are marked with a caret (^). The word frequency is not presented in number here, but they are presented previously in Figure 4.1 in Chapter 4. The family members of the complete SAWL are based on the BNC/COCA word family list (Nation, 2012). The members are published in alphabetical order. It should be remarked that some members of the SAWL word families (approximately 14%) do not occur in the SAJ corpus.

This complete list can be used in at least three different methods. First, the users may primarily focus on the 421 AWL words before learning the 432 non-AWL words. Second, the users may pay more attention to the most frequent family members because they are used more often in natural science disciplines than other members and provide high learning return. Finally, the main emphasis could be placed on learning the headwords because they provide the central meaning concept of the word family members.

Figure 5.1
Structure of headwords and word family members of the complete SAWL



5.3.2 The frequency-based SAWL sublists

The 853 SAWL word families are divided into 15 sublists based on their frequency, as shown in Appendix C. Sublists 1-14 contain 60 headwords each and the last sublist has 13 headwords. The words in each list are also arranged by frequency. That means the first word in the sublist is the most frequent word. According to Table 5.1, the first five sublists provide 11% coverage of the whole SAJ corpus while the remaining 10 sublists altogether account for 4.25% coverage only. Thus, it is recommended that the users pay attention to the first five sublists because knowing these 300 words makes substantial returns.

Table 5.1
The frequency-based SAWL sublists and coverage in the SAJ corpus

Levels	No. of Words	Coverage in the SAJ
Sublist 01	60	4.75%
Sublist 02	60	2.41%
Sublist 03	60	1.70%
Sublist 04	60	1.31%
Sublist 05	60	1.06%
Sublist 06	60	0.88%
Sublist 07	60	0.72%
Sublist 08	60	0.60%
Sublist 09	60	0.50%
Sublist 10	60	0.46%
Sublist 11	60	0.35%
Sublist 12	60	0.29%
Sublist 13	60	0.24%
Sublist 14	60	0.19%
Sublist 15	13	0.02%
Total	**853**	**15.50%**

5.3.3 The CEFR-mapped SAWL sublists

The 853 most frequent members in the complete SAWL (also marked with a caret ^ in Appendix B) have been mapped with 6 levels of the CEFR. It is not assumed that the identified level is the level of the whole word family, but it can be used as a rough guideline for choosing the words to teach. Nonetheless, the 853 words are classified into seven sublists, i.e. A1, A2, B1, B2, C1, C2, and UL (unlisted), as shown in Appendix E. As mentioned in Chapter 4, the CEFR-mapped sublists are beneficial for the users who aim to focus on specialized academic words that are suitable for the target proficiency levels. For example, for B1 learners, the attention might be primarily paid to 158 words in the B1 sublist and their family members. Appendix D, which shows the parts of speech of the target words frequently used in the SAJ corpus, might be consulted. If the learners become confident in using these words, the attention can be turned to 350 words in the B2 sublist and so on.

5.3.4 The SAWL collocation list

To emphasize the importance of learning collocations, the list of SAWL collocations are provided in Appendix F. Part 1 shows the list of those 234 SAWL words. Every word is marked with any of the symbols - “1st”, “2nd”, or “1st/2nd” - to show its position in the collocations. Part 2 contains 1,390 collocations

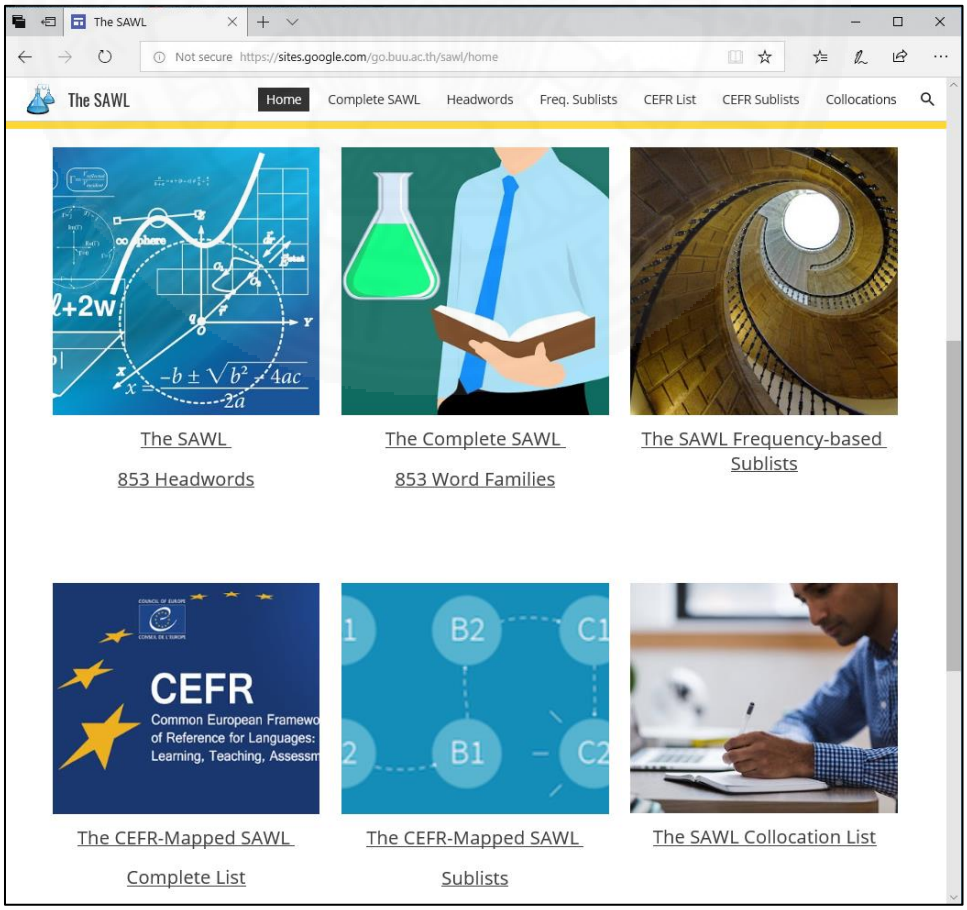
of the 234 SAWL words coincided with the Academic Collocation List (ACL) (Ackermann & Chen, 2013).

To use this list, the users might look for the target word in Part 1. If the word exists, they should look at the given symbol and search for the collocations in Part 2. For example, the word *"fraction"* is in Part 1 of the collocation list and it is tagged with "2nd", which means this word is used in the second position of the collocations. The collocation of this word is presented in Part 2 of the collocation list and it shows that this word is a noun and co-occurs with the adjective *"small"*. The combinations that also exist in the SAJ corpus are tagged with "sci" shown in the last column of the list. The SAWL collocation list can be used as one guideline for learning collocations frequently used in science disciplines.

5.3.5 The SAWL website

The complete SAWL and its descendant word lists are available online at the SAWL website (https://sites.google.com/go.buu.ac.th/sawl). In the first version of the website, users can download six different types of SAWL, i.e. the complete SAWL, the SAWL headwords, the SAWL frequency-based sublists, the CEFR-mapped SAWL, the CEFR-mapped SAWL sublists, and the SAWL collocation list. Interactive learning tools, such as online flashcard program and Moodle, will be added to the website soon. It is expected that the SAWL website will increase usefulness of the lists and support learners, teachers, and other researchers. The interface of the web is shown in Figure 5.2.

Figure 5.2
The SAWL website (https://sites.google.com/go.buu.ac.th/sawl)



5.4 Summary

This chapter presents the conclusions of the study and provides recommendations for future research in word list making and vocabulary learning area. Although several discipline-specific word lists have been continuously developed, there is still a need for more research in this area. This study ends with several pedagogical implications of the SAWL in order to insist that 853 word families in the core SAWL list and other supplementary lists are beneficial and easy to use.

REFERENCES

Ackermann, K., & Chen, Y. H. (2013). Developing the Academic Collocation List (ACL) – A corpus-driven and expert-judged approach. *Journal of English for Academic Purposes, 12*(4), 235-247. doi:https://doi.org/10.1016/j.jeap.2013.08.002

Ahmad, J. (2012). Stylistic features of scientific English: A study of scientific research articles. *English Language and Literature Studies, 2*(1), 47-55.

Anthony, L. (2014). AntWordProfiler (Version 1.4.1) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.laurenceanthony.net/.

Anthony, L. (2015). TagAnt (Version 1.1.0) [Computer Software]. Tokyo: Waseda University. Available from http://www.laurenceanthony.net/software.

Anthony, L. (2018). AntConc (Version 3.5.7) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.laurenceanthony.net/.

Aydın, B., Unver, M. M., Alan, B., & Saglam, S. (2017). Combining the old and the new: Designing a curriculum based on the Taba model and the global scale of English. *Journal of Language & Linguistics Studies, 13*(1), 304-320.

Bauer, L., & Nation, P. (1993). Word Families. *International Journal of Lexicography, 6*(4), 253-279. doi:10.1093/ijl/6.4.253

Benigno, V., & de Jong, J. (2017). *Developing the GSE Vocabulary*. London: Pearson.

Bennett, G. R. (2010). *Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers*. Michigan: University of Michigan Press.

Biber, D. (2006). *University Language : A Corpus-based Study of Spoken and Written Registers*. Amsterdam: John Benjamins Publishing Co.

Boonyos, K. (2014). *Establishment of a Microbiology Academic Word List.* (Thesis), Thammasat University.

Brezina, V., & Gablasova, D. (2015). Is There a Core General Vocabulary? Introducing the New General Service List. *Applied Linguistics, 36*(1), 1-22. doi:10.1093/applin/amt018

British Council. (2013). *The English Effect*. London: British Council.

Browne, C. (2013). The new general service list version 1.01: Getting better all the time. *Korea TESOL Journal, 11*(1), 35-50.

Cambridge University Press. (2013). *Introductory Guide to the Common European Framework of Reference (CEFR) for English Language Teachers*. Cambridge: Cambridge University Press.

Campion, M. E. (1971). *The academic vocabulary list / [by] Mary E. Campion and Warwick B. Elley*. [Wellington]: New Zealand Council for Educational Research.

Capel, A. (2010). A1–B2 vocabulary: insights and issues arising from the English Profile Wordlists project. *English Profile Journal, 1*(1), 1-11. doi:10.1017/S2041536210000048

Capel, A. (2012). Completing the English Vocabulary Profile: C1 and C2 vocabulary. *English Profile Journal, 3*(1), 1-14. doi:10.1017/S2041536212000013

Caro, K., & Mendinueta, N. R. (2017). Lexis, lexical competence and lexical knowledge: a review *Journal of Language Teaching and Research, 8*(2), 205-213.

Chanasattru, S., & Tangkiengsirisin, S. (2016). Developing of a high frequency word list in Social Sciences. *Journal of English Studies, 11*, 41-87.

Chung, T. M., & Nation, P. (2004). Identifying technical vocabulary. *System, 32*(2), 251-263. doi:https://doi.org/10.1016/j.system.2003.11.008

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Council of Europe. (2014, 9 January 2017). Reference Level Descriptions (RLD) for national and regional languages. Retrieved from https://www.coe.int/t/dg4/linguistic/DNR_EN.asp

Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly, 34*(2), 213-238. doi:doi:10.2307/3587951

Coxhead, A. (2011). The Academic Word List 10 Years On: Research and Teaching Implications. *TESOL Quarterly, 45*(2), 355-362. doi:10.5054/tq.2011.254528

Coxhead, A. (2014). *New Ways in Teaching Vocabular*. Virginia: TESOL Press

Coxhead, A. (2016). Reflecting on Coxhead (2000), “A New Academic Word List”. *TESOL Quarterly, 50*(1), 181-185. doi:10.1002/tesq.287

Coxhead, A. (2018). *Vocabulary and English for Specific Purposes Research: Quantitative and Qualitative Perspectives*. Oxon: Taylor & Francis.

Coxhead, A., & Demecheleer, M. (2018). Investigating the technical vocabulary of Plumbing. *English for Specific Purposes, 51*, 84-97. doi:https://doi.org/10.1016/j.esp.2018.03.006

Coxhead, A., & Hirsh, D. (2007). A pilot science-specific word list. [A pilot science-specific word list]. *Revue française de linguistique appliquée, XII*(2), 65-78.

Coxhead, A., & Nation, P. (2001). The specialised vocabulary of English for academic purposes. In J. Flowerdew & M. Peacock (Eds.), *Research Perspectives on English for Academic Purposes* (pp. 252-267). Cambridge: Cambridge University Press.

Coxhead, A., Yen Dang, T. N., & Mukai, S. (2017). Single and multi-word unit vocabulary in university tutorials and laboratories: Evidence from corpora and textbooks. *Journal of English for Academic Purposes, 30*, 66-78. doi:https://doi.org/10.1016/j.jeap.2017.11.001

Crystal, D. (2012). *English as a Global Language*. Cambridge: Cambridge University Press.

Dang, T. N. Y., Coxhead, A., & Webb, S. (2017). The Academic Spoken Word List. *Language Learning*, n/a-n/a. doi:10.1111/lang.12253

Dang, T. N. Y., & Webb, S. (2016). Making an essential word list. In Nation (Ed.), *Making and using word lists for language learning and testing* (pp. 153–167). Amsterdam: John Benjamins.

Davis, M., Davis, K. J., & Dunagan, M. M. (2012). *Scientific Papers and Presentations*. London: Academic Press.

Figueras, N. (2012). The impact of the CEFR. *ELT Journal, 66*(4), 477-485. doi:10.1093/elt/ccs037

Folse, K. (2010). Is Explicit Vocabulary Focus the Reading Teacher's Job? *Reading in a Foreign Language, 22*(1), 139-160.

Fries, C. C. (1940). *English Word Lists: A Study of Their Adaptability for Instruction*. Washington, D.C.: American Council on Education.

Galloway, N., Kriukow, J., & Numajiri, T. (2017). *Internationalisation, Higher Education and the Growing Demand for English: an Investigation into the English Medium of Instruction (EMI) Movement in China and Japan*. London: British Council.

Gardner, D., & Davies, M. (2014). A New Academic Vocabulary List. *Applied Linguistics, 35*(3), 305-327. doi:10.1093/applin/amt015

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

Gholami, J., & Zeinolabedini, M. (2018). Learnability and Teachability Hypothesis. In J. I. Liontas & M. DelliCarpini (Eds.), *The TESOL Encyclopedia of English Language Teaching*. New Jersey: John Wiley & Sons.

Gilner, L. (2011). A primer on the General Service List. *Reading in a Foreign Language, 23*(1), 65-83.

Goldhahn, D., Eckart, T., & Quasthoff, U. (2012). *Building large monolingual dictionaries at the Leipzig Corpora Collection: From 100 to 200 languages.* Paper presented at the In: Proceedings of the 8th International Language Ressources and Evaluation (LREC'12).

Graddol, D. (2006). *English Next*. London: British Council.

Ha, A. Y. H., & Hyland, K. (2017). What is technicality? A Technicality Analysis Model for EAP vocabulary. *Journal of English for Academic Purposes, 28*, 35-49. doi:https://doi.org/10.1016/j.jeap.2017.06.003

Halliday, M. A. K. (2004). *The Language of Science*. London: Continuum.

Hasselgreen, A. (2011). Empowering language professionals: Current ECML projects. *Language Teaching, 44*(3), 400-401. doi:10.1017/S0261444811000231

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). Range and Frequency programs. Retrieved from https://www.victoria.ac.nz/lals/about/staff/paul-nation

Higher Educational Information. (2017). University Students in Thailand of Academic Year 2016 Retrieved from http://www.info.mua.go.th/information/index.php

Hindmarsh, R. (1980). *Cambridge English lexicon*. Cambridge: Cambridge University Press.

Hopkins, A., & Dudley-Evans, T. (1988). A genre-based investigation of the discussion sections in articles and dissertations. *English for Specific Purposes, 7*(2), 113-121. doi:https://doi.org/10.1016/0889-4906(88)90029-4

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. doi:10.1177/1362168813494121

Hsu, W. (2014). Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes, 33*(Supplement C), 54-65. doi:https://doi.org/10.1016/j.esp.2013.07.001

IELTS. (2017). IELTS and the CEFR. Retrieved from https://www.ielts.org/ielts-for-organisations/common-european-framework

Introducing the Global Scale of English Learning Objectives. (2015). London: Pearson Education Limited

Ishikawa, S. (2013). The ICNALE and sophisticated contrastive interlanguage analysis of Asian learners of English. In S. Ishikawa (Ed.), *Learner Corpus Studies in Asia and the World* (Vol. 1, pp. 91-118). Kobe, Japan: Kobe University.

Jones, N. (2014). *Multilingual Frameworks: The Construction and Use of Multilingual Proficiency Frameworks*. Cambridge: Cambridge University Press.

Jones, R., & Tschirner, E. (2015). *A Frequency Dictionary of German: Core Vocabulary for Learners*. Oxon: Taylor & Francis.

Kaeding, F. W. (1898). *Häufigkeitswörterbuch der deutschen Sprache*. Berlin: Selbstverlag des Herausgebers.

Kramsch, C. (2004). Language, thought, and culture. In A. Davies & C. Elder (Eds), *The Handbook of Applied Linguistics*. Oxford: Blackwell.

Koda, K. (2005). *Insights into Second Language Reading: A Cross-Linguistic Approach*. Cambridge: Cambridge University Press.

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.

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*(Supplement C), 42-53. doi:https://doi.org/10.1016/j.jeap.2016.01.008

Lindquist, H. (2009). *Corpus Linguistics and the Description of English*. Edinburgh: Edinburgh University Press.

Liu, J., & Han, L. (2015). A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes, 39*, 1-11. doi:https://doi.org/10.1016/j.esp.2015.03.001

Lynn, R. W. (1973). Preparing Word-Lists: A Suggested Method. *RELC Journal, 4*(1), 25-28. doi:10.1177/003368827300400103

Martyniuk, W., & Noijons, J. (2007). *Executive Summary of Results of a Survey on the Use of the CEFR at National Level in the Council of Europe Member States*. Strasbourg: Council of Europe.

Matsuoka, W. (2012). Searching for the Right Words: Creating Word Lists to Inform EFL Learning. In D. Hirsh (Ed.), *Current Perspectives in Second Language Vocabulary Research*: Peter Lang.

Mauranen, A., Hynninen, N., & Ranta, E. (2010). English as an academic lingua franca: The ELFA project. *English for Specific Purposes, 29*(3), 183-190. doi:https://doi.org/10.1016/j.esp.2009.10.001

McComas, W. F. (2014). *The Language of Science Education: An Expanded Glossary of Key Terms and Concepts in Science Teaching and Learning*. Rotterdam: Sense Publishers.

McEnery, T., & Wilson, A. (1996). *Corpus linguistics*. Edinburgh: Edinburgh University Press.

Miller, D. (2011). ESL reading textbooks vs. university textbooks: Are we giving our students the input they may need? *Journal of English for Academic Purposes, 10*(1), 32-46. doi:https://doi.org/10.1016/j.jeap.2010.12.002

Milton, J. (2009). *Measuring Second Language Vocabulary Acquisition*. Bristol: Multilingual Matters.

Motteram, G. (2013). *Innovations in Learning Technologies for English Language Teaching*. London: British Council.

Nation, P. (2006). How Large a Vocabulary is Needed For Reading and Listening? *Canadian Modern Language Review, 63*(1), 59-82.

Nation, P. (2012). Notes on the BNC/COCA Lists. Retrieved from https://www.victoria.ac.nz/lals/about/staff/paul-nation

Nation, P. (2013). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

Nation, P. (2014). How much input do you need to learn the most frequent 9,000 words? *Reading in a Foreign Language, 26*(2), 1-16.

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

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

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

Nation, P., & Heatley, A. (2004). The Range Programme with British National Corpus word liist. Retrieved from https://www.victoria.ac.nz/lals/about/staff/paul-nation

Nation, P., & Kyongho, H. (1995). Where would general service vocabulary stop and special purposes vocabulary begin? *System, 23*(1), 35-41. doi:https://doi.org/10.1016/0346-251X(94)00050-G

Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.

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

North, B. (2014). *The CEFR in Practice*. Cambridge: Cambridge University Press.

North, B., Ortega, A., & Sheehan, S. (2010). *A Core Inventory for General English*. London: British Council/EAQUALS.

Northrup, D. (2013). *How English Became the Global Language*. New York: Palgrave Macmillan.

Office of The Basic Education Commision. (2014). *Practical Guidelines for Notification of Ministry of Education of English Education Reform*. Bangkok: Office of The Basic Education Commision, Ministry of Education.

Office of the Education Council. (2017). *National Education Plan B.E. 2560 - 2579*. Bangkok: Office of the Education Council, Ministry of Education.

Office of the Higher Education Commission. (2014). *Study in Thailand 2014*. Bangkok: Office of the Higher Education Commission, Ministry of Education.

Office of the National Economic and Social Development Board. (1961). *the First National Economic and Social Development Plan B.E. 2504-2506* Bangkok: Office of the National Economic and Social Development Board, The Office of the Prime Minister.

Office of the National Education Commission. (2003). *Education in Thailand 2002/2003*. Bangkok: Amarin Printing and Publishing.

Palmer, H. E. (1931). *Second Interim Roport [sic] on Vocabulary Selection submitted to the Eighth Annual Conference of English Teachers*. Tokyo: Institute for Research in English Teaching, Department of Education.

Papageorgiou, S., Tannenbaum, R. J., Bridgeman, B., & Cho, Y. (2015). *The Association Between TOEFL iBT® Test Scores and the Common European Framework of Reference (CEFR) Levels (Research Memorandum No. RM-15-06)*. Princeton, NJ: Educational Testing Service.

Paquot, M. (2007). Towards a productively-oriented academic word list. In J. Walinski, K. Kredens, & S. Gozdz-Roszkowski (Eds.), *Practical Applications in Language and Computers 2005* (pp. 127-140). Frankfurt: Peter Lang.

Philip, G. A. (2007). The Imperial Tongue: English as the Dominating Academic Language. *Economic and Political Weekly*(36), 3608.

Praninskas, J. (1972). *American University Word List*. London: Longman.

Pugsee, P., Limgomolvilas, S., Wudthayagorn, J., & Janpugdee, P. (2017). *A framework for generating ICT word lists.* Paper presented at the 2017 10th International Conference on Ubi-media Computing and Workshops (Ubi-Media), Pattaya, Thailand, 2017.

Quero, B. (2015). *Estimating the vocabulary size of L1 Spanish ESP learners and the vocabulary load of medical textbooks.* Vicotia University of Wellington, New Zealand.

Reeves, C. (2005). *The Language of Science*. Oxon: Routledge.

Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. London: Palgrave Macmillan.

Schmitt, N., Cobb, T., Horst, M., & Schmitt, D. (2015). How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2007). *Language Teaching, 50*(2), 212-226. doi:10.1017/S0261444815000075

Schmitt, N., & Schmitt, D. (2012). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching, 47*(4), 484-503. doi:10.1017/S0261444812000018

Short, D. J., Vogt, M. E., & Echevarria, J. (2010). *The SIOP Model for Teaching Science to English Learners*. New York: Pearson.

Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.

Templeton, S., Bear, D. R., & Johnston, F. R. (2010). *Vocabulary Their Way: Word Study with Middle and Secondary Students*. New York: Pearson.

Thompson, D. K. (1993). Arguing for Experimental “Facts” in Science:A Study of Research Article Results Sections in Biochemistry. *Written Communication, 10*(1), 106-128. doi:10.1177/0741088393010001004

Thorndike, E. L. (1921). *The Teachers Word Book*. New York City: Teachers College, Columbia University.

Thorndike, E. L. (1931). *A Teacher’s Word Book of 20,000 Words*. New York: Teachers College, Columbia University.

Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Teachers College, Columbia University.

Tongpoon-Patanasorn, A. (2018). Developing a frequent technical words list for finance: A hybrid approach. *English for Specific Purposes, 51*, 45-54. doi:https://doi.org/10.1016/j.esp.2018.03.002

UCLES/CUP. (2011). *English Profile: Introducing the CEFR for English*. Cambridge: Cambridge University Press.

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. doi:https://doi.org/10.1016/j.jeap.2013.07.001

Valldepérez Castillo, M. S. (2017). *Oral Discourse in Foreign Language Education. The Speaking Activities in EFL Textbooks and the CEFR.* (Dissertation/Thesis), Universitat de Barcelona.

van Zeeland, H., & Schmitt, N. (2013). Lexical Coverage in L1 and L2 Listening Comprehension: The Same or Different from Reading Comprehension? *Applied Linguistics, 34*(4), 457-479. doi:10.1093/applin/ams074

Wang, J., Liang, S.-l., & Ge, G.-c. (2008). Establishment of a Medical Academic Word List. *English for Specific Purposes, 27*(4), 442-458. doi:https://doi.org/10.1016/j.esp.2008.05.003

Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes, 28*(3), 170-182. doi:https://doi.org/10.1016/j.esp.2009.04.001

Watson Todd, R. (2017). An opaque engineering word list: Which words should a teacher focus on? *English for Specific Purposes, 45*, 31-39. doi:https://doi.org/10.1016/j.esp.2016.08.003

Webb, S., & Nation, P. (2017). *How Vocabulary is Learned*. Oxford: Oxford University Press.

Webb, S., & Rodgers, M. P. H. (2009). The Lexical Coverage of Movies. *Applied Linguistics, 30*(3), 407-427. doi:10.1093/applin/amp010

West, M. (1953). *A General Service List of English Words with Semantic Frequencies*. London: Longman.

Wood, A. (2001). International scientific English: The language of research scientists around the world. In J. Flowerdew & M. Peacock (Eds.), *Research Perspectives on English for Academic Purposes* (pp. 71-83). Cambridge: Cambridge University Press.

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

Yang, M.-N. (2015). A nursing academic word list. *English for Specific Purposes, 37*, 27-38. doi:https://doi.org/10.1016/j.esp.2014.05.003

APPENDICES

APPENDIX A
HEADWORDS OF THE SCIENCE ACADEMIC WORD LIST (SAWL)

absorb
abstract
abundant
academy
access
accommodate
accompany
accurate
acetate
acetic
acetone
achieve
acid
acquire
activate
acute
adapt
additives
adequate
adhesion
adjacent
adjust
administrate
adult
adverse
affect
affinity
agar
agarose
aggregate
aid
albumin
alcohol
align
alkaline
allele
allocate
alter
alternative
ambience
amino
ammonia
amplify
amplitude
anaerobic
analogy
analyse
anneal
annual
antimicrobial
apoptosis
apparent
append
approach
appropriate
approximate
aqueous
arbitrary
architecture
area
aromatic
array
aspect
assay
assemble
assess
assign
assist
assume
atmosphere
atom
attach
attain
attitude
attribute
author
automate
available
aware
bacillus
bacterium
barrier
basal
baseline
batch
bead
benefit
bias
bioactive
biochemical
biology
biomass
biomedical
biosynthetic
biotechnology
bond
bovine
breakdown
breast
brief
buffer
bulk
calcium
calibrate
cancer
candidate
capable
capacity
capillary
capture
carbohydrate
carbon
cardiac
cardiovascular
cascade
catalyse
category
cavity
cellular
cellulose
centrifuge
cerevisiae
challenge
chamber
channel
chemical
chemistry
chip
chloride
cholesterol
chromatography
chromosome
chronic
clarify
classic
climate
clinic
clone
coating
code
coefficient
coincide
colon
column
comment
communicate
community
compatible
compensate
complement
complex
component
composition
compound
comprehensive
comprise
compute
concept
conclude
concurrent
conduct
confer
configure
confirm
confocal
conform
consequent
conserve
considerable
consist
constant
constitute

constrain
construct
consume
contact
contaminate
context
contrast
contribute
convene
converse
convert
cooperate
coordinate
core
correlate
correspond
couple
covalent
create
criteria
crucial
crude
crystal
culture
cumulative
cycle
cysteine
dash
data
database
dataset
decade
decline
decompose
defect
deficiency
define
definite
degrade
dehydrogenased
demonstrate
denote
dense
dependence
depict
deplete
deposit
deposition
depress
design
despite
detect
developmental
device
devote
diagram
diet
differential
differentiate
diffract
diffuse
digest
digital
dilute
dimension
discrepancy
discrete
discriminate
disperse
displace
display
disrupt
dissolution
dissolve
distil
distinct
distribute
diverge
diverse
document
domain
dominate
donor
downstream
drama
droplet
drug
dual
duration
dye
dynamic
ecological
economy
ecosystem
edit
efficiently
electrode
electron
electrophoresis
electrostatic
element
elemental
elevate
eliminate
elongate
embed
emerge
emit
emphasis
empirical
enable
encode
encounter
endogenous
energy
engineered
enrich
ensure
entity
environment
enzymatic
enzyme
epithelial
equate
equip
equivalent
error
establish
ester
estimate
ethanol
evaluate
evaporate
eventual
evident
evolve
exceed
excitation
exclude
exhibit
exogenous
expand
expert
explicit
exploit
exponential
expose
external
extract
fabricate
facilitate
factor
favorable
favorably
feasible
feature
feedback
feeding
ferment
fibre
file
filter
final
finite
flexible
fluctuate
fluid
fluorescent
flux
focus
formula
fraction
fragment
framework
frequencies
function
functionalize
fundamental
fungus
furthermore
fuse
gel
gene
generate
generation
genetic
genome
genotype

genus
geography
geometry
germ
globe
glucose
glycerol
goal
grade
graph
grid
guarantee
gut
height
hence
hierarchy
highlight
hybrid
hydrogen
hydrolysis
hydrophilic
hydrophobic
hydroxyl
hypothesis
identical
identify
illustrate
image
imaged
immobilize
immune
impact
implement
implicate
imply
impose
incidence
incorporate
incubate
index
indicate
individual
induce
infect
infer
inflame
infrared
inherent
inhibit
initial
initiate
inject
injure
innovate
input
insert
inset
insight
instance
instruct
intact
intake
integral
integrate
intense
interact
interestingly
interface
interior
intermediate
internal
interpret
interval
intervene
intestine
intrinsic
invasive
inverse
investigate
involve
ion
isolate
item
kernel
kidney
kinase
kinetic
kit
label
laboratory
lactic
laser
latent
lateral
lattice
layer
linear
link
lipid
liver
localise
locate
locus
longitudinal
loop
lysine
magnesium
magnet
magnify
magnitude
maintain
major
manipulate
manual
mapping
margin
marine
mathematics
matrix
mature
maximal
maximise
mechanism
media
median
mediate
medical
medium
membrane
mesh
metabolic
metabolism
metabolite
methanol
methionine
method
micro
microbe
microorganism
microscope
migrate
minimal
minimise
minimum
minor
mitochondria
mobile
mode
modify
molar
molecule
monitor
morphology
mortal
motive
mount
muscle
mutant
mutate
nanoparticle
negate
negligible
network
neural
neutral
nevertheless
nitrogen
node
norm
normal
notion
novel
nuclear
nucleotide
nucleus
null
nutrient
nutrition
objective
obtain
obvious
occupy
occur
odd
online
onset
optic
optimal

optimise
optimum
option
oral
organic
organism
orient
outcome
output
oven
overall
overlap
overnight
overview
oxidant
oxide
oxidise
oxygen
panel
parallel
parameter
participate
partner
passive
pathogen
pathogenic
pathway
patients
peak
penetrate
peptide
perceive
percent
period
periphery
peroxide
persist
perspective
pharmaceutical
phase
phenomenon
phenotypic
phosphate
phylogenetic
physical
physiological
plasma
plastic
platform
plot
plus
polar
poly
polymer
polymerase
polynomial
pooled
pore
portion
pose
positive
posterior
potassium
potent
potential
precede
precipitate
precise
precursor
predict
predominant
preliminary
presume
prevalent
previous
primary
prime
primer
principal
principle
prior
probe
proceed
process
profile
progression
project
proliferate
proline
promote
propagate
proportion
protease
protein
protocol
publication
publish
pulse
purchase
purify
purity
pursue
putative
quantify
radical
random
range
ratio
rational
react
reagent
receptor
recover
redox
reference
refine
regime
region
regress
regulate
reject
relax
release
relevant
rely
remove
replicate
require
research
reservoir
reside
residue
resolve
resonance
resource
respond
restore
restrict
resuspend
retain
reveal
reverse
revise
robust
role
rotate
route
routine
saline
saturate
scaling
scan
scenario
scheme
score
seasonal
secrete
section
sector
seek
segment
select
sensing
sensor
sequence
series
serine
serum
setup
sex
shift
significant
silica
silicon
similar
simulate
simultaneous
site
skeletal
sodium
software
sole
soluble
solvent
somewhat
source
spatial
species

specific
specify
spectra
spectre
sphere
spontaneous
stable
static
statistic
status
storage
straightforward
strain
strand
strategy
stress
structure
subsequent
subset
substitute
successor
sufficient
sum
summary
superior
supernatants
supplement
suppress
survey
survive
susceptible
suspend
sustain
switch
symbol
symmetry
symptom
synergistic
synthesis
synthetic
tank
target
task
taxonomy
team
technical
technique
technology
template
temporal
terminate
tertiary
text
theory
therapeutic
therapy
thereby
thermal
threshold
tissue
tolerance
topic
toxic
toxin
trace
tract
tradition
transcript
transfer
transform
transient
transit
transmit
transport
trend
triangle
trigger
triple
triplicate
tumour
tyrosine
ultimate
unclear
undergo
underlie
uniform
unique
untreated
uptake
urea
utilise
valid
vary
vascular
velocity
verify
version
versus
vertical
via
viable
video
violate
virtual
virus
vitamin
volatile
volt
volume
wavelength
weighted
whereas
widespread
worldwide
yeast
zinc
zone

APPENDIX B
THE COMPLETE 853 WORD FAMILIES OF SAWL

*Note: Modified from Nation, 2012; ^ highest frequency word in the family; * coinciding with AWL (Coxhead, 2000)*

absorb*
- absorbance
- absorbed
- absorbency
- absorbent
- absorber
- absorbers
- absorbing
- absorbs
- absorption^
- absorptions
- absorptive

abstract^
- abstraction
- abstractions
- abstractly
- abstracts

abundant
- abundance^
- abundances
- abundantly

academy*
- academia
- academic^
- academically
- academics
- academies

access^*
- accessed
- accesses
- accessibility
- accessible
- accessing
- inaccessible

accommodate^
- accommodated
- accommodates
- accommodating
- accommodation

accompany*
- accompanied^
- accompanies
- accompaniment
- accompanying
- unaccompanied

accurate*
- accuracy^
- accurately
- inaccuracy
- inaccuracies
- inaccurate

acetate^
- acetates

acetic^

acetone^
- acetonin

achieve*
- achievable
- achieved^
- achievement
- achievements
- achieves
- achieving

acid^
- acidic
- acidity
- acidly
- acids

acquire*
- acquired
- acquires
- acquiring
- acquisition^
- acquisitions

activate
- activated
- activates
- activating
- activation^
- activator
- activators
- inactivation
- unactivated

acute^
- acutely
- acuteness

adapt*
- adaptability
- adaptable
- adaptation
- adaptations
- adapted^
- adapting
- adaptive
- adapts

additives^*

adequate^
- adequacy
- adequately
- inadequacies
- inadequacy
- inadequate
- inadequately

adhesion^*
- adhesions

adjacent^

adjust*
- adjusted^
- adjusting
- adjustment
- adjustments
- adjusts
- readjust
- readjusted
- readjusting
- readjustment
- readjustments
- readjusts

administrate*
- administrates
- administration^
- administrations
- administrative
- administratively
- administrator
- administrators

adult^
- adulthood
- adults

adverse^*
- adversely
- adverseness
- adversities
- adversity

affect^
- affected
- affecting
- affective
- affectively
- affects
- unaffected

affinity^
- affinities

agar^

agarose^

aggregate*
- aggregated
- aggregates
- aggregating
- aggregation^

aid^
- aided
- aiding
- aids
- unaided

albumin^
- albuminous
- albumins

alcohol^
- alcoholic
- alcoholics
- alcoholism
- alcohols

align
- aligned
- aligner
- aligners
- aligning
- alignment^
- alignments
- aligns

alkaline^
- alkalinities
- alkalinity

allele^
- alleles
- allelic

allocate*
- allocated
- allocates
- allocating
- allocation^
- allocations

alter*
- alterable
- alteration
- alterations
- altered^
- altering
- alternate
- alternating
- alters
- unalterable
- unaltered

alternative^
- alternatively
- alternatives

ambience
- ambiences
- ambient^

amino^

ammonia^
- ammonium

amplify
- amplification^
- amplifications
- amplified
- amplifier
- amplifiers
- amplifies
- amplifying

amplitude^
- amplitudes

anaerobic^
- anaerobically

analogy*
- analogies
- analogous^

analyse*
- analysed
- analyser
- analysers
- analyses
- analysing

analysis^

analyst
- analysts

analytic
- analytical
- analytically

analyze
- analyzed
- analyzes
- analyzing

anneal*
- annealed
- annealer
- annealers
- annealing^
- anneals

annual^
- annually

antimicrobial^

apoptosis^*

apparent^
- apparently

append*
- appendix^
- appended
- appends
- appending
- appendices
- appendixes

approach^*
- approachable
- approached
- approaches
- approaching
- unapproachable

appropriate^
- appropriacy
- appropriately
- appropriateness
- inappropriacy
- inappropriate
- inappropriately

approximate*

approximated
approximately^
approximates
approximating
approximation
approximations
aqueous^*
arbitrary^
arbitrariness
arbitrarily
architecture^*
architectural
architecturally
architectures
area^
areas
aromatic^
aromatically
array^
arrayed
arraying
arrays
aspect*
aspects^
assay^
assayed
assayer
assayers
assaying
assays
assemble*
assembled
assembles
assemblies
assembling
assembly^
assess*
assessable
assessed^
assesses
assessing
assessment
assessments
reassess
reassessed
reassessing
reassessment
unassessed
assign*
assigned^
assigning
assignment
assignments
assigns
reassign
reassigned
reassigning
reassigns
unassigned
assist*
assistance
assistant
assistants
assisted^
assisting
assists
unassisted
assume^
assumed
assumes
assuming
assumption
assumptions
atmosphere^
atmospheres
atmospheric
atom
atomic
atomise
atomised
atomiser
atomisers
atomises
atomising
atomism
atomize
atomized
atomizer
atomizers
atomizes
atomizing
atoms^
subatomic
attach*
attached^
attaches
attaching
attachment
attachments
unattached
attain*
attainable
attained^
attaining
attainment
attainments
attains
unattainable
attitude^
attitudes
attribute*
attributable
attributed^
attributes
attributing
attribution
author*
authored
authoring
authors^
authorship
automate*
automatic
automated^
automates
automating
automatically
automation
available^*
availability
unavailable
aware^
awareness
unaware
bacillus^
bacilli
bacterium
bacteria^
bacterial
barrier^
barriers

basal^
basally
baseline^
baselines
batch^
batches
bead
beaded
beadier
beadiest
beading
beads^
beady
benefit*
beneficial^
beneficiary
beneficiaries
benefited
benefiting
benefits
bias^
biased
biases
biasing
unbiased
bioactive^
biochemical^
biochemically
biochemicals
biology^
biologies
biologist
biologists
biomass^
biomasses
biomedical^
biosynthetic
biosynthesis^
biosynthetically
biotechnology^
biotechnologies
bond*
bonded
bonding
bonds^
bovine^
breakdown^
breakdowns
breast^
breasted
breasting
breasts
brief*
brevity
briefed
briefing
briefly^
briefs
buffer^*
buffered
buffering
buffers
bulk^
bulky
calcium^
calibrate^
calibrated
calibrates
calibrating
calibration
calibrations
cancer^
cancerlike
cancerous
cancers
candidate^*
candidacies
candidacy
candidates
candidature
capable^*
capabilities
capability
incapable
capacity^
capacities
incapacitate
incapacitated
capillary^
capillaries
capture^
captured
captures
capturing
uncaptured
carbohydrate^
carbohydrates
carbon^
carbons
cardiac^
cardiovascular^
cascade^
cascaded
cascades
cascading
catalyse
catalysed
catalyses
catalysing
catalytic^
catalyze
catalyzed
catalyzes
catalyzing
category*
categories^
categorisation
categorise
categorised
categorises
categorising
categorization
categorized
categorizes
categorizing
cavity^
cavities
cellular^
extracellular
intercellular
intracellular
cellulose^
celluloses
centrifuge
centrifugal
centrifugally
centrifugals
centrifugation^
centrifuged
centrifuges
centrifuging

cerevisiae^*
challenge^
    challenged
    challenger
    challengers
    challenges
    challenging
chamber^*
    chambered
    chambers
channel^*
    channelled
    channelling
    channels
chemical^
    chemically
    chemicals
chemistry^
chip^
    chipped
    chipping
    chippings
    chips
chloride^
    chlorides
cholesterol^
chromatography^
    chromatographic
    chromatographically
    chromatographies
chromosome^
    chromosomal
    chromosomes
chronic^*
    chronically
clarify^
    clarification
    clarified
    clarifies
    clarifying
    clarity
classic*
    classical^
    classics
climate^
    climates
    climatic
    climatically
    climatological
clinic
    clinical^
    clinically
    clinician
    clinicians
    clinics
    nonclinical
clone
    cloned
    cloner
    cloners
    clones^
    cloning
coating^*
    coatings
code^
    coded
    codes
    coding
coefficient^*
    coefficients
coincide^
    coincided
    coincides
    coinciding
    coincidence
    coincidences
    coincident
    coincidental
colon^
    colons
    semicolon
    semicolons
column^
    col
    cols
    columnar
    columned
    columns
comment*
    commentaries
    commentary
    commentator
    commentators
    commented
    commenting
    comments^
communicate*
    communicable
    communicated
    communicates
    communicating
    communication^
    communications
    communicative
    communicatively
    uncommunicative
community^*
    communities
compatible^*
    compatibility
    incompatibility
    incompatible
compensate^
    compensated
    compensates
    compensating
    compensation
    compensations
    compensatory
complement*
    complementary^
    complemented
    complementing
    complements
complex^
    complexities
    complexity
component*
    componentry
    components^
composition^
    compositional
    compositions
compound*
    compounded
    compounding
    compounds^
comprehensive^
    comprehensively
comprise*
    comprised^

comprises
comprising
compute*
computation
computational^
computations
computable
computer
computed
computerised
computers
computing
concept^
conception
concepts
conceptual
conceptualisation
conceptualise
conceptualised
conceptualises
conceptualising
conceptually
conclude*
concluded
concludes
concluding
conclusion^
conclusions
conclusive
conclusively
inconclusive
inconclusively
concurrent^
concurrently
conduct*
conducted^
conducting
conducts
confer*
conference^
conferences
conferred
conferring
confers
configure
configurable
configuration^
configurations
configured
configures
configuring
confirm*
confirmation
confirmed^
confirming
confirms
confocal^
conform*
conformable
conformability
conformance
conformation^
conformed
conforming
conformist
conformists
conformity
conforms
nonconformist
nonconformists
nonconformity
consequent*
consequence
consequences
consequently^
conserve*
conservation
conservationism
conservationist
conservationists
conservator
conservators
conserved^
conserves
conserving
considerable^
considerably
consist*
consisted
consistency
consistent^
consistently
consisting
consists
inconsistencies
inconsistency
inconsistent
constant^
constancy
constantly
constants
inconstancy
inconstantly
constitute*
constituencies
constituency
constituent
constituents^
constituted
constitutes
constituting
constitution
constitutions
constitutional
constitutionally
constitutive
unconstitutional
constrain*
constrained
constraining
constrains
constraint
constraints^
unconstrained
construct^
constructed
constructing
construction
constructions
constructive
constructs
reconstruct
reconstructed
reconstructing
reconstruction
reconstructs
consume*
consumed
consumer
consumers
consumes

consuming
consumption^
contact^
contactable
contacted
contacting
contacts
contaminate*
contaminated
contaminates
contaminating
contamination^
contaminations
uncontaminated
context^*
contexts
contextual
contextualise
contextualised
contextualising
uncontextualised
contextualize
contextualized
contextualizing
uncontextualized
contrast^*
contrasted
contrasting
contrastive
contrasts
contribute^
contributed
contributes
contributing
contribution
contributions
contributor
contributors
convene*
convention
convenes
convened
convening
conventional^
conventionally
conventions
unconventional
converse*
conversely^
convert*
conversion^
conversions
converted
convertible
converting
converts
cooperate*
cooperated
cooperates
cooperating
cooperation^
cooperative
cooperatively
coordinate*
coordinated
coordinates^
coordinating
coordination
coordinator
coordinators
core^
cores
coring
cored
correlate
correlated
correlates
correlating
correlation^
correlational
correlations
uncorrelated
correspond*
corresponded
correspondence
corresponding^
correspondingly
corresponds
couple*
coupled^
coupling
couples
covalent^
create*
created^
creates
creating
creation
creations
creative
creatively
creativity
creator
creators
recreate
recreated
recreates
recreating
criteria^*
criterion
crucial^
crucially
crude^
crudely
crudeness
cruder
crudest
crudities
crudity
crystal^*
crystallisation
crystallise
crystallised
crystallises
crystallising
crystallization
crystallize
crystallized
crystallizes
crystallizing
crystals
culture^
cultural
culturally
cultured
cultures
uncultured
cumulative^*
cumulatively
noncumulative
cycle^

- cycled
- cycles
- cyclic
- cyclical
- cycling

cysteine^
- cysteines

dash*
- dashed^
- dasher
- dashers
- dashes
- dashing

data^

database^
- databases
- dbase
- dbases

dataset^
- datasets

decade*
- decades^

decline^
- declined
- declines
- declining

decompose
- decomposed
- decomposer
- decomposers
- decomposes
- decomposing
- decomposition^

defect
- defective
- defectively
- defectiveness
- defects^

deficiency
- deficiencies
- deficient^

define*
- definable
- defined^
- defines
- defining
- definition
- definitions
- redefine
- redefined
- redefines
- redefining
- undefined

definite^
- definitely
- definitive
- indefinite
- indefinitely

degrade
- degradation^
- degradations
- degraded
- degrades
- degrading

dehydrogenased
- dehydrogenase^

demonstrate*
- demonstrable
- demonstrably
- demonstrated^
- demonstrates
- demonstrating
- demonstration
- demonstrations
- demonstrative
- demonstratively
- demonstrator
- demonstrators

denote^
- denotation
- denotations
- denoted
- denotes
- denoting

dense
- densely
- denseness
- denser
- densest
- densities
- density^

dependence^
- dependences
- interdependence
- interdependences

depict
- depicted^
- depicting
- depiction
- depictions
- depicts

deplete
- depleted
- depletes
- depleting
- depletion^

deposit
- deposited^
- depositing
- depositor
- depositors
- depository
- deposits

deposition^
- depositional
- depositions

depress*
- depressed
- depresses
- depressing
- depression^

design^*
- designed
- designer
- designers
- designing
- designs

despite^

detect*
- detectable
- detected
- detecting
- detection^
- detective
- detectives
- detector
- detectors
- detects

developmental^*

device^
- devices

devote*
- devoted^
- devotedly
- devotes
- devoting
- devotion
- devotions

diagram^
- diagrammatic
- diagrammatically
- diagrammed
- diagramming
- diagrams

diet^
- dietary
- dieted
- dieter
- dieters
- dietician
- dieticians
- dieting
- dietitian
- dietitians
- diets

differential^
- diff
- differentially
- differentials
- diffs

differentiate*
- differentiated
- differentiates
- differentiating
- differentiation^

diffract
- diffractions
- diffraction^
- diffractive

diffuse
- diffused
- diffuser
- diffusers
- diffuses
- diffusing
- diffusion^
- diffusions

digest
- digested
- digestible
- digesting
- digestion^
- digestive
- digestives
- digests
- indigestible
- undigested

digital^
- digitally

dilute
- diluted^
- dilutes
- diluting
- dilution
- dilutions
- undiluted

dimension*
- dimensional^
- dimensions
- multidimensional

discrepancy^*
- discrepancies
- discrepant

discrete^
- discretely
- discretion
- discretionary
- indiscrete
- indiscretion

discriminate*
- discriminated
- discriminates
- discriminating
- discrimination^

disperse
- dispersal
- dispersals
- dispersed
- disperses
- dispersing
- dispersion^

displace*
- displaced
- displacement^
- displaces
- displacing

display^
- displayed
- displaying
- displays

disrupt
- disrupted
- disrupting
- disruption^
- disruptions
- disruptive
- disruptively
- disrupts

dissolution^
- dissolutions

dissolve
- dissolved^
- dissolves
- dissolving

distil*
- distilation
- distilations
- distill
- distillation
- distillations
- distilled^
- distiller
- distilleries
- distillers
- distillery
- distilling
- distills
- distils

distinct^
- distinction
- distinctions
- distinctive
- distinctively
- distinctly
- indistinct
- indistinctly

distribute*
- distributed
- distributing
- distribution^
- distributional
- distributions

- distributive
- distributor
- distributors
- redistribute
- redistributed
- redistributes
- redistributing
- redistribution

diverge
- diverged
- divergence^
- divergences
- divergent
- diverges
- diverging

diverse*
- diversely
- diversification
- diversified
- diversifies
- diversify
- diversifying
- diversity^

document*
- documentation
- documented^
- documenting
- documents

domain^
- domains

dominate*
- dominance
- dominant^
- dominated
- dominates
- dominating
- domination

donor^
- donee
- donees
- donors

downstream^

drama*
- dramas
- dramatic
- dramatically^
- dramatise
- dramatised
- dramatising
- dramatises
- dramatisation
- dramatisations
- dramatist
- dramatists
- dramatization
- dramatizations
- dramatize
- dramatized
- dramatizes
- dramatizing

droplet^
- droplets

drug^
- drugged
- druggie
- druggies
- drugging
- drugs

dual^*
- dualism
- dualisms
- dualist
- dualistic
- dualists
- duality

duration^

dye^
- dyed
- dyeing
- dyer
- dyers
- dyes

dynamic*
- dynamically
- dynamics^

ecological^
- ecologically

economy*
- economic^
- economical
- economically
- economics
- economies
- economist
- economists
- uneconomical

ecosystem
- ecosystems^

edit*
- edited^
- editing
- edition
- editions
- editor
- editorial
- editorials
- editors
- edits

efficiently^

electrode^
- electrodes

electron^
- electrons

electrophoresis^

electrostatic^
- electrostatically
- electrostatics

element*
- elements^

elemental^
- elementally

elevate
- elevated^
- elevates
- elevating
- elevation
- elevations
- elevator
- elevators

eliminate*
- eliminated
- eliminates
- eliminating
- elimination^

elongate
- elongated
- elongates
- elongating
- elongation^

embed
- embedded^

- embedding
- embeds

emerge*
- emerged
- emergence^
- emergent
- emerges
- emerging

emit*
- emission^
- emissions
- emits
- emitted
- emitter
- emitters
- emitting

emphasis^*
- emphasise
- emphasised
- emphasising
- emphasize
- emphasized
- emphasizes
- emphasizing
- emphatic
- emphatically

empirical^*
- empirically
- empiricism

enable^
- enabled
- enables
- enabling

encode
- encoded
- encoder
- encoders
- encodes
- encoding^

encounter*
- encountered^
- encountering
- encounters

endogenous^*
- endogenously

energy^
- energetic
- energetically
- energies

engineered^

enrich*
- enriched
- enriches
- enriching
- enrichment^
- enrichments

ensure^
- ensured
- ensures
- ensuring

entity*
- entities^

environment*
- environmental^
- environmentalist
- environmentalists
- environmentally
- environments

enzymatic^

enzyme^
- enzymes

epithelial^
- epithelia
- epithelium
- epitheliums

equate*
- equated
- equates
- equating

equation^
- equations

equip*
- equipment
- equipped^
- equipping
- equips

equivalent^*
- equivalence

error^
- erroneous
- erroneously
- errors

establish*
- disestablish
- disestablished
- disestablishes
- disestablishing
- disestablishment
- established^
- establishes
- establishing
- establishment
- establishments

ester
- esterase
- esters^

estimate*
- estimated^
- estimates
- estimating
- estimation
- estimations
- overestimate
- overestimated
- overestimates
- overestimating
- underestimate
- underestimated
- underestimates
- underestimating

ethanol^
- ethanols

evaluate*
- evaluated^
- evaluates
- evaluating
- evaluation
- evaluations
- evaluative

evaporate
- evaporated
- evaporates
- evaporating
- evaporation^
- evaporator
- evaporators

eventual*
- eventuality
- eventually^

evident*
- evidenced

evidence^
evidential
evidently
evolve*
evolution^
evolved
evolving
evolves
evolutionary
evolutionist
evolutionists
exceed^
exceeded
exceeding
exceeds
excitation^
excitations
exclude*
excluded^
excludes
excluding
exclusion
exclusionary
exclusionist
exclusions
exclusive
exclusively
exhibit*
exhibited^
exhibiting
exhibition
exhibitions
exhibits
exogenous^
expand*
expanded^
expanding
expands
expansion
expansionism
expansive
expert*
expertise
expertly
experts^
explicit^
explicitly
exploit*
exploitation
exploited^
exploiting
exploits
exponential^
exponentially
expose*
exposed
exposes
exposing
exposure^
exposures
external^
externalisation
externalise
externalised
externalises
externalising
externality
externalization
externalize
externalized
externalizes
externalizing
externally
extract*
extracted
extracting
extraction^
extracts
fabricate*
fabricated^
fabricates
fabricating
fabrication
fabrications
facilitate^
facilitated
facilitates
facilities
facilitating
facilitation
facilitator
facilitators
facility
factor*
factored
factoring
factors^
favorable^
favorably^
feasible^
feasibilities
feasibility
infeasible
unfeasible
feature*
featured
features^
featuring
feedback^
feedbacks
feeding^
ferment
fermentation^
fermentations
fermented
fermenting
ferments
fibre*
fiber^
fibers
fibres
fibrous
file^
filed
files
filing
filter^
filtered
filtering
filters
filtrate
filtrated
filtrating
filtration
filtrations
unfiltered
final*
finalise
finalised
finalises
finalising

finalize
finalized
finalizes
finalizing
finality
finally^
finals
finite^*
infinite
infinitely
flexible^
flexibility
inflexible
inflexibility
fluctuate*
fluctuated
fluctuates
fluctuating
fluctuation
fluctuations^
fluid^
fluidities
fluidity
fluids
fluorescent
fluoresce
fluoresced
fluorescence^
fluorescences
fluorescently
fluoresces
fluorescing
flux^*
fluxes
focus^*
focused
focuses
focusing
focussed
focussing
refocus
refocused
refocuses
refocusing
refocussed
refocusses
refocussing
formula^
formulae
formulas
formulate
formulated
formulating
formulation
formulations
reformulate
reformulated
reformulating
reformulation
reformulations
fraction^
fractional
fractionally
fractions
fragment*
fragmentary
fragmentation
fragmented
fragmenting
fragments^
framework^
frameworks
frequencies^*
function^
functional
functionally
functioned
functioning
functions
functionalize*
functionalized^
functionalizes
functionalizing
fundamental^
fundamentally
fungus*
fungal^
fungi
funguses
furthermore^
fuse
fused
fuses
fusing
fusion^
fusions
gel^
gelled
gelling
gels
gene^
genes
generate*
generated^
generates
generating
generation^
generations
genetic^
genetically
geneticist
geneticists
genetics
genome^
genomes
genomic
genotype
genotypes^
genus^
genuses
geography
geographer
geographers
geographic^
geographical
geographically
geographies
geometry^
geometric
geometrical
geometrically
geometries
germ^
germs
globe*
global^
globally
globalisation
globalization
glucose^
glucoses

glycerol^*
goal^*
goals
grade^
graded
grades
grading
graph^
graphed
graphing
graphs
grid^*
grids
guarantee^
guaranteed
guaranteeing
guarantees
gut^
gutless
gutlessness
guts
gutsier
gutsiest
gutsy
gutted
gutting
height^*
heights
hence^
hierarchy*
hierarchical^
hierarchies
highlight*
highlighted^
highlighting
highlights
hybrid^
hybridisation
hybridisations
hybridise
hybridised
hybridises
hybridising
hybridization
hybridizations
hybridize
hybridized
hybridizes
hybridizing
hybrids
hydrogen^
hydrogenous
hydrogens
hydrolysis^
hydrophilic^
hydrophobic^
hydroxyl^*
hydroxyls
hypothesis^*
hypotheses
hypothesise
hypothesised
hypothesises
hypothesising
hypothesize
hypothesized
hypothesizes
hypothesizing
hypothetical
hypothetically
identical^
identically
identify*
identifiable
identification
identified^
identifies
identifying
identities
identity
unidentifiable
illustrate*
illustrated^
illustrates
illustrating
illustration
illustrations
illustrative
image^
imagery
images
imaged
imaging^
immobilize
immobilisation
immobilise
immobilised
immobiliser
immobilisers
immobilises
immobilising
immobilization
immobilized^
immobilizes
immobilizing
immune^*
immunisation
immunisations
immunise
immunised
immunising
immunities
immunity
immunization
immunizations
immunize
immunized
immunizing
immuno
impact^
impacted
impacting
impacts
implement*
implementation
implemented^
implementing
implements
implicate*
implicated
implicates
implicating
implication
implications^
imply*
implied
implies^
implying
impose*
imposed^
imposes

imposing
imposition
incidence^
incident
incidentally
incidents
incorporate*
incorporated^
incorporates
incorporating
incorporation
incubate*
incubated^
incubates
incubating
incubation
incubations
incubator
incubators
preincubated
preincubation
index^
indexed
indexes
indexing
indicate*
indicated^
indicates
indicating
indication
indications
indicative
indicator
indicators
individual^
individualised
individuality
individualism
individualist
individualists
individualistic
individually
individuals
induce*
induced^
induces
inducing
induction
infect
infected
infecting
infection^
infections
infectious
infectiously
infectiousness
infective
infects
noninfectious
reinfect
reinfected
reinfecting
reinfection
reinfections
reinfects
uninfected
infer*
inference^
inferences
inferred
inferring
infers
inflame
inflamed
inflames
inflaming
inflammation
inflammations
inflammatory^
noninflammatory
infrared^*
inherent^
inherently
inhibit*
inhibited
inhibiting
inhibition^
inhibitions
inhibits
initial^
initially
initiate*
initiated
initiates
initiating
initiation^
initiations
initiative
initiatives
initiator
initiators
inject
injected
injecting
injection^
injections
injector
injectors
injects
injure*
injured
injures
injuries
injuring
injury^
uninjured
innovate*
innovation^
innovated
innovates
innovating
innovations
innovative
innovator
innovators
input^
inputs
insert*
inserted
inserting
insertion^
inserts
inset^
insets
insight*
insightful
insights^
instance^
instances
instruct*
instruction

instructed
instructing
instructions^
instructive
instructor
instructors
instructs
intact^
intactness
intake^*
intakes
integral^
integrate*
integrated^
integrates
integrating
integration
intense*
intensely
intenseness
intensification
intensified
intensifies
intensify
intensifying
intension
intensity^
intensive
intensively
interact*
interacted
interacting
interaction
interactions^
interactive
interactively
interacts
interestingly^
interface^
interfaced
interfaces
interfacing
interior^*
interiors
intermediate^*
internal^
internalise
internalised
internalises
internalising
internalize
internalized
internalizes
internalizing
internally
interpret*
interpretation^
interpretations
interpretative
interpreted
interpreting
interpretive
interprets
misinterpret
misinterpretation
misinterpretations
misinterpreted
misinterpreting
misinterprets
reinterpret
reinterpreted
reinterprets
reinterpreting
reinterpretation
reinterpretations
interval^
intervals
intervene*
intervened
intervenes
intervening
intervention^
interventions
intestine*
intestinal^
intestines
intrinsic^
intrinsically
invasive^
noninvasive
inverse^
inversely
investigate*
investigated^
investigates
investigating
investigation
investigations
investigative
investigator
investigators
involve*
involved^
involvement
involves
involving
uninvolved
ion^
ionisation
ionise
ionised
ioniser
ionisers
ionises
ionising
ionization
ionize
ionized
ionizer
ionizers
ionizes
ionizing
ions
isolate*
isolated^
isolates
isolating
isolation
isolationism
item*
itemisation
itemise
itemised
itemises
itemising
items^
kernel^
kernels
kidney^
kidneys
kinase^

kinetic^
kinetically
kinetics
kit^
kits
kitted
kitting
label*
labeled^
labeling
labelled
labelling
labels
laboratory^
lab
laboratories
labs
lactic^
laser^
lase
lased
lasers
lases
lasing
latent^
latencies
latency
lateral^
laterality
laterally
laterals
lattice^*
lattices
layer^
layered
layering
layers
linear^
linearities
linearity
linearly
nonlinear
link*
linkage
linkages
linked^
linking
links
lipid^
lipids
liver^
livers
localise
localisation
localised
localises
localising
localism
localization^
localize
localized
localizes
localizing
locate*
located
locating
location^
locations
relocate
relocated
relocates
relocating
relocation
locus^
loci
locuses
longitudinal^
longitudinally
loop^
looped
looping
loops
lysine^
magnesium^
magnesiums
magnet
magnetic^
magnetically
magnetics
magnetise
magnetised
magnetises
magnetising
magnetism
magnetization
magnetize
magnetized
magnetizes
magnetizing
magnets
nonmagnetic
magnify
magnification^
magnifications
magnified
magnifies
magnifying
magnitude^
magnitudes
maintain*
maintained^
maintaining
maintains
maintenance
major^
majorities
majority
manipulate*
manipulated
manipulates
manipulating
manipulation^
manipulations
manipulative
manual*
manually^
manuals
mapping^
mappings
margin*
marginal^
marginally
margins
marine^
mathematics^
math
mathematical
mathematically
mathematician
mathematicians
maths

matrix^*
- matrices
- matrixes

mature^
- immature
- immaturity
- maturation
- maturational
- matured
- matures
- maturing
- maturity

maximal^
- maximally

maximise*
- max
- maximised
- maximises
- maximising
- maximisation
- maximize
- maximized
- maximizes
- maximizing
- maximization
- maximum^

mechanism^*
- mechanisms

media^

median^
- medians

mediate*
- mediated^
- mediates
- mediating
- mediation

medical^*
- medically

medium^

membrane^
- membranes
- membranous

mesh^
- enmesh
- enmeshed
- enmeshes
- enmeshing
- enmeshment
- enmeshments
- meshed
- meshes
- meshing

metabolic^
- metabolically

metabolism^
- metabolisms

metabolite
- metabolites^

methanol^
- methanols

methionine^*

method^
- methodical
- methodological
- methodologies
- methodology
- methods

micro^
- micros

microbe
- microbes
- microbial^

microorganism
- microorganisms^

microscope
- microscopes
- microscopic
- microscopical
- microscopically
- microscopy^

migrate*
- migrant
- migrants
- migrated
- migrates
- migrating
- migration^
- migrations
- migratory

minimal^
- minimalisation
- minimalise
- minimalises
- minimalised
- minimalising
- minimalist
- minimalists
- minimalistic
- minimalization
- minimalize
- minimalized
- minimalizes
- minimalizing
- minimally

minimise*
- minimised
- minimises
- minimising
- minimize^
- minimized
- minimizes
- minimizing

minimum^*

minor^
- minorities
- minority
- minors

mitochondria
- mitochondrial^
- mitochondrion

mobile^*
- immobile
- immobility
- mobiles
- mobilities
- mobility

mode^
- modes

modify*
- modification
- modifications
- modified^
- modifies
- modifying
- unmodified

molar^
- molars

molecule
- molecular^
- molecules

monitor*

monitored
monitoring^
monitors
unmonitored
morphology^
morphological
morphologically
morphologies
morphologist
morphologists
mortal
immortal
immortalise
immortalised
immortalises
immortalising
immortalities
immortality
immortalize
immortalized
immortalizes
immortalizing
immortally
immortals
mortalities
mortality^
mortally
mortals
motive*
motivate
motivated^
motivates
motivating
motivation
motivations
motives
unmotivated
mount
mounted^
mounting
mountings
mounts
muscle^
muscled
muscles
muscling
mutant^
mutants
mutate
mutated
mutates
mutating
mutation
mutational
mutations^
nanoparticle
nanoparticles^
negate*
negative^
negated
negates
negating
negatively
negatives
negligible^*
negligibly
network^
networked
networking
networks
neural^*
neurally
neutral^*
neutralisation
neutralise
neutralised
neutralises
neutralising
neutrality
neutralization
neutralize
neutralized
neutralizes
neutralizing
nevertheless^
nitrogen^
nitrogenous
nitrogens
node*
nodal
nodes^
norm^*
norms
normal^*
abnormal
abnormally
normalisation
normalise
normalised
normalises
normalising
normalization
normalize
normalized
normalizes
normalizing
normality
normally
notion^
notions
novel^*
novelist
novelistic
novelists
novels
nuclear^
nucleotide^
nucleotides
nucleus^
nuclei
nucleuses
null^
nullification
nullifications
nullified
nullifies
nullify
nullifying
nullities
nullity
nutrient^
nutrients
nutrition*
nutritional^
nutritionally
nutritionist
nutritionists
nutritions
objective^
objectively
objectivity

obtain*
- obtainable
- obtained^
- obtaining
- obtains
- unobtainable

obvious^
- obviously

occupy*
- occupancy
- occupant
- occupants
- occupation
- occupational
- occupations
- occupied^
- occupier
- occupiers
- occupies
- occupying

occur^*
- occurred
- occurrence
- occurrences
- occurring
- occurs
- reoccur
- reoccurred
- reoccurring
- reoccurs

odd^
- odds

online^

onset^
- onsets

optic
- optical^
- optically
- optician
- opticians
- optics

optimal^
- optimality
- optimally

optimise
- optimisation
- optimisations
- optimised
- optimises
- optimising
- optimization^
- optimizations
- optimize
- optimized
- optimizes
- optimizing

optimum^*

option^
- optional
- options

oral^
- orally

organic^
- inorganic
- inorganically
- organically
- organics

organism
- organisms^

orient*
- orientate
- orientated
- orientates
- orientation^
- orientating
- oriented
- orienting
- orients
- reorient
- reorientation

outcome*
- outcomes^

output^
- outputs

oven^*
- ovens

overall^*

overlap^
- overlapped
- overlapping
- overlaps

overnight^

overview^
- overviews

oxidant
- antioxidant^
- antioxidants
- oxidants

oxide^
- oxides

oxidise
- oxidisation
- oxidised
- oxidises
- oxidising
- oxidization
- oxidize
- oxidized^
- oxidizes
- oxidizing

oxygen^*
- oxygens

panel^*
- panelled
- panelling
- panels

parallel^
- paralleled
- parallelled
- parallelling
- parallels
- unparalleled

parameter*
- parameters^

participate*
- participant
- participants^
- participated
- participates
- participating
- participation
- participatory

partner*
- partners^
- partnership
- partnerships

passive^
- passively
- passivity

pathogen^
- pathogens

pathogenic^
pathogenically
pathway^
pathways
patients^
peak^
peaked
peaking
peaks
penetrate
interpenetration
interpenetrations
penetrated
penetrates
penetrating
penetration^
penetrations
penetrative
peptide^
peptides
perceive*
perceived
perceives
perceiving
perception^
perceptions
percent*
percentage^
percentages
period^
periodic
periodical
periodically
periodicals
periods
periphery
peripheral^
peripherally
peripherals
peripheries
peroxide^
peroxides
persist*
persisted
persistence^
persistent
persistently
persisting
persists
perspective^
perspectives
pharmaceutical^*
pharmaceuticals
phase^*
phased
phases
phasing
phenomenon^
phenomena
phenomenal
phenotypic^
phosphate^
phosphates
phylogenetic^*
physical^
physically
physiological^
physiologically
plasma^
plasmas
plastic^
plasticity
plastics
platform^
platforms
plot^*
plots
plotted
plotter
plotters
plotting
plus^
pluses
polar^
bipolar
bipolarities
bipolarity
polarisation
polarise
polarised
polarises
polarising
polarities
polarity
polarization
polarize
polarized
polarizes
polarizing
poly^
polymer^
polymerisation
polymerise
polymerised
polymerising
polymerization
polymerize
polymerized
polymerizes
polymerizing
polymers
polymerase^
polymerases
polynomial^
polynomials
pooled^
pooling
pore^*
pores
portion^*
portions
pose^*
posed
poses
posing
positive^
positively
posterior^
potassium^
potassiums
potent^*
impotent
impotently
potently
potential^
potentially
precede*
preceded
precedence
precedent
precedes

preceding^
unprecedented
precipitate
precipitated
precipitately
precipitates
precipitating
precipitation^
precise*
imprecise
precisely
precision^
precursor^
precursors
precursory
predict*
predictability
predictable
predictably
predicted^
predicting
prediction
predictions
predicts
unpredictability
unpredictable
predominant*
predominance
predominantly^
predominate
predominated
predominates
predominating
preliminary^
preliminaries
presume*
presumably^
presumed
presumes
presuming
presumption
presumptions
prevalent*
prevalence^
prevalences
previous^*
previously
primary^*
primarily
prime^
primacy
primer*
primers^
principal^*
principally
principle^*
principled
principles
unprincipled
prior^
probe^
probed
probes
probing
proceed*
procedural
procedure^
procedures
proceeded
proceeding
proceedings
proceeds
process^
processed
processes
processing
profile^
profiled
profiles
profiling
progression^
progressions
project*
projected
projecting
projection^
projections
projects
proliferate
proliferated
proliferates
proliferating
proliferation^
proliferations
proliferative
proline^*
promote^
promoted
promoter
promoters
promotes
promoting
promotion
promotions
propagate*
propagated
propagates
propagating
propagation^
propagations
propagator
propagators
proportion^
disproportion
disproportionate
disproportionately
proportional
proportionally
proportionate
proportionately
proportions
protease^
proteases
protein^*
proteins
protocol^*
protocols
publication^
publications
publish*
published^
publisher
publishers
publishes
publishing
unpublished
pulse^
pulsed
pulses
pulsing
purchase*

- purchased^
- purchaser
- purchasers
- purchases
- purchasing

purify
- purification
- purifications
- purified^
- purifies
- purifying

purity^

pursue*
- pursued^
- pursues
- pursuing
- pursuit
- pursuits

putative^

quantify*
- quantifiable
- quantification^
- quantified
- quantifies
- quantifying
- unquantifiable

radical^*
- radically
- radicals

random^*
- randomly
- randomness

range^*
- ranged
- ranges
- ranging

ratio^*
- ratios

rational^
- irrational
- rationalisation
- rationalisations
- rationalise
- rationalised
- rationalises
- rationalising
- rationalism
- rationality
- rationalization
- rationalizations
- rationalize
- rationalized
- rationalizes
- rationally

react*
- reacted
- reacts
- reacting
- reaction^
- reactionaries
- reactionary
- reactions
- reactive
- reactivate
- reactivation
- reactor
- reactors

reagent
- reagents^

receptor^
- receptors

recover*
- recoverable
- recovered
- recovering
- recovers
- recovery^

redox^

reference^
- referenced
- references
- referencing

refine*
- refined
- refinement^
- refinements
- refines
- refining

regime^*
- regimes

region^
- regional
- regionally
- regions

regress
- regressed
- regresses
- regressing
- regression^
- regressions
- regressive

regulate*
- deregulated
- deregulates
- deregulating
- deregulation
- regulated
- regulates
- regulating
- regulation^
- regulations
- regulator
- regulators
- regulatory
- unregulated

reject*
- rejected
- rejecting
- rejection^
- rejects
- rejections

relax*
- relaxation^
- relaxed
- relaxes
- relaxing

release^*
- released
- releases
- releasing

relevant^
- irrelevance
- irrelevant
- relevance

rely*
- reliability
- reliable^
- reliably
- reliance
- reliant
- relied

relies
relying
unreliable
remove*
removable
removal
removals
removed^
removes
removing
replicate
replicated
replicates
replicating
replication^
replications
require*
required^
requirement
requirements
requires
requiring
research^
researched
researcher
researchers
researches
researching
reservoir
reservoirs^
reside*
resided
residence^
resident
residential
residents
resides
residing
residue
nonresidual
residual
residuals
residuary
residues^
resolve*
resolution^
resolved
resolves
resolving
unresolved
resonance^
resonances
resource*
resourced
resourceful
resources^
resourcing
unresourceful
respond*
responded
respondent
respondents
responding
responds
response^
responses
responsive
responsiveness
unresponsive
restore*
restoration^
restored
restores
restoring
restrict*
restricted^
restricting
restriction
restrictions
restrictive
restrictively
restricts
unrestricted
unrestrictive
resuspend
resuspended^
resuspending
resuspends
retain*
retained
retaining
retainer
retainers
retains
retention^
retentive
reveal*
revealed^
revealing
reveals
revelation
revelations
reverse^
reversal
reversed
reverses
reversible
reversing
reversals
irreversible
revise*
revised
revises
revising
revision^
revisions
robust^*
robuster
robustly
robustness
role^
roles
rotate*
rotated
rotates
rotating
rotation^
rotational
rotations
rotatory
route^
routed
routes
routing
routine^
routinely
routines
subroutine
subroutines
saline
salinities

salinity^
saturate
saturated
saturates
saturating
saturation^
saturations
unsaturated
scaling^
scaled
scan^*
scanned
scanner
scanners
scanning
scans
scenario^*
scenarios
scheme^
schematic
schematically
schemed
schemes
scheming
score^
scored
scorer
scorers
scores
scoring
seasonal^
seasonality
seasonally
unseasonal
secrete*
secreted
secretes
secreting
secretion^
secretions
secretor
secretors
secretory
section^*
sectioned
sectioning
sections
sector^*
sectors
seek^
seeking
seeks
sought
segment^
segmental
segmentary
segmentation
segmentations
segmented
segmenting
segments
select*
selected^
selecting
selection
selections
selective
selectively
selector
selectors
selects
sensing^
sensings
sensor^*
sensors
sequence^*
sequenced
sequences
sequencing
sequential
sequentially
series^
serine^
serines
serum^
antisera
antiserum
antiserums
serums
setup^*
setups
sex^*
sexes
sexism
sexual
sexuality
sexually
shift^*
shifted
shifting
shifts
significant^
insignificant
insignificantly
significance
significantly
signified
signifies
signify
signifying
silica^
silicas
silicon^*
silicone
similar^
dissimilar
similarities
similarity
similarly
simulate*
simulated
simulates
simulating
simulation^
simultaneous*
simultaneity
simultaneously^
site^
sites
skeletal^
sodium^
sodiums
software^
softwares
sole*
solely^
soluble^
insolubility
insoluble
solubility
solvent^*

solvents
somewhat^*
source^
sourced
sources
sourcing
spatial^
spatially
species^*
sp
specie
spp
subspecies
specific^
specifically
specification
specifications
specificity
specifics
specify*
specifiable
specified^
specifies
specifying
unspecified
spectra^
spectre*
specter
specters
spectral^
spectres
sphere^
spheres
spherical
spherically
spontaneous^
spontaneously
stable*
instability
stabilisation
stabilise
stabilised
stabilises
stabilising
stabilization
stabilize
stabilized
stabilizes
stabilizing
stability^
unstable
static^
statically
statistic*
statistician
statisticians
statistical^
statistically
statistics
status^
storage^*
straightforward^
strain
strained
strainer
strainers
straining
strains^
strand^*
strands
strategy^*
strategic
strategies
strategically
strategist
strategists
stress^*
stressed
stresses
stressful
stressing
unstressed
structure^*
restructure
restructured
restructures
restructuring
structural
structurally
structured
structures
structuring
unstructured
subsequent^
subsequently
subset^
subsets
substitute*
substituted
substitutes
substituting
substitution^
successor*
succession
successions
successive^
successively
successors
sufficient^*
sufficiency
insufficient
insufficiently
sufficiently
sum^*
summation
summed
summing
sums
summary^
summaries
summarise
summarised
summarises
summarising
summarisation
summarisations
summarization
summarizations
summarize
summarized
summarizes
summarizing
superior^
superiority
superiors
supernatants
supernatant^
supplement*
supplementary^
supplemented
supplementing

supplements
suppress*
suppressed
suppresses
suppressing
suppression^
suppressor
suppressors
survey^
surveyed
surveying
surveys
survive*
survival^
survived
survives
surviving
survivor
survivors
susceptible^
insusceptible
susceptibilities
susceptibility
suspend*
suspended
suspending
suspends
suspension^
sustain*
sustainable^
sustainability
sustained
sustaining
sustains
sustenance
unsustainable
switch^
switchable
switched
switches
switching
symbol*
symbolic
symbolically
symbolise
symbolises
symbolised
symbolising
symbolism
symbolize
symbolized
symbolizes
symbolizing
symbols^
symmetry^
symmetric
symmetrical
symmetrically
symmetries
unsymmetrical
symptom
symptomatic
symptomatically
symptoms^
synergistic^
synergistically
synthesis^
synth
syntheses
synthesise
synthesised
synthesiser
synthesisers
synthesises
synthesising
synthesize
synthesized
synthesizer
synthesizers
synthesizes
synthesizing
synths
synthetic^
synthetically
synthetics
tank^*
antitank
tankful
tankless
tanks
target^*
targeted
targeting
targets
task^
tasks
taxonomy*
taxonomic^
taxonomies
taxonomist
taxonomists
team^*
teamed
teaming
teams
technical^
technically
technique*
techniques^
technology^
technological
technologically
template^
templates
temporal^
temporality
temporally
terminate*
terminal^
terminals
terminated
terminates
terminating
termination
terminations
tertiary^*
text^*
texts
textual
theory^
theoretical
theoretically
theories
theorist
theorists
therapeutic^
therapeutically
therapy^*
therapies
therapist
therapists

thereby^
thermal^
    thermally
    thermals
threshold^
    thresholds
tissue^
    tissues
tolerance^*
    tolerances
topic^
    topical
    topics
toxic^
    nontoxic
    toxicities
    toxicity
toxin^*
    antitoxin
    antitoxins
    toxins
trace^
    traceable
    traced
    traces
    tracing
tract^
    tracts
tradition*
    traditional^
    traditionalist
    traditionally
    traditions
transcript*
    transcription^
    transcriptional
    transcriptions
    transcripts
transfer^
    transferable
    transference
    transferred
    transferring
    transfers
transform*
    transformation^
    transformations
    transformed
    transforming
    transforms
transient^
    intransience
    transience
    transiently
    transients
transit*
    transited
    transiting
    transition^
    transitional
    transitions
    transitory
    transits
transmit*
    transmission^
    transmissions
    transmitted
    transmitting
    transmits
transport^*
    transportation
    transported
    transporter
    transporters
    transporting
    transports
trend^
    trends
triangle^*
    triangles
    triangular
trigger^
    triggered
    triggering
    triggers
triple^
    tripled
    triples
    tripling
    triply
triplicate^
    triplicates
tumour
    tumor^
    tumors
    tumours
tyrosine^
    tyrosines
ultimate*
    ultimately^
unclear^*
undergo^
    undergoes
    undergoing
    undergone
    underwent
underlie*
    underlay
    underlies
    underlying^
uniform^*
    uniformity
    uniformly
unique^
    uniquely
    uniqueness
untreated^
uptake^
    uptakes
urea^
    ureas
utilise*
    utilisation
    utilised
    utilises
    utilising
    utiliser
    utilisers
    utility
    utilities
    utilization
    utilize
    utilized^
    utilizes
    utilizing
valid*
    invalidate
    invalidity
    validate
    validated
    validating

validation^
validity
validly
vary*
invariable
invariably
variability
variable
variables^
variably
variance
variant
variants
variation
variations
varied
varies
varying
vascular^
velocity^
velocities
verify^*
verifiable
verification
verifications
verified
verifier
verifiers
verifies
verifying
version^
versions
versus
vs^
vertical^*
vertically
verticals
via^
viable
nonviable
viabilities
viability^
video^
videoed
videoing
videos
violate*
violated
violates
violating
violation^
violations
virtual*
virtually^
virus^
viruses
vitamin^
vitamins
volatile^
nonvolatile
volatiles
volatilities
volatility
volt*
voltage^
voltages
voltmeter
voltmeters
volts
volume^
volumes
vol
wavelength^
wavelengths
weighted^*
weighting
whereas^*
widespread^
worldwide^
yeast^
yeastier
yeastiest
yeasts
yeasty
zinc^
zincs
zone^
zonal
zoned
zones
zoning

APPENDIX C
THE FREQUENCY-BASED SAWL SUBLISTS (HEADWORDS)

Sublist 1

1. analyse
2. protein
3. method
4. species
5. image
6. function
7. data
8. vary
9. process
10. acid
11. structure
12. gene
13. significant
14. obtain
15. distribute
16. indicate
17. specific
18. similar
19. sequence
20. detect
21. identify
22. factor
23. estimate
24. define
25. mathematics
26. react
27. interact
28. phase
29. region
30. approach
31. range
32. parameter
33. require
34. select
35. potential
36. extract
37. molecule
38. respond
39. strain
40. matrix
41. site
42. area
43. section
44. energy
45. previous
46. target
47. ion
48. culture
49. induce
50. stable
51. dense
52. occur
53. activate
54. regulate
55. complex
56. linear
57. correspond
58. chemical
59. consist
60. infect

Sublist 2

1. environment
2. dimension
3. volume
4. mechanism
5. investigate
6. layer
7. demonstrate
8. component
9. role
10. tissue
11. isolate
12. research
13. domain
14. assume
15. coating
16. ratio
17. evaluate
18. bacterium
19. approximate
20. equate
21. generate
22. design
23. involve
24. positive
25. individual
26. enzyme
27. compound
28. strategy
29. modify
30. normal
31. available
32. pathway
33. affect
34. cycle
35. cellular
36. construct
37. final
38. peak
39. maximise
40. assay
41. technique
42. statistic
43. proceed
44. carbon
45. column
46. predict
47. compute
48. inhibit
49. correlate
50. source
51. composition
52. synthesis
53. negate
54. initial
55. lipid
56. theory
57. major
58. remove
59. transfer
60. fluorescent

Sublist 3

1. period
2. locate
3. expose
4. error
5. random
6. dynamic
7. via
8. residue
9. contribute
10. transform
11. achieve
12. assess
13. fungus
14. medium
15. element
16. release
17. amino
18. label
19. promote
20. cancer
21. constant
22. genetic
23. genome
24. network
25. muscle
26. contact
27. exhibit
28. plasma
29. pathogen
30. conclude
31. reveal
32. spectra
33. electron
34. imaged
35. transit
36. spatial
37. incubate
38. formula
39. stress
40. impact
41. whereas
42. simulate
43. diverse
44. finite
45. evolve
46. membrane
47. contrast
48. intense
49. fraction
50. magnet
51. organic
52. primary
53. feature
54. peptide
55. focus
56. differentiate
57. coefficient
58. mode
59. supplement
60. resolve

Sublist 4

1. maintain
2. receptor
3. consume
4. confirm
5. buffer
6. laboratory
7. nanoparticle
8. abundant
9. transcript
10. establish
11. reference
12. evident
13. virus
14. diffuse
15. microscope
16. alternative
17. link
18. optic
19. subsequent
20. absorb
21. mutate
22. morphology
23. hence
24. transport
25. drug
26. scheme
27. glucose
28. degrade
29. capacity
30. plot
31. accurate
32. core
33. metabolite
34. denote
35. metabolic
36. survive
37. symmetry
38. furthermore
39. illustrate
40. percent
41. profile
42. community
43. replicate
44. proportion
45. display
46. filter
47. concept
48. interval
49. ferment
50. hypothesis
51. mutant
52. recover
53. diet
54. utilise
55. sufficient
56. consequent
57. integrate
58. fibre
59. create
60. version

Sublist 5

1. adapt
2. versus
3. microbe
4. biomass
5. conduct
6. media
7. device
8. yeast
9. optimise
10. quantify
11. couple
12. node
13. probe
14. generation
15. shift
16. optimal
17. overall
18. mediate
19. technology
20. oxide
21. distinct
22. oxygen
23. alter
24. thermal
25. nutrient
26. summary
27. challenge
28. enable
29. kinase
30. tumour
31. abstract
32. suppress
33. storage
34. crystal
35. valid
36. convert
37. hybrid
38. emerge
39. inflame
40. atom
41. scan
42. clinic
43. author
44. attribute
45. triangle
46. series
47. sum
48. prior
49. enrich
50. criteria
51. unique
52. electrode
53. solvent
54. dilute
55. benefit
56. project
57. monitor
58. sensor
59. nitrogen
60. retain

Sublist 6

1. regress
2. globe
3. patients
4. relevant
5. inject
6. reverse
7. dual
8. vascular
9. bond
10. physical
11. laser
12. dominate
13. convene
14. droplet
15. publish
16. imply
17. transmit
18. spectre
19. immune
20. terminate
21. restrict
22. gel
23. implement
24. metabolism
25. expand
26. geometry
27. acquire
28. novel
29. primer
30. soluble
31. susceptible
32. fragment
33. disperse
34. serum
35. phylogenetic
36. score
37. bias
38. purify
39. phosphate
40. appropriate
41. deposition
42. facilitate
43. rely
44. interface
45. index
46. polymer
47. survey
48. ethanol
49. context
50. equivalent
51. array
52. catalyse
53. organism
54. amplify
55. zone
56. software
57. oxidant
58. highlight
59. lattice
60. aspect

Sublist 7

1. adult
2. graph
3. incorporate
4. sodium
5. defect
6. precise
7. kinetic
8. mature
9. feeding
10. suspend
11. protocol
12. encode
13. segment
14. uniform
15. polar
16. instance
17. simultaneous
18. panel
19. genotype
20. assemble
21. internal
22. deposit
23. discrete
24. chromatography
25. interpret
26. anneal
27. fuse
28. exclude
29. centrifuge
30. hydrogen
31. loop
32. minimum
33. emit
34. genus
35. mobile
36. mortal
37. threshold
38. capable
39. initiate
40. ensure
41. clone
42. mitochondria
43. robust
44. coordinate
45. adhesion
46. adjust
47. capture
48. affinity
49. viable
50. category
51. digest
52. rotate
53. conserve
54. trend
55. economy
56. access
57. channel
58. aggregate
59. flux
60. align

Sublist 8

1. biosynthetic
2. classic
3. constrain
4. localise
5. parallel
6. sex
7. despite
8. intermediate
9. wavelength
10. tradition
11. elevate
12. fabricate
13. phenomenon
14. scaling
15. minimise
16. differential
17. dataset
18. input
19. liver
20. grid
21. physiological
22. dependence
23. sphere
24. diagram
25. posterior
26. secrete
27. aqueous
28. microorganism
29. obvious
30. proliferate
31. assign
32. epithelial
33. contaminate
34. pulse
35. code
36. locus
37. nuclear
38. ammonia
39. framework
40. synthetic
41. apparent
42. intestine
43. infer
44. margin
45. mesh
46. eliminate
47. considerable
48. grade
49. principle
50. insert
51. scenario
52. hydrolysis
53. empirical
54. reagent
55. constitute
56. fluid
57. toxic
58. brief
59. comprise
60. edit

Sublist 9

1. null
2. pore
3. switch
4. height
5. attach
6. tolerance
7. propagate
8. online
9. saturate
10. candidate
11. resource
12. bulk
13. configure
14. identical
15. marine
16. calibrate
17. endogenous
18. polynomial
19. volt
20. underlie
21. volatile
22. instruct
23. verify
24. dye
25. thereby
26. uptake
27. cellulose
28. migrate
29. zinc
30. orient
31. participate
32. temporal
33. therapeutic
34. external
35. velocity
36. dissolve
37. chloride
38. interestingly
39. trigger
40. exponential
41. platform
42. insight
43. precipitate
44. integral
45. biology
46. chamber
47. precursor
48. alcohol
49. symptom
50. implicate
51. option
52. task
53. barrier
54. decompose
55. fundamental
56. nevertheless
57. saline
58. bead
59. donor
60. magnitude

Sublist 10

1. substitute
2. tank
3. disrupt
4. acetate
5. transient
6. sensing
7. cerevisiae
8. incidence
9. deficiency
10. outcome
11. reservoir
12. batch
13. output
14. assist
15. enzymatic
16. ecosystem
17. gut
18. biochemical
19. kit
20. allele
21. decade
22. taxonomy
23. glycerol
24. neutral
25. minimal
26. decline
27. goal
28. prevalent
29. excitation
30. supernatants
31. chip
32. equip
33. atmosphere
34. vertical
35. subset
36. nucleus
37. plastic
38. chromosome
39. file
40. capillary
41. silica
42. database
43. fluctuate
44. overlap
45. regime
46. trace
47. undergo
48. automate
49. crucial
50. complement
51. climate
52. immobilize
53. resonance
54. diverge
55. specify
56. persist
57. latent
58. perspective
59. route
60. methanol

Sublist 11

1. agar
2. therapy
3. apoptosis
4. flexible
5. predominant
6. radical
7. depict
8. kernel
9. status
10. explicit
11. purchase
12. nucleotide
13. objective
14. micro
15. nutrition
16. symbol
17. dash
18. sustain
19. relax
20. topic
21. triple
22. manipulate
23. ecological
24. geography
25. diffract
26. document
27. poly
28. toxin
29. mapping
30. silicon
31. weighted
32. duration
33. emphasis
34. ambience
35. chemistry
36. oxidise
37. cardiac
38. infrared
39. ester
40. vitamin
41. video
42. hydrophobic
43. putative
44. compatible
45. carbohydrate
46. lateral
47. annual
48. penetrate
49. plus
50. restore
51. setup
52. cholesterol
53. inverse
54. evaporate
55. exceed
56. overnight
57. seasonal
58. intrinsic
59. adjacent
60. bacillus

Sublist 12

1. adequate
2. potassium
3. overview
4. rational
5. norm
6. engineered
7. portion
8. embed
9. feasible
10. append
11. occupy
12. superior
13. redox
14. alkaline
15. downstream
16. reside
17. arbitrary
18. perceive
19. exploit
20. drama
21. deplete
22. minor
23. bovine
24. communicate
25. acetic
26. refine
27. technical
28. anaerobic
29. ultimate
30. amplitude
31. frequencies
32. spontaneous
33. dehydrogenased
34. elongate
35. maximal
36. principal
37. template
38. tyrosine
39. neural
40. injure
41. mount
42. accompany
43. distil
44. acute
45. administrate
46. invasive
47. lysine
48. eventual
49. item
50. optimum
51. strand
52. cascade
53. displace
54. attain
55. comprehensive
56. periphery
57. cysteine
58. cavity
59. publication
60. hierarchy

Sublist 13

1. progression
2. resuspend
3. magnify
4. median
5. architecture
6. basal
7. discrepancy
8. pathogenic
9. prime
10. baseline
11. discriminate
12. odd
13. untreated
14. aromatic
15. routine
16. breast
17. conform
18. motive
19. intervene
20. agarose
21. calcium
22. developmental
23. protease
24. analogy
25. inset
26. skeletal
27. lactic
28. comment
29. functionalize
30. reject
31. efficiently
32. confocal
33. worldwide
34. confer
35. oral
36. innovate
37. static
38. clarify
39. negligible
40. tract
41. cardiovascular
42. notion
43. cooperate
44. biotechnology
45. encounter
46. intact
47. aid
48. crude
49. converse
50. exogenous
51. attitude
52. manual
53. triplicate
54. unclear
55. passive
56. phenotypic
57. text
58. chronic
59. definite
60. favorable

Sublist 14

1. bioactive
2. dissolution
3. expert
4. kidney
5. pharmaceutical
6. pooled
7. urea
8. elemental
9. compensate
10. intake
11. adverse
12. academy
13. methionine
14. preliminary
15. coincide
16. polymerase
17. digital
18. hydroxyl
19. impose
20. colon
21. cumulative
22. medical
23. purity
24. potent
25. germ
26. acetone
27. aware
28. biomedical
29. partner
30. widespread
31. additives
32. straightforward
33. interior
34. proline
35. revise
36. sector
37. synergistic
38. antimicrobial
39. guarantee
40. seek
41. electrostatic
42. tertiary
43. precede
44. sole
45. electrophoresis
46. successor
47. inherent
48. somewhat
49. feedback
50. allocate
51. breakdown
52. covalent
53. longitudinal
54. molar
55. onset
56. peroxide
57. presume
58. albumin
59. pose
60. serine

Sublist 15

1. magnesium
2. team
3. oven
4. hydrophilic
5. depress
6. accommodate
7. virtual
8. pursue
9. concurrent
10. devote
11. entity
12. violate
13. favorably

APPENDIX D
THE CEFR MAPPING RESULTS

No.	Word	EVP	GSE	Result	Parts of Speech
001	absorption	UL	C1	C1	[noun, singular or mass]
002	abstract	B2	B2	B2	[noun, singular or mass]
003	abundance	UL	C1	C1	[noun, singular or mass]
004	academic	B2	B1	B1	[adjective]
005	access	B1	B1	B1	[noun, singular or mass]; [verb, base form]
006	accommodate	UL	B2	B2	[verb, base form]
007	accompanied	B1	B2	B1	[verb, past participle]
008	accuracy	B2	B2	B2	[noun, singular or mass]
009	acetate	UL	UL	UL	[noun, singular or mass]
010	acetic	UL	UL	UL	[adjective]
011	acetone	UL	UL	UL	[noun, singular or mass]
012	achieved	B1	B1	B1	[verb, past participle]
013	acid	B2	A1	A1	[noun, singular or mass]
014	acquisition	UL	C1	C1	[noun, singular or mass]
015	activation	UL	UL	UL	[noun, singular or mass]
016	acute	UL	B2	B2	[adjective]
017	adapted	B2	B2	B2	[verb, past participle]
018	additive	UL	UL	UL	[adjective]
019	adequate	B2	B2	B2	[adjective]
020	adhesion	UL	C1	C1	[noun, singular or mass]
021	adjacent	C2	B2	B2	[adjective]
022	adjusted	B2	B2	B2	[verb, past participle]
023	administration	C1	B2	B2	[noun, singular or mass]
024	adult	A1	A2	A1	[adjective]; [noun, singular or mass]
025	adverse	C2	B2	B2	[adjective]
026	affect	B2	B1	B1	[verb, base form]
027	affinity	UL	C1	C1	[noun, singular or mass]
028	agar	UL	UL	UL	[noun, singular or mass]
029	agarose	UL	UL	UL	[noun, singular or mass]
030	aggregation	UL	UL	UL	[noun, singular or mass]
031	aid	C1	B2	B2	[noun, singular or mass]; [verb, base form]
032	albumin	UL	UL	UL	[noun, singular or mass]
033	alcohol	A2	B1	A2	[noun, singular or mass]
034	alignment	UL	C1	C1	[noun, singular or mass]
035	alkaline	UL	C1	C1	[adjective]
036	allele	UL	UL	UL	[noun, singular or mass]
037	allocation	C2	C1	C1	[noun, singular or mass]
038	altered	B2	B2	B2	[adjective]
039	alternative	B2	B1	B1	[adjective]; [noun, singular or mass]

No.	Word	EVP	GSE	Result	Parts of Speech
040	ambient	UL	UL	UL	[adjective]
041	amino	UL	UL	UL	[noun, singular or mass]
042	ammonia	UL	C1	C1	[noun, singular or mass]
043	amplification	UL	C1	C1	[noun, singular or mass]
044	amplitude	UL	C1	C1	[noun, singular or mass]
045	anaerobic	UL	UL	UL	[adjective]
046	analogous	C2	C1	C1	[adjective]
047	analysis	B2	B1	B1	[noun, singular or mass]
048	annealing	UL	UL	UL	[verb, gerund/participle]
049	annual	B1	B1	B1	[adjective]
050	antimicrobial	UL	UL	UL	[adjective]
051	antioxidant	UL	C1	C1	[noun, singular or mass]
052	apoptosis	UL	UL	UL	[noun, singular or mass]
053	apparent	B2	B2	B2	[adjective]
054	appendix	UL	B2	B2	[proper noun, singular]
055	approach	B1	B1	B1	[noun, singular or mass]; [verb, base form]
056	appropriate	B2	B1	B1	[adjective]
057	approximately	B1	B1	B1	[adverb]
058	aqueous	UL	C1	C1	[adjective]
059	arbitrary	C2	UL	C2	[adjective]
060	architecture	B1	B1	B1	[noun, singular or mass]
061	area	A2	A2	A2	[noun, singular or mass]
062	aromatic	UL	C1	C1	[adjective]
063	array	UL	C1	C1	[noun, singular or mass]
064	aspects	B2	B2	B2	[noun plural]
065	assay	UL	C1	C1	[noun, singular or mass]
066	assembly	C2	B2	B2	[noun, singular or mass]
067	assessed	B2	B2	B2	[verb, past participle]
068	assigned	C1	B2	B2	[verb, past participle]
069	assisted	B2	B2	B2	[adjective]; [verb, past participle]
070	assume	B2	B2	B2	[verb, base form]; [verb, present, non-3rd p.]
071	atmosphere	B1	B1	B1	[noun, singular or mass]
072	atoms	B2	B2	B2	[noun plural]
073	attached	B1	B2	B1	[verb, past participle]
074	attained	C1	B2	B2	[verb, past participle]
075	attitude	B1	B1	B1	[noun, singular or mass]
076	attributed	C2	B2	B2	[verb, past participle]; [verb, past tense]
077	authors	B1	A2	A2	[noun plural]
078	automated	UL	B2	B2	[adjective]; [verb, past participle]
079	available	A2	A2	A2	[adjective]
080	aware	B2	B2	B2	[adjective]
081	bacillus	UL	UL	UL	[noun, singular or mass]
082	bacteria	C1	B1	B1	[noun plural]

No.	Word	EVP	GSE	Result	Parts of Speech
083	barrier	B2	B2	B2	[noun, singular or mass]
084	basal	UL	UL	UL	[adjective]
085	baseline	UL	C1	C1	[noun, singular or mass]; [adjective]
086	batch	UL	C1	C1	[noun, singular or mass]
087	beads	UL	C1	C1	[noun plural]
088	beneficial	B2	B2	B2	[adjective]
089	bias	C2	B2	B2	[noun, singular or mass]; [verb, base form]
090	bioactive	UL	UL	UL	[adjective]
091	biochemical	UL	C1	C1	[adjective]
092	biology	A2	B2	A2	[noun, singular or mass]
093	biomass	UL	C1	C1	[noun, singular or mass]
094	biomedical	UL	C1	C1	[adjective]
095	biosynthesis	UL	UL	UL	[noun, singular or mass]
096	biotechnology	UL	C1	C1	[noun, singular or mass]
097	bonds	B2	B2	B2	[noun plural]
098	bovine	UL	C1	C1	[adjective]
099	breakdown	B2	B2	B2	[noun, singular or mass]
100	breast	B1	B2	B1	[noun, singular or mass]
101	briefly	B1	B1	B1	[adverb]
102	buffer	UL	C1	C1	[noun, singular or mass]; [verb, base form]
103	bulk	C1	B2	B2	[adjective]; [noun, singular or mass]
104	calcium	UL	B2	B2	[noun, singular or mass]
105	calibrate	UL	C1	C1	[verb, base form]
106	cancer	B1	B1	B1	[noun, singular or mass]
107	candidate	B2	B1	B1	[noun, singular or mass]
108	capable	B2	B2	B2	[adjective]
109	capacity	B2	B2	B2	[noun, singular or mass]
110	capillary	UL	C1	C1	[adjective]; [noun, singular or mass]
111	capture	B2	B2	B2	[noun, singular or mass]; [verb, base form]
112	carbohydrate	UL	C1	C1	[noun, singular or mass]
113	carbon	B2	B2	B2	[noun, singular or mass]
114	cardiac	UL	B2	B2	[adjective]
115	cardiovascular	UL	C1	C1	[adjective]
116	cascade	UL	UL	UL	[noun, singular or mass]
117	catalytic	UL	C1	C1	[adjective]
118	categories	B2	B1	B1	[noun plural]
119	cavity	UL	C1	C1	[noun, singular or mass]
120	cellular	UL	C1	C1	[adjective]
121	cellulose	UL	C1	C1	[noun, singular or mass]
122	centrifugation	UL	UL	UL	[noun, singular or mass]
123	cerevisiae	UL	UL	UL	[noun, singular or mass]; [noun plural]
124	challenge	B1	B1	B1	[noun, singular or mass]; [verb, base form]
125	chamber	UL	B2	B2	[noun, singular or mass]

No.	Word	EVP	GSE	Result	Parts of Speech
126	channel	A2	A2	A2	[noun, singular or mass]; [verb, base form]
127	chemical	B2	B1	B1	[adjective]; [noun, singular or mass]
128	chemistry	A2	B2	A2	[noun, singular or mass]
129	chip	B2	B2	B2	[noun, singular or mass]
130	chloride	UL	C1	C1	[noun, singular or mass]
131	cholesterol	C1	B2	B2	[noun, singular or mass]
132	chromatography	UL	UL	UL	[noun, singular or mass]
133	chromosome	UL	UL	UL	[noun, singular or mass]
134	chronic	C2	B2	B2	[adjective]
135	clarify	C1	B2	B2	[verb, base form]
136	classical	A2	B2	A2	[adjective]
137	climate	B1	B1	B1	[noun, singular or mass]
138	clinical	C1	B2	B2	[adjective]
139	clones	UL	B2	B2	[noun plural]
140	coating	UL	C1	C1	[noun, singular or mass]; [verb, gerund/ participle]
141	code	B2	B1	B1	[noun, singular or mass]
142	coefficient	UL	C1	C1	[noun, singular or mass]
143	coincide	C2	B2	B2	[verb, base form]; [verb, present, non-3rd p.]
144	colon	B2	B2	B2	[noun, singular or mass]
145	column	B2	B1	B1	[adjective]; [noun, singular or mass]
146	comments	B1	B1	B1	[noun plural]
147	communication	B1	B1	B1	[noun, singular or mass]
148	community	B2	A2	A2	[noun, singular or mass]
149	compatible	C1	B2	B2	[adjective]
150	compensate	C1	B2	B2	[verb, base form]
151	complementary	UL	C1	C1	[adjective]
152	complex	B2	B1	B1	[adjective]; [noun, singular or mass]
153	components	C1	B2	B2	[noun plural]
154	composition	C2	B2	B2	[noun, singular or mass]
155	compounds	UL	B2	B2	[noun, singular or mass]; [verb, present 3d p. sing.]
156	comprehensive	C1	B2	B2	[adjective]
157	comprised	C1	B2	B2	[verb, past participle]; [verb, past tense]
158	computational	UL	C1	C1	[adjective]
159	concept	B2	B2	B2	[noun, singular or mass]
160	conclusion	B1	B1	B1	[noun, singular or mass]
161	concurrent	UL	UL	UL	[adjective]
162	conducted	B2	B2	B2	[verb, past participle]; [verb, past tense]
163	conference	B1	B1	B1	[noun, singular or mass]
164	configuration	UL	C1	C1	[noun, singular or mass]
165	confirmed	B1	B1	B1	[verb, past participle]; [verb, past tense]
166	confocal	UL	UL	UL	[adjective]
167	conformation	UL	UL	UL	[noun, singular or mass]

No.	Word	EVP	GSE	Result	Parts of Speech
168	consequently	B2	B2	B2	[adverb]
169	conserved	UL	UL	UL	[verb, past participle]; [verb, past tense]
170	considerable	B2	B2	B2	[adjective]
171	consistent	C2	B2	B2	[adjective]
172	constant	B2	B2	B2	[adjective]
173	constituents	UL	B2	B2	[noun plural]
174	constraints	C2	B2	B2	[noun plural]
175	construct	B2	B2	B2	[noun, singular or mass]; [verb, base form]
176	consumption	C1	B2	B2	[noun, singular or mass]
177	contact	A2	A2	A2	[noun, singular or mass]; [verb, base form]
178	contamination	UL	B2	B2	[noun, singular or mass]
179	context	B2	B1	B1	[noun, singular or mass]
180	contrast	B2	B2	B2	[noun, singular or mass]; [verb, base form]
181	contribute	B2	B2	B2	[verb, base form]; [verb, present, non-3rd p.]
182	conventional	B2	B2	B2	[adjective]
183	conversely	UL	C1	C1	[adverb]
184	conversion	C2	B2	B2	[noun, singular or mass]
185	cooperation	B2	B2	B2	[noun, singular or mass]
186	coordinates	UL	B2	B2	[noun plural]; [verb, present 3d p. sing.]
187	core	C2	B2	B2	[noun, singular or mass]
188	correlation	UL	C1	C1	[noun, singular or mass]
189	corresponding	B2	B2	B2	[adjective]; [verb, gerund/participle]
190	coupled	UL	C1	C1	[verb, past participle]; [adjective]
191	covalent	UL	UL	UL	[adjective]
192	created	B1	B1	B1	[verb, past tense]; [verb, past participle]
193	criteria	C1	UL	C1	[noun plural]
194	crucial	B2	B2	B2	[adjective]
195	crude	C2	B2	B2	[adjective]; [noun, singular or mass]
196	crystal	C1	B2	B2	[adjective]; [noun, singular or mass]
197	culture	B1	A2	A2	[noun, singular or mass]; [verb, base form]
198	cumulative	UL	UL	UL	[adjective]
199	cycle	B1	B2	B1	[noun, singular or mass]
200	cysteine	UL	UL	UL	[noun, singular or mass]
201	dashed	B2	UL	B2	[verb, past participle]
202	data	B2	B1	B1	[noun plural]
203	database	B2	B1	B1	[noun, singular or mass]
204	dataset	UL	C1	C1	[noun, singular or mass]
205	decades	B2	B1	B1	[noun plural]
206	decline	B2	B2	B2	[noun, singular or mass]; [verb, base form]
207	decomposition	UL	UL	UL	[noun, singular or mass]
208	defects	C1	B2	B2	[noun plural]
209	deficient	UL	C1	C1	[adjective]; [proper noun, singular]
210	defined	B2	B1	B1	[adjective]; [verb, past participle]

No.	Word	EVP	GSE	Result	Parts of Speech
211	definite	B2	B1	B1	[adjective]
212	degradation	UL	C1	C1	[noun, singular or mass]
213	dehydrogenase	UL	UL	UL	[noun, singular or mass]
214	demonstrated	B2	B1	B1	[verb, past participle]; [verb, past tense]
215	denote	UL	C1	C1	[verb, base form]; [verb, present, non-3rd p.]
216	density	C1	B2	B2	[noun, singular or mass]
217	dependence	C2	C1	C1	[noun, singular or mass]
218	depicted	C2	B2	B2	[verb, past participle]
219	depletion	UL	C1	C1	[noun, singular or mass]
220	deposited	C1	B2	B2	[verb, past participle]; [adjective]
221	deposition	UL	C1	C1	[noun, singular or mass]
222	depression	B2	B1	B1	[noun, singular or mass]
223	design	B1	B1	B1	[noun, singular or mass]; [verb, base form]
224	despite	B1	B1	B1	[preposition/subord. conj.]
225	detection	UL	C1	C1	[noun, singular or mass]
226	developmental	UL	B2	B2	[adjective]
227	device	B2	B2	B2	[noun, singular or mass]
228	devoted	B2	B2	B2	[verb, past participle]
229	diagram	B1	B2	B1	[noun, singular or mass]
230	diet	B1	B2	B1	[noun, singular or mass]
231	differential	UL	B1	B1	[adjective]
232	differentiation	UL	C1	C1	[noun, singular or mass]
233	diffraction	UL	C1	C1	[noun, singular or mass]
234	diffusion	UL	C1	C1	[noun, singular or mass]
235	digestion	C1	B2	B2	[noun, singular or mass]
236	digital	A2	B1	A2	[adjective]
237	diluted	UL	UL	UL	[verb, past participle]; [adjective]
238	dimensional	UL	B2	B2	[adjective]
239	discrepancy	UL	B2	B2	[noun, singular or mass]
240	discrete	UL	C1	C1	[adjective]
241	discrimination	C1	B2	B2	[adjective]
242	dispersion	UL	UL	UL	[noun, singular or mass]
243	displacement	C1	C1	C1	[noun, singular or mass]
244	display	B1	B1	B1	[noun, singular or mass]; [verb, base form]; [verb, present, non-3rd p.]
245	disruption	C1	B2	B2	[noun, singular or mass]
246	dissolution	UL	C1	C1	[noun, singular or mass]
247	dissolved	C1	B2	B2	[verb, past participle]
248	distilled	UL	C1	C1	[adjective]
249	distinct	C1	B2	B2	[adjective]
250	distribution	C1	B2	B2	[noun, singular or mass]
251	divergence	UL	C1	C1	[noun, singular or mass]
252	diversity	C1	B2	B2	[noun, singular or mass]

No.	Word	EVP	GSE	Result	Parts of Speech
253	documented	UL	B2	B2	[verb, past participle]
254	domain	C1	UL	C1	[noun, singular or mass]
255	dominant	C1	B2	B2	[adjective]
256	donor	C2	B2	B2	[noun, singular or mass]
257	downstream	UL	C1	C1	[adjective]; [adverb]
258	dramatically	B2	B2	B2	[adverb]
259	droplet	UL	C1	C1	[noun, singular or mass]
260	drug	B2	B1	B1	[noun, singular or mass]
261	dual	C1	B2	B2	[adjective]
262	duration	C1	B2	B2	[noun, singular or mass]
263	dye	UL	C1	C1	[noun, singular or mass]
264	dynamics	UL	C1	C1	[noun plural]
265	ecological	B2	B2	B2	[adjective]
266	economic	B2	B1	B1	[adjective]
267	ecosystems	UL	UL	UL	[noun plural]
268	edited	B2	B2	B2	[verb, past participle]; [adjective]; [noun, singular or mass]
269	efficiently	B2	B2	B2	[adverb]
270	electrode	UL	C1	C1	[noun, singular or mass]
271	electron	UL	C1	C1	[noun, singular or mass]
272	electrophoresis	UL	UL	UL	[noun, singular or mass]
273	electrostatic	UL	UL	UL	[adjective]
274	elemental	UL	UL	UL	[adjective]
275	elements	B2	B2	B2	[noun plural]
276	elevated	UL	C1	C1	[adjective]; [verb, past participle]
277	elimination	C2	B2	B2	[noun, singular or mass]
278	elongation	UL	UL	UL	[noun, singular or mass]
279	embedded	UL	C1	C1	[verb, past participle]
280	emergence	UL	C1	C1	[noun, singular or mass]
281	emission	C1	UL	C1	[noun, singular or mass]
282	emphasis	B2	B1	B1	[noun, singular or mass]
283	empirical	C2	C1	C1	[adjective]
284	enable	B2	B2	B2	[verb, base form]
285	encoding	UL	UL	UL	[verb, gerund/participle]
286	encountered	B2	B2	B2	[verb, past tense]
287	endogenous	UL	UL	UL	[adjective]
288	energy	B1	A2	A2	[noun, singular or mass]
289	engineered	UL	B2	B2	[adjective]; [verb, past participle]
290	enrichment	UL	C1	C1	[noun, singular or mass]
291	ensure	B2	B2	B2	[verb, base form]
292	entities	C2	UL	C2	[noun plural]
293	environmental	B1	B1	B1	[adjective]
294	enzymatic	UL	UL	UL	[adjective]

No.	Word	EVP	GSE	Result	Parts of Speech
295	enzyme	UL	B2	B2	[noun, singular or mass]
296	epithelial	UL	UL	UL	[adjective]
297	equation	C1	B2	B2	[noun, singular or mass]
298	equipped	B2	B2	B2	[verb, past tense]
299	equivalent	C1	B2	B2	[adjective]
300	error	B2	B1	B1	[noun, singular or mass]
301	established	B2	B1	B1	[verb, past participle]; [adjective]
302	esters	UL	UL	UL	[noun plural]
303	estimated	B2	B2	B2	[verb, past participle]; [adjective]
304	ethanol	UL	C1	C1	[noun, singular or mass]
305	evaluated	C1	B2	B2	[verb, past participle]
306	evaporation	UL	C1	C1	[noun, singular or mass]
307	eventually	B2	B1	B1	[adverb]
308	evidence	B2	B1	B1	[noun, singular or mass]
309	evolution	B2	B2	B2	[noun, singular or mass]
310	exceed	C1	B2	B2	[verb, base form]
311	excitation	UL	UL	UL	[noun, singular or mass]
312	excluded	C1	B2	B2	[verb, past participle]
313	exhibited	C1	B2	B2	[verb, past tense]; [verb, past participle]
314	exogenous	UL	UL	UL	[adjective]
315	expanded	B2	B2	B2	[verb, past participle]; [adjective]
316	experts	B1	B1	B1	[noun plural]
317	explicit	C2	B2	B2	[adjective]
318	exploited	B2	B2	B2	[verb, past participle]
319	exponential	UL	UL	UL	[adjective]; [noun, singular or mass]
320	exposure	C1	B2	B2	[noun, singular or mass]
321	external	B2	B2	B2	[adjective]
322	extraction	UL	C1	C1	[noun, singular or mass]
323	fabricated	UL	UL	UL	[verb, past participle]; [adjective]
324	facilitate	C1	B2	B2	[verb, base form]; [verb, present, non-3rd p.]
325	factors	B2	B2	B2	[noun plural]
326	favorable	B2	B2	B2	[adjective]
327	favorably	UL	UL	UL	[adverb]
328	feasible	C1	C1	C1	[adjective]
329	features	B2	B2	B2	[noun plural]; [verb, present 3d p. sing.]
330	feedback	B2	B1	B1	[noun, singular or mass]
331	feeding	B1	B2	B1	[noun, singular or mass]; [verb, gerund/participle]; [adjective]
332	fermentation	UL	C1	C1	[noun, singular or mass]
333	fiber	UL	B2	B2	[noun, singular or mass]
334	file	A2	B1	A2	[noun, singular or mass]
335	filter	C2	B2	B2	[noun, singular or mass]; [verb, base form]
336	finally	A2	B1	A2	[adverb]

No.	Word	EVP	GSE	Result	Parts of Speech
337	finite	C2	C1	C1	[adjective]
338	flexible	B2	B2	B2	[adjective]
339	fluctuations	UL	C1	C1	[noun plural]
340	fluid	C2	B2	B2	[adjective]; [noun, singular or mass]
341	fluorescence	UL	UL	UL	[noun, singular or mass]
342	flux	UL	UL	UL	[noun, singular or mass]
343	focus	B2	B1	B1	[noun, singular or mass]; [verb, present, non-3rd p.]
344	formula	C1	B2	B2	[noun, singular or mass]
345	fraction	C2	B2	B2	[noun, singular or mass]
346	fragments	UL	B2	B2	[noun plural]; [verb, present 3d p. sing.]
347	framework	C2	B2	B2	[noun, singular or mass]
348	frequencies	UL	B2	B2	[noun plural]
349	function	B2	B1	B1	[noun, singular or mass]; [verb, present, non-3rd p.]
350	functionalized	UL	UL	UL	[verb, past participle]; [adjective]
351	fundamental	C2	B2	B2	[adjective]
352	fungal	UL	C1	C1	[adjective]
353	furthermore	B2	B2	B2	[adverb]
354	fusion	UL	B2	B2	[noun, singular or mass]
355	gel	UL	B2	B2	[noun, singular or mass]
356	gene	C1	B2	B2	[noun, singular or mass]
357	generated	B2	B2	B2	[verb, past participle]; [adjective]
358	generation	B1	B2	B1	[noun, singular or mass]
359	genetic	C1	B2	B2	[adjective]
360	genome	UL	C1	C1	[noun, singular or mass]
361	genotypes	UL	UL	UL	[noun plural]
362	genus	UL	C1	C1	[noun, singular or mass]
363	geographic	UL	B2	B2	[adjective]
364	geometry	UL	B2	B2	[noun, singular or mass]
365	germ	C1	B2	B2	[noun, singular or mass]
366	global	B2	A2	A2	[adjective]
367	glucose	UL	B2	B2	[noun, singular or mass]
368	glycerol	UL	UL	UL	[noun, singular or mass]
369	goal	A2	A2	A2	[noun, singular or mass]
370	grade	B1	A2	A2	[noun, singular or mass]; [adjective]
371	graph	B2	B1	B1	[noun, singular or mass]
372	grid	UL	B2	B2	[noun, singular or mass]
373	guarantee	B2	B2	B2	[verb, base form]; [noun, singular or mass]
374	gut	UL	B2	B2	[noun, singular or mass]; [adjective]
375	height	B1	B1	B1	[noun, singular or mass]
376	hence	C1	B2	B2	[adverb]
377	hierarchical	C2	C1	C1	[adjective]
378	highlighted	B2	B1	B1	[verb, past participle]; [verb, past tense]

No.	Word	EVP	GSE	Result	Parts of Speech
379	hybrid	UL	C1	C1	[adjective]; [noun, singular or mass]
380	hydrogen	UL	B2	B2	[noun, singular or mass]
381	hydrolysis	UL	UL	UL	[noun, singular or mass]
382	hydrophilic	UL	UL	UL	[adjective]
383	hydrophobic	UL	UL	UL	[adjective]
384	hydroxyl	UL	UL	UL	[noun, singular or mass]
385	hypothesis	C2	B2	B2	[noun, singular or mass]
386	identical	B2	B1	B1	[adjective]
387	identified	B2	B1	B1	[verb, past participle]; [adjective]
388	illustrated	B2	B2	B2	[verb, past participle]
389	image	B2	B1	B1	[noun, singular or mass]
390	imaging	UL	B2	B2	[noun, singular or mass]; [verb, gerund/participle]
391	immobilized	UL	UL	UL	[verb, past participle]
392	immune	C2	B2	B2	[adjective]
393	impact	B2	B2	B2	[noun, singular or mass]; [verb, base form]
394	implemented	B2	B2	B2	[verb, past participle]
395	implications	C1	B2	B2	[noun plural]
396	implies	C2	B2	B2	[verb, present 3d p. sing.]
397	imposed	C1	B2	B2	[verb, past participle]
398	incidence	UL	C1	C1	[noun, singular or mass]
399	incorporated	C2	B2	B2	[verb, past participle]
400	incubated	UL	C1	C1	[verb, past participle]
401	index	C1	B1	B1	[noun, singular or mass]
402	indicated	B2	B1	B1	[verb, past participle]; [adjective]
403	individual	B1	B1	B1	[adjective]; [noun, singular or mass]
404	induced	UL	B2	B2	[adjective]; [verb, past participle]
405	infection	B2	B1	B1	[noun, singular or mass]
406	inference	UL	C1	C1	[noun, singular or mass]
407	inflammatory	UL	C1	C1	[adjective]
408	infrared	UL	UL	UL	[adjective]
409	inherent	C2	UL	C2	[adjective]
410	inhibition	C2	C1	C1	[noun, singular or mass]
411	initial	B1	B1	B1	[adjective]
412	initiation	UL	C1	C1	[noun, singular or mass]
413	injection	C2	B2	B2	[noun, singular or mass]
414	injury	B2	B1	B1	[noun, singular or mass]
415	innovation	C1	B2	B2	[noun, singular or mass]
416	input	B2	B2	B2	[noun, singular or mass]
417	insertion	UL	C1	C1	[noun, singular or mass]
418	inset	UL	UL	UL	[noun, singular or mass]; [verb, past participle]
419	insights	C1	B2	B2	[noun plural]
420	instance	B1	B1	B1	[noun, singular or mass]

No.	Word	EVP	GSE	Result	Parts of Speech
421	instructions	A2	B1	A2	[noun plural]
422	intact	C2	B2	B2	[adjective]
423	intake	UL	B2	B2	[noun, singular or mass]
424	integral	C1	B2	B2	[adjective]; [noun, singular or mass]
425	integrated	C1	B2	B2	[adjective]; [verb, past participle]
426	intensity	C2	UL	C2	[noun, singular or mass]
427	interactions	C1	B2	B2	[noun plural]
428	interestingly	UL	B2	B2	[adverb]
429	interface	UL	C1	C1	[noun, singular or mass]
430	interior	B2	B2	B2	[noun, singular or mass]; [adjective]
431	intermediate	B1	B1	B1	[adjective]; [noun, singular or mass]
432	internal	B2	B2	B2	[adjective]
433	interpretation	C2	B2	B2	[noun, singular or mass]
434	interval	B1	B2	B1	[noun, singular or mass]
435	intervention	C2	B2	B2	[noun, singular or mass]
436	intestinal	UL	C1	C1	[adjective]
437	intrinsic	C2	UL	C2	[adjective]
438	invasive	UL	C1	C1	[adjective]
439	inverse	UL	UL	UL	[adjective]; [noun, singular or mass]
440	investigated	B2	B2	B2	[verb, past participle]; [verb, past tense]
441	involved	B1	B2	B1	[verb, past participle]
442	ion	UL	C1	C1	[noun, singular or mass]
443	isolated	C1	B2	B2	[verb, past participle]
444	items	B1	B1	B1	[noun plural]
445	kernel	UL	C1	C1	[noun, singular or mass]
446	kidney	C2	B2	B2	[noun, singular or mass]
447	kinase	UL	UL	UL	[noun, singular or mass]
448	kinetic	UL	UL	UL	[adjective]
449	kit	B1	B1	B1	[noun, singular or mass]
450	labeled	C1	B2	B2	[verb, past participle]
451	laboratory	B1	B1	B1	[noun, singular or mass]
452	lactic	UL	UL	UL	[adjective]
453	laser	B2	B2	B2	[noun, singular or mass]
454	latent	UL	UL	UL	[adjective]
455	lateral	UL	C1	C1	[adjective]
456	lattice	UL	UL	UL	[noun, singular or mass]
457	layer	B2	B2	B2	[noun, singular or mass]
458	linear	UL	B2	B2	[adjective]
459	linked	B2	B1	B1	[verb, past participle]
460	lipid	UL	UL	UL	[noun, singular or mass]
461	liver	B2	B2	B2	[noun, singular or mass]
462	localization	UL	UL	UL	[noun, singular or mass]
463	location	B1	B1	B1	[noun, singular or mass]

No.	Word	EVP	GSE	Result	Parts of Speech
464	locus	UL	UL	UL	[noun, singular or mass]
465	longitudinal	UL	C1	C1	[adjective]
466	loop	UL	B2	B2	[noun, singular or mass]; [verb, present, non-3rd p.]
467	lysine	UL	UL	UL	[noun, singular or mass]
468	magnesium	UL	C1	C1	[noun, singular or mass]
469	magnetic	C1	B2	B2	[adjective]
470	magnification	UL	C1	C1	[noun, singular or mass]
471	magnitude	UL	C1	C1	[noun, singular or mass]
472	maintained	B2	B2	B2	[verb, past participle]
473	major	B2	B1	B1	[adjective]
474	manipulation	UL	B2	B2	[noun, singular or mass]
475	manually	UL	B2	B2	[adverb]
476	mapping	UL	B2	B2	[noun, singular or mass]; [verb, gerund/participle]
477	marginal	C2	C1	C1	[adjective]
478	marine	UL	B2	B2	[adjective]
479	mathematics	UL	B1	B1	[noun plural]
480	matrix	UL	B1	B1	[noun, singular or mass]
481	mature	B2	B2	B2	[adjective]; [verb, present, non-3rd p.]
482	maximal	UL	UL	UL	[adjective]
483	maximum	B1	B1	B1	[adjective]; [noun, singular or mass]
484	mechanism	C1	B2	B2	[noun, singular or mass]
485	media	B2	B2	B2	[noun plural]
486	median	UL	C1	C1	[adjective]; [noun, singular or mass]
487	mediated	UL	C1	C1	[adjective]; [verb, past participle]
488	medical	B2	A2	A2	[adjective]
489	medium	B1	B1	B1	[noun, singular or mass]
490	membrane	UL	C1	C1	[noun, singular or mass]
491	mesh	UL	C1	C1	[noun, singular or mass]
492	metabolic	UL	C1	C1	[adjective]
493	metabolism	UL	C1	C1	[noun, singular or mass]
494	metabolites	UL	UL	UL	[noun plural]
495	methanol	UL	UL	UL	[noun, singular or mass]
496	methionine	UL	UL	UL	[noun, singular or mass]
497	method	B1	B1	B1	[noun, singular or mass]
498	micro	UL	B2	B2	[noun, singular or mass]
499	microbial	UL	UL	UL	[adjective]
500	microorganisms	UL	UL	UL	[noun plural]
501	microscopy	UL	UL	UL	[noun, singular or mass]
502	migration	C1	B2	B2	[noun, singular or mass]
503	minimal	C1	B2	B2	[adjective]
504	minimize	C1	B2	B2	[verb, base form]
505	minimum	B1	B1	B1	[adjective]; [noun, singular or mass]

No.	Word	EVP	GSE	Result	Parts of Speech
506	minor	B2	B1	B1	[adjective]
507	mitochondrial	UL	UL	UL	[adjective]
508	mobile	UL	B2	B2	[adjective]
509	mode	C1	B2	B2	[noun, singular or mass]
510	modified	C1	B2	B2	[verb, past participle]; [adjective]
511	molar	UL	C1	C1	[adjective]
512	molecular	UL	C1	C1	[adjective]
513	monitoring	C1	B2	B2	[verb, gerund/participle]; [noun, singular or mass]
514	morphology	UL	C1	C1	[noun, singular or mass]
515	mortality	C2	B2	B2	[noun, singular or mass]
516	motivated	B2	B2	B2	[verb, past participle]
517	mounted	C2	B2	B2	[verb, past participle]; [adjective]
518	muscle	B2	B1	B1	[noun, singular or mass]
519	mutant	UL	C1	C1	[noun, singular or mass]; [adjective]
520	mutations	UL	C1	C1	[noun plural]
521	nanoparticles	UL	UL	UL	[noun plural]
522	negative	A2	B1	A2	[adjective]
523	negligible	C2	C1	C1	[adjective]
524	network	B2	B1	B1	[noun, singular or mass]
525	neural	UL	C1	C1	[adjective]
526	neutral	C1	B2	B2	[adjective]
527	nevertheless	B2	B2	B2	[adverb]
528	nitrogen	UL	B2	B2	[noun, singular or mass]
529	nodes	UL	C1	C1	[noun plural]
530	norm	C1	B2	B2	[noun, singular or mass]
531	normal	A2	A1	A1	[adjective]
532	notion	C1	C1	C1	[noun, singular or mass]
533	novel	UL	C1	C1	[adjective]
534	nuclear	B2	B1	B1	[adjective]
535	nucleotide	UL	UL	UL	[noun, singular or mass]
536	nucleus	UL	B2	B2	[noun, singular or mass]
537	null	UL	C1	C1	[adjective]
538	nutrient	UL	B2	B2	[adjective]; [noun, singular or mass]
539	nutritional	C1	C1	C1	[adjective]
540	objective	B2	B1	B1	[noun, singular or mass]
541	obtained	B2	B1	B1	[verb, past participle]
542	obvious	B1	B1	B1	[adjective]
543	occupied	B2	B2	B2	[verb, past participle]
544	occur	B2	B1	B1	[verb, present, non-3rd p.]
545	odd	B2	B1	B1	[adjective]
546	online	A2	A2	A2	[adjective]
547	onset	UL	C1	C1	[noun, singular or mass]

No.	Word	EVP	GSE	Result	Parts of Speech
548	optical	UL	B2	B2	[adjective]
549	optimal	UL	C1	C1	[adjective]
550	optimization	UL	C1	C1	[noun, singular or mass]
551	optimum	UL	C1	C1	[adjective]
552	option	B1	B1	B1	[noun, singular or mass]
553	oral	B2	B2	B2	[adjective]
554	organic	B2	B2	B2	[adjective]
555	organisms	UL	B2	B2	[noun plural]
556	orientation	C2	B2	B2	[noun, singular or mass]
557	outcomes	C1	B2	B2	[noun plural]
558	output	C2	B2	B2	[noun, singular or mass]
559	oven	B1	B1	B1	[noun, singular or mass]
560	overall	B2	B2	B2	[adjective]; [adverb]
561	overlap	C2	B2	B2	[verb, present, non-3rd p.]
562	overnight	B1	B2	B1	[adverb]; [adjective]
563	overview	C1	B2	B2	[noun, singular or mass]
564	oxide	UL	C1	C1	[noun, singular or mass]
565	oxidized	UL	UL	UL	[verb, past participle]
566	oxygen	B2	B1	B1	[noun, singular or mass]
567	panel	C1	B2	B2	[noun, singular or mass]
568	parallel	C2	B2	B2	[adjective]; [noun, singular or mass]; [verb, base form]
569	parameters	UL	UL	UL	[noun plural]
570	participants	C1	B2	B2	[noun plural]
571	partners	A2	B1	A2	[noun plural]
572	passive	B1	B2	B1	[adjective]
573	pathogen	UL	C1	C1	[noun, singular or mass]
574	pathogenic	UL	UL	UL	[adjective]
575	pathway	UL	B2	B2	[noun, singular or mass]
576	patients	B1	B1	B1	[noun plural]
577	peak	B2	B1	B1	[adjective]; [noun, singular or mass]; [verb, base form]
578	penetration	UL	UL	UL	[noun, singular or mass]
579	peptide	UL	UL	UL	[noun, singular or mass]
580	percentage	B2	B1	B1	[noun, singular or mass]
581	perception	C2	B2	B2	[noun, singular or mass]
582	period	B1	B1	B1	[noun, singular or mass]
583	peripheral	UL	C1	C1	[adjective]
584	peroxide	UL	UL	UL	[noun, singular or mass]
585	persistence	C2	C1	C1	[noun, singular or mass]
586	perspective	C1	B2	B2	[noun, singular or mass]
587	pharmaceutical	UL	B2	B2	[adjective]
588	phase	B2	B2	B2	[noun, singular or mass]
589	phenomenon	C1	B2	B2	[noun, singular or mass]

No.	Word	EVP	GSE	Result	Parts of Speech
590	phenotypic	UL	UL	UL	[adjective]
591	phosphate	UL	C1	C1	[noun, singular or mass]
592	phylogenetic	UL	UL	UL	[adjective]
593	physical	B2	B1	B1	[adjective]
594	physiological	UL	B2	B2	[adjective]
595	plasma	UL	C1	C1	[noun, singular or mass]
596	plastic	A2	A2	A2	[adjective]; [noun, singular or mass]
597	platform	A2	A2	A2	[noun, singular or mass]
598	plot	B2	B2	B2	[noun, singular or mass]; [verb, present, non-3rd p.]
599	plus	A2	B1	A2	[coordinating conjunction]; [noun, singular or mass]; [preposition/subord. conj.]
600	polar	UL	B2	B2	[adjective]
601	poly	UL	UL	UL	[noun, singular or mass]
602	polymer	UL	UL	UL	[noun, singular or mass]
603	polymerase	UL	UL	UL	[noun, singular or mass]
604	polynomial	UL	UL	UL	[noun, singular or mass]; [adjective]
605	pooled	UL	UL	UL	[verb, past participle]; [adjective]
606	pore	UL	C1	C1	[noun, singular or mass]
607	portion	B2	B1	B1	[noun, singular or mass]
608	pose	C1	B2	B2	[verb, base form]; [verb, present, non-3rd p.]
609	positive	B1	B1	B1	[adjective]
610	posterior	UL	UL	UL	[adjective]; [noun, singular or mass]
611	potassium	UL	C1	C1	[noun, singular or mass]
612	potent	UL	C1	C1	[adjective]
613	potential	B2	B2	B2	[adjective]; [noun, singular or mass]
614	preceding	C2	B2	B2	[adjective]; [verb, gerund/participle]
615	precipitation	UL	C1	C1	[noun, singular or mass]
616	precision	UL	C1	C1	[noun, singular or mass]
617	precursor	UL	UL	UL	[noun, singular or mass]
618	predicted	B1	B2	B1	[verb, past participle]
619	predominantly	C2	B2	B2	[adverb]
620	preliminary	C1	B2	B2	[adjective]
621	presumably	B2	B2	B2	[adverb]
622	prevalence	UL	C1	C1	[noun, singular or mass]
623	previous	B1	B1	B1	[adjective]
624	primary	B2	B1	B1	[adjective]
625	prime	C2	B2	B2	[adjective]
626	primers	UL	C1	C1	[noun plural]
627	principal	B1	B1	B1	[adjective]
628	principle	C1	B2	B2	[noun, singular or mass]
629	prior	C1	B2	B2	[adjective]; [adverb]
630	probe	UL	B2	B2	[noun, singular or mass]; [verb, base form]
631	procedure	B2	B2	B2	[noun, singular or mass]

No.	Word	EVP	GSE	Result	Parts of Speech
632	process	B2	B1	B1	[noun, singular or mass]
633	profile	B2	B2	B2	[noun, singular or mass]; [adjective]; [verb, base form]
634	progression	C1	B2	B2	[noun, singular or mass]
635	projection	C1	B2	B2	[noun, singular or mass]
636	proliferation	UL	C1	C1	[noun, singular or mass]
637	proline	UL	UL	UL	[noun, singular or mass]
638	promote	B1	B1	B1	[verb, base form]
639	propagation	UL	C1	C1	[noun, singular or mass]
640	proportion	C1	B2	B2	[noun, singular or mass]
641	protease	UL	UL	UL	[noun, singular or mass]
642	protein	C1	B2	B2	[noun, singular or mass]
643	protocol	UL	B2	B2	[noun, singular or mass]
644	publication	B2	B2	B2	[noun, singular or mass]
645	published	B1	B2	B1	[verb, past participle]
646	pulse	C1	B2	B2	[noun, singular or mass]
647	purchased	B2	B1	B1	[verb, past participle]
648	purified	UL	C1	C1	[verb, past participle]
649	purity	C1	B2	B2	[noun, singular or mass]
650	pursued	C1	B2	B2	[verb, past participle]
651	putative	UL	UL	UL	[adjective]
652	quantification	UL	UL	UL	[noun, singular or mass]
653	radical	C1	B2	B2	[adjective]
654	random	C1	B2	B2	[adjective]
655	range	B1	B1	B1	[noun, singular or mass]
656	ratio	C1	B2	B2	[noun, singular or mass]
657	rational	C1	B2	B2	[adjective]
658	reaction	B2	B1	B1	[noun, singular or mass]
659	reagents	UL	UL	UL	[noun plural]
660	receptor	UL	C1	C1	[noun, singular or mass]
661	recovery	B2	B2	B2	[noun, singular or mass]
662	redox	UL	UL	UL	[noun, singular or mass]
663	reference	B2	B2	B2	[noun, singular or mass]; [preposition/subord. conj.]
664	refinement	UL	UL	UL	[noun, singular or mass]
665	regime	C2	B2	B2	[noun, singular or mass]
666	region	B1	B1	B1	[noun, singular or mass]
667	regression	UL	UL	UL	[noun, singular or mass]
668	regulation	B2	B2	B2	[noun, singular or mass]
669	rejection	C2	B2	B2	[noun, singular or mass]
670	relaxation	B1	B2	B1	[noun, singular or mass]
671	release	B2	B2	B2	[noun, singular or mass]; [verb, present, non-3rd p.]
672	relevant	B2	B2	B2	[adjective]

No.	Word	EVP	GSE	Result	Parts of Speech
673	reliable	B1	B2	B1	[adjective]
674	removed	B1	B1	B1	[verb, past participle]
675	replication	UL	C1	C1	[verb, past participle]
676	required	B1	B2	B1	[verb, past participle]
677	research	B1	B1	B1	[noun, singular or mass]
678	reservoirs	UL	B2	B2	[noun plural]
679	residence	C2	B2	B2	[noun, singular or mass]
680	residues	UL	C1	C1	[noun plural]
681	resolution	C2	B2	B2	[noun, singular or mass]
682	resonance	UL	UL	UL	[noun, singular or mass]
683	resources	B2	B2	B2	[noun plural]
684	response	B2	B2	B2	[noun, singular or mass]
685	restoration	UL	B2	B2	[noun, singular or mass]
686	restricted	C1	B2	B2	[verb, past participle]; [adjective]
687	resuspended	UL	UL	UL	[verb, past participle]; [adjective]
688	retention	C2	C1	C1	[noun, singular or mass]
689	revealed	B2	B2	B2	[verb, past tense]; [verb, past participle]
690	reverse	B2	B2	B2	[adjective]; [verb, base form]; [noun, singular or mass]
691	revision	B1	B2	B1	[noun, singular or mass]
692	robust	UL	C1	C1	[adjective]
693	role	B1	B1	B1	[noun, singular or mass]
694	rotation	UL	B2	B2	[noun, singular or mass]
695	route	B1	A2	A2	[noun, singular or mass]
696	routine	B1	B1	B1	[adjective]; [noun, singular or mass]
697	salinity	UL	UL	UL	[noun, singular or mass]
698	saturation	UL	C1	C1	[noun, singular or mass]
699	scaling	UL	UL	UL	[noun, singular or mass]; [verb, gerund/participle]
700	scan	C1	B2	B2	[noun, singular or mass]; [verb, base form]
701	scenario	C2	B2	B2	[noun, singular or mass]
702	scheme	B2	B2	B2	[noun, singular or mass]
703	score	B1	A2	A2	[noun, singular or mass]
704	seasonal	B2	B2	B2	[adjective]
705	secretion	UL	C1	C1	[noun, singular or mass]
706	section	B1	B1	B1	[noun, singular or mass]
707	sector	C1	B2	B2	[noun, singular or mass]
708	seek	B2	B2	B2	[verb, base form]; [verb, present, non-3rd p.]
709	segment	UL	B2	B2	[noun, singular or mass]; [verb, base form]
710	selected	B1	B1	B1	[adjective]; [verb, past participle]
711	sensing	C2	B2	B2	[verb, gerund/participle]
712	sensor	UL	B2	B2	[noun, singular or mass]
713	sequence	C1	B2	B2	[noun, singular or mass]
714	series	B1	B2	B1	[noun, singular or mass]

No.	Word	EVP	GSE	Result	Parts of Speech
715	serine	UL	UL	UL	[noun, singular or mass]
716	serum	UL	C1	C1	[noun, singular or mass]
717	setup	UL	UL	UL	[noun, singular or mass]
718	sex	B1	A2	A2	[noun, singular or mass]
719	shift	B2	B2	B2	[noun, singular or mass]; [verb, base form]
720	significant	B2	B1	B1	[adjective]
721	silica	UL	UL	UL	[noun, singular or mass]
722	silicon	UL	B2	B2	[noun, singular or mass]
723	similar	B1	B1	B1	[adjective]
724	simulation	C1	B2	B2	[noun, singular or mass]
725	simultaneously	B2	B2	B2	[adverb]
726	site	A2	B1	A2	[noun, singular or mass]
727	skeletal	UL	C1	C1	[adjective]
728	sodium	UL	C1	C1	[noun, singular or mass]
729	software	A2	B1	A2	[noun, singular or mass]
730	solely	C1	B2	B2	[adverb]
731	soluble	UL	C1	C1	[adjective]
732	solvent	UL	C1	C1	[adjective]; [noun, singular or mass]
733	somewhat	C1	B2	B2	[adverb]
734	source	B2	B1	B1	[noun, singular or mass]
735	spatial	UL	C1	C1	[adjective]
736	species	B2	B2	B2	[adjective]
737	specific	B2	B1	B1	[adjective]
738	specified	B2	B2	B2	[verb, past participle]; [adjective]
739	spectra	UL	UL	UL	[noun, singular or mass]
740	spectral	UL	C1	C1	[adjective]
741	sphere	C1	B2	B2	[noun, singular or mass]
742	spontaneous	UL	B2	B2	[adjective]
743	stability	C1	B2	B2	[noun, singular or mass]
744	static	UL	B2	B2	[adjective]
745	statistical	C1	B2	B2	[adjective]
746	status	C1	B2	B2	[noun, singular or mass]
747	storage	B2	B2	B2	[noun, singular or mass]
748	straightforward	B2	B2	B2	[adjective]
749	strains	B2	B2	B2	[noun plural]
750	strand	C2	B2	B2	[noun, singular or mass]
751	strategy	B2	B2	B2	[noun, singular or mass]
752	stress	B1	A2	A2	[noun, singular or mass]
753	structure	B2	B2	B2	[noun, singular or mass]; [verb, base form]
754	subsequent	C1	B2	B2	[adjective]
755	subset	UL	C1	C1	[noun, singular or mass]
756	substitution	C1	B2	B2	[noun, singular or mass]
757	successive	C2	B2	B2	[adjective]

No.	Word	EVP	GSE	Result	Parts of Speech
758	sufficient	B2	B2	B2	[adjective]
759	sum	B1	B1	B1	[noun, singular or mass]
760	summary	B2	B1	B1	[noun, singular or mass]
761	superior	C1	B2	B2	[adjective]
762	supernatant	UL	UL	UL	[adjective]
763	supplementary	C2	B2	B2	[adjective]
764	suppression	UL	C1	C1	[noun, singular or mass]
765	survey	B2	B1	B1	[noun, singular or mass]; [verb, present, non-3rd p.]
766	survival	B2	B1	B1	[noun, singular or mass]
767	susceptible	UL	B2	B2	[adjective]
768	suspension	UL	B2	B2	[noun, singular or mass]
769	sustainable	C1	B2	B2	[adjective]
770	switch	B2	B2	B2	[noun, singular or mass]; [verb, base form]
771	symbols	B2	B1	B1	[noun plural]
772	symmetry	UL	B2	B2	[noun, singular or mass]
773	symptoms	B2	B1	B1	[noun plural]
774	synergistic	UL	UL	UL	[adjective]
775	synthesis	UL	C1	C1	[noun, singular or mass]
776	synthetic	UL	B2	B2	[adjective]
777	tank	C2	B2	B2	[noun, singular or mass]
778	target	B2	B1	B1	[noun, singular or mass]; [verb, present, non-3rd p.]
779	task	B2	A2	A2	[noun, singular or mass]
780	taxonomic	UL	UL	UL	[adjective]
781	team	A2	B1	A2	[noun, singular or mass]
782	technical	B2	B1	B1	[adjective]
783	techniques	B1	B1	B1	[noun plural]
784	technology	B1	A2	A2	[noun, singular or mass]
785	template	UL	B2	B2	[noun, singular or mass]
786	temporal	UL	C1	C1	[adjective]
787	terminal	B2	B1	B1	[adjective]; [noun, singular or mass]
788	tertiary	C1	UL	C1	[adjective]
789	text	A2	A2	A2	[noun, singular or mass]
790	theory	B2	B2	B2	[noun, singular or mass]
791	therapeutic	C1	B2	B2	[adjective]
792	therapy	B2	B2	B2	[noun, singular or mass]
793	thereby	C1	B2	B2	[adverb]
794	thermal	UL	B2	B2	[adjective]
795	threshold	C2	C1	C1	[noun, singular or mass]
796	tissue	UL	B2	B2	[noun, singular or mass]
797	tolerance	C2	UL	C2	[noun, singular or mass]
798	topic	B1	B1	B1	[noun, singular or mass]
799	toxic	B2	B2	B2	[adjective]

No.	Word	EVP	GSE	Result	Parts of Speech
800	toxin	UL	B2	B2	[noun, singular or mass]
801	trace	B2	B2	B2	[noun, singular or mass]; [verb, base form]
802	tract	UL	C1	C1	[noun, singular or mass]
803	traditional	B1	B1	B1	[adjective]
804	transcription	UL	UL	UL	[noun, singular or mass]
805	transfer	B1	B2	B1	[noun, singular or mass]; [verb, base form]
806	transformation	C1	B2	B2	[noun, singular or mass]
807	transient	UL	C1	C1	[adjective]
808	transition	C2	B2	B2	[noun, singular or mass]
809	transmission	C2	B2	B2	[noun, singular or mass]
810	transport	B1	B1	B1	[noun, singular or mass]; [verb, base form]
811	trend	B1	A2	A2	[noun, singular or mass]
812	triangle	B2	B1	B1	[noun, singular or mass]
813	trigger	C1	B2	B2	[verb, base form]; [noun, singular or mass]
814	triple	UL	B2	B2	[adjective]
815	triplicate	UL	UL	UL	[noun, singular or mass]; [adjective]
816	tumor	UL	B2	B2	[noun, singular or mass]
817	tyrosine	UL	UL	UL	[noun, singular or mass]
818	ultimately	C1	B2	B2	[adverb]
819	unclear	B2	B1	B1	[adjective]
820	undergo	C1	B2	B2	[verb, base form]; [verb, present, non-3rd p.]
821	underlying	C2	UL	C2	[verb, gerund/participle]
822	uniform	A2	A2	A2	[adjective]; [noun, singular or mass]
823	unique	B2	B2	B2	[adjective]
824	untreated	UL	B2	B2	[adjective]
825	uptake	UL	UL	UL	[noun, singular or mass]
826	urea	UL	UL	UL	[noun, singular or mass]
827	utilized	UL	UL	UL	[verb, past participle]; [verb, past tense]
828	validation	UL	B2	B2	[noun, singular or mass]
829	variables	C1	B2	B2	[noun plural]
830	vascular	UL	C1	C1	[adjective]
831	velocity	UL	C1	C1	[noun, singular or mass]
832	verify	C1	B2	B2	[verb, base form]
833	version	B2	B2	B2	[noun, singular or mass]
834	vertical	C1	B2	B2	[adjective]
835	via	B1	B1	B1	[preposition/subord. conj.]
836	viability	UL	UL	UL	[noun, singular or mass]
837	video	A2	A2	A2	[noun, singular or mass]
838	violation	UL	C1	C1	[noun, singular or mass]
839	virtually	B2	B2	B2	[adverb]
840	virus	B2	B1	B1	[noun, singular or mass]
841	vitamin	B2	B2	B2	[noun, singular or mass]
842	volatile	UL	C1	C1	[adjective]

No.	Word	EVP	GSE	Result	Parts of Speech
843	voltage	UL	C1	C1	[noun, singular or mass]
844	volume	B1	B1	B1	[noun, singular or mass]
845	vs	C1	B2	B2	[coordinating conjunction]; [preposition/subord. conj.]; [proper noun, singular]
846	wavelength	C2	C1	C1	[noun, singular or mass]
847	weighted	UL	UL	UL	[adjective]; [verb, past participle]
848	whereas	B2	B2	B2	[preposition/subord. conj.]
849	widespread	C1	B2	B2	[adjective]
850	worldwide	B2	B1	B1	[adverb]; [adjective]
851	yeast	UL	C1	C1	[noun, singular or mass]
852	zinc	UL	C1	C1	[noun, singular or mass]
853	zone	B1	B2	B1	[noun, singular or mass]

APPENDIX E
THE CEFR-MAPPED SAWL SUBLISTS
(HIGH FREQUENCY WORDS)

Sublist 1 [A1]

acid	adult	normal

Sublist 2 [A2]

alcohol	file	route
area	finally	score
authors	global	sex
available	goal	site
biology	grade	software
channel	instructions	stress
chemistry	medical	task
classical	negative	team
community	online	technology
contact	partners	text
culture	plastic	trend
digital	platform	uniform
energy	plus	video

Sublist 3 [B1]

academic	code	drug
access	column	economic
accompanied	comments	emphasis
achieved	communication	environmental
affect	complex	error
alternative	conclusion	established
analysis	conference	eventually
annual	confirmed	evidence
approach	context	experts
appropriate	created	feedback
approximately	cycle	feeding
architecture	data	focus
atmosphere	database	function
attached	decades	generation
attitude	defined	graph
bacteria	definite	height
breast	demonstrated	highlighted
briefly	depression	identical
cancer	design	identified
candidate	despite	image
categories	diagram	index
challenge	diet	indicated
chemical	differential	individual
climate	display	infection

initial
injury
instance
intermediate
interval
involved
items
kit
laboratory
linked
location
major
mathematics
matrix
maximum
medium
method
minimum
minor
muscle
network
nuclear
objective
obtained
obvious
occur
odd
option
oven
overnight
oxygen
passive
patients
peak
percentage
period
physical
portion
positive
predicted
previous
primary
principal
process
promote
published
purchased
range
reaction
region
relaxation
reliable
removed
required
research
revision
role
routine
section
selected
series
significant
similar
source
specific
sum
summary
survey
survival
symbols
symptoms
target
technical
techniques
terminal
topic
traditional
transfer
transport
triangle
unclear
via
virus
volume
worldwide
zone

Sublist 4 [B2]

abstract
accommodate
accuracy
acute
adapted
adequate
adjacent
adjusted
administration
adverse
aid
altered
apparent
appendix
aspects
assembly
assessed
assigned
assisted
assume
atoms
attained
attributed
automated
aware
barrier
beneficial
bias
bonds
breakdown
bulk
calcium
capable
capacity
capture
carbon
cardiac
chamber
chip
cholesterol
chronic
clarify
clinical
clones
coincide
colon
compatible
compensate

components
composition
compounds
comprehensive
comprised
concept
conducted
consequently
considerable
consistent
constant
constituents
constraints
construct
consumption
contamination
contrast
contribute
conventional
conversion
cooperation
coordinates
core
corresponding
crucial
crude
crystal
dashed
decline
defects
density
depicted
deposited
developmental
device
devoted
digestion
dimensional
discrepancy
discrimination
disruption
dissolved
distinct
distribution
diversity
documented
dominant
donor
dramatically
dual
duration
ecological
edited
efficiently
elements
elimination
enable
encountered
engineered
ensure
enzyme
equation
equipped
equivalent
estimated
evaluated
evolution
exceed
excluded
exhibited
expanded
explicit
exploited
exposure
external
facilitate
factors
favorable
features
fiber
filter
flexible
fluid
formula
fraction
fragments
framework
frequencies
fundamental
furthermore
fusion
gel
gene
generated
genetic
geographic
geometry
germ
glucose
grid
guarantee
gut
hence
hydrogen
hypothesis
illustrated
imaging
immune
impact
implemented
implications
implies
imposed
incorporated
induced
injection
innovation
input
insights
intact
intake
integral
integrated
interactions
interestingly
interior
internal
interpretation
intervention
investigated
isolated
kidney
labeled
laser
layer
linear
liver
loop
magnetic
maintained
manipulation

manually
mapping
marine
mature
mechanism
media
micro
migration
minimal
minimize
mobile
mode
modified
monitoring
mortality
motivated
mounted
neutral
nevertheless
nitrogen
norm
nucleus
nutrient
occupied
optical
oral
organic
organisms
orientation
outcomes
output
overall
overlap
overview
panel
parallel
participants
pathway
perception
perspective
pharmaceutical
phase
phenomenon
physiological
plot
polar
pose
potential
preceding
predominantly
preliminary
presumably
prime
principle
prior
probe
procedure
profile
progression
projection
proportion
protein
protocol
publication
pulse
purity
pursued
radical
random
ratio
rational
recovery
reference
regime
regulation
rejection
release
relevant
reservoirs
residence
resolution
resources
response
restoration
restricted
revealed
reverse
rotation
scan
scenario
scheme
seasonal
sector
seek
segment
sensing
sensor
sequence
shift
silicon
simulation
simultaneously
solely
somewhat
species
specified
sphere
spontaneous
stability
static
statistical
status
storage
straightforward
strains
strand
strategy
structure
subsequent
substitution
successive
sufficient
superior
supplementary
susceptible
suspension
sustainable
switch
symmetry
synthetic
tank
template
theory
therapeutic
therapy
thereby
thermal
tissue
toxic
toxin
trace

transformation
transition
transmission
trigger
triple
tumor
ultimately
undergo
unique
untreated
validation
variables
verify
version
vertical
virtually
vitamin
vs
whereas
widespread

Sublist 5 [C1]

absorption
abundance
acquisition
adhesion
affinity
alignment
alkaline
allocation
ammonia
amplification
amplitude
analogous
antioxidant
aqueous
aromatic
array
assay
baseline
batch
beads
biochemical
biomass
biomedical
biotechnology
bovine
buffer
calibrate
capillary
carbohydrate
cardiovascular
catalytic
cavity
cellular
cellulose
chloride
coating
coefficient
complementary
computational
configuration
conversely
correlation
coupled
criteria
dataset
deficient
degradation
denote
dependence
depletion
deposition
detection
differentiation
diffraction
diffusion
discrete
displacement
dissolution
distilled
divergence
domain
downstream
droplet
dye
dynamics
electrode
electron
elevated
embedded
emergence
emission
empirical
enrichment
ethanol
evaporation
extraction
feasible
fermentation
finite
fluctuations
fungal
genome
genus
hierarchical
hybrid
incidence
incubated
inference
inflammatory
inhibition
initiation
insertion
interface
intestinal
invasive
ion
kernel
lateral
longitudinal
magnesium
magnification
magnitude
marginal
median
mediated
membrane
mesh
metabolic
metabolism
molar
molecular
morphology
mutant
mutations

negligible
neural
nodes
notion
novel
null
nutritional
onset
optimal
optimization
optimum
oxide
pathogen
peripheral
persistence
phosphate
plasma
pore
potassium
potent
precipitation
precision
prevalence
primers
proliferation
propagation
purified
receptor
replication
residues
retention
robust
saturation
secretion
serum
skeletal
sodium
soluble
solvent
spatial
spectral
subset
suppression
synthesis
temporal
tertiary
threshold
tract
transient
vascular
velocity
violation
volatile
voltage
wavelength
yeast
zinc

Sublist 6 [C2]

arbitrary
entities
inherent
intensity
intrinsic
tolerance
underlying

Sublist 7 [UL]

acetate
acetic
acetone
activation
additive
agar
agarose
aggregation
albumin
allele
ambient
amino
anaerobic
annealing
antimicrobial
apoptosis
bacillus
basal
bioactive
biosynthesis
cascade
centrifugation
cerevisiae
chromatography
chromosome
concurrent
confocal
conformation
conserved
covalent
cumulative
cysteine
decomposition
dehydrogenase
diluted
dispersion
ecosystems
electrophoresis
electrostatic
elemental
elongation
encoding
endogenous
enzymatic
epithelial
esters
excitation
exogenous
exponential
fabricated
favorably
fluorescence
flux
functionalized
genotypes
glycerol
hydrolysis
hydrophilic
hydrophobic
hydroxyl
immobilized
infrared
inset

inverse
kinase
kinetic
lactic
latent
lattice
lipid
localization
locus
lysine
maximal
metabolites
methanol
methionine
microbial
microorganisms
microscopy
mitochondrial
nanoparticles
nucleotide
oxidized
parameters
pathogenic
penetration
peptide
peroxide
phenotypic
phylogenetic
poly
polymer
polymerase
polynomial
pooled
posterior
precursor
proline
protease
putative
quantification
reagents
redox
refinement
regression
resonance
resuspended
salinity
scaling
serine
setup
silica
spectra
supernatant
synergistic
taxonomic
transcription
triplicate
tyrosine
uptake
urea
utilized
viability
weighted

APPENDIX F
THE SAWL COLLOCATION LIST
(HIGH FREQUENCY WORDS)

Part 1

The list of 234 SAWL coinciding with the ACL (Ackermann & Chen, 2013)

*Note: 1st=1st Component (eg. **abstract** concept);2nd=2nd Component (eg.allow **access**)*

1.abstract 1st
2.academic 1st
3.access 2nd
4.accuracy 2nd
5.acute 2nd
6.administration 2nd
7.adverse 1st
8.affect 2nd
9.alternative 1st/2nd
10.analysis 2nd
11.annual 1st
12.apparent 2nd
13.approach 2nd
14.appropriate 1st/2nd
15.area 2nd
16.array 2nd
17.aspects 2nd
18.assigned 2nd
19.assume 1st
20.attitude 2nd
21.available 1st/2nd
22.aware 2nd
23.beneficial 1st
24.briefly 1st
25.capacity 2nd
26.challenge 2nd
27.chemical 1st
28.classical 1st
29.climate 1st/2nd
30.communication 2nd
31.community 2nd
32.complex 1st/2nd
33.comprehensive 1st
34.concept 2nd
35.conclusion 2nd
36.conference 2nd
37.considerable 1st
38.consistent 1st
39.constant 1st/2nd
40.constraints 2nd
41.construct 2nd
42.contact 2nd
43.context 2nd
44.contrast 2nd
45.contribute 1st
46.conventional 1st
47.core 1st/2nd
48.correlation 2nd
49.created 2nd
50.criteria 2nd
51.crucial 1st
52.culture 2nd
53.data 1st/2nd
54.database 2nd
55.decades 2nd
56.defined 2nd
57.degradation 2nd
58.demonstrated 2nd
59.design 2nd
60.developmental 1st
61.digital 1st
62.discrimination 2nd
63.display 2nd
64.distinct 1st
65.distribution 2nd
66.diversity 2nd
67.documented 2nd
68.domain 2nd
69.dominant 1st
70.dramatically 2nd
71.duration 2nd
72.economic 1st
73.elements 2nd
74.embedded 2nd
75.emphasis 2nd
76.empirical 1st
77.encountered 2nd
78.energy 2nd
79.environmental 1st
80.equivalent 2nd
81.error 2nd
82.established 1st/2nd
83.evidence 2nd
84.evolution 2nd
85.explicit 2nd
86.external 1st
87.facilitate 1st
88.factors 2nd
89.features 2nd
90.feedback 2nd
91.finite 1st
92.flexible 1st
93.focus 1st/2nd
94.formula 2nd
95.fraction 2nd
96.framework 2nd
97.function 2nd
98.fundamental 1st
99.generation 2nd
100.genetic 1st
101.geographic 1st/2nd
102.global 1st
103.goal 2nd
104.hierarchical 1st
105.identical 2nd
106.identified 2nd
107.image 2nd
108.impact 2nd
109.implications 2nd
110.incidence 2nd
111.individual 1st/2nd
112.initial 1st
113.innovation 2nd
114.instructions 2nd
115.integral 1st
116.integrated 1st/2nd
117.intensity 2nd
118.internal 1st
119.interpretation 2nd
120.intervention 2nd

121.intrinsic 1st
122.involved 2nd
123.linear 1st
124.linked 2nd
125.location 2nd
126.longitudinal 1st
127.major 1st
128.maximum 1st
129.media 1st/2nd
130.medical 1st
131.method 2nd
132.minimum 1st
133.minor 1st/2nd
134.modified 1st
135.motivated 2nd
136.negative 1st
137.network 2nd
138.norm 2nd
139.normal 1st
140.nuclear 1st
141.objective 1st/2nd
142.obvious 1st/2nd
143.occur 1st
144.online 1st
145.optimal 1st
146.oral 1st
147.orientation 2nd
148.overall 1st
149.overview 2nd
150.panel 2nd
151.parameters 2nd
152.peak 2nd
153.percentage 2nd
154.perception 2nd
155.period 2nd
156.perspective 2nd
157.phase 2nd
158.phenomenon 2nd
159.physical 1st
160.portion 2nd
161.pose 1st
162.positive 1st
163.potential 1st/2nd
164.preceding 1st/2nd
165.preliminary 1st
166.previous 1st
167.primary 1st
168.prime 1st
169.principal 1st
170.principle 2nd
171.prior 1st
172.procedure 2nd
173.process 1st/2nd
174.profile 2nd
175.promote 1st
176.proportion 2nd
177.published 1st
178.radical 1st
179.random 1st
180.range 2nd
181.reaction 2nd
182.reference 2nd
183.relevant 1st/2nd
184.reliable 1st
185.removed 2nd
186.research 1st/2nd
187.resolution 2nd
188.resources 2nd
189.response 2nd
190.role 2nd
191.score 2nd
192.section 2nd
193.sector 2nd
194.seek 1st
195.selected 2nd
196.sex 2nd
197.shift 2nd
198.significant 1st/2nd
199.similar 1st/2nd
200.source 1st/2nd
201.species 2nd
202.specific 1st/2nd
203.sphere 2nd
204.stability 2nd
205.statistical 1st
206.status 2nd
207.straightforward 2nd
208.strategy 2nd
209.stress 1st/2nd
210.structure 2nd
211.subsequent 1st
212.sufficient 1st
213.summary 2nd
214.superior 1st
215.survey 1st/2nd
216.target 1st/2nd
217.task 2nd
218.technical 1st
219.technology 2nd
220.text 2nd
221.theory 2nd
222.topic 2nd
223.traditional 1st
224.transformation 2nd
225.transition 2nd
226.transport 1st/2nd
227.trend 2nd
228.unclear 2nd
229.undergo 1st
230.underlying 1st
231.unique 1st
232.version 2nd
233.virtually 1st
234.widespread 1st/2nd

Part 2

2.1 Component 1 *Modified from Ackermann and Chen (2013)*

#	AD	Component I	POS I	Component II	POS II	AD	Sci
1		**abstract**	adj	concept	n		sci
2		**academic**	adj	achievement	n		sci
3		**academic**	adj	career	n		
4	(in)	**academic**	adj	circles	n		
5		**academic**	adj	community	n		
6		**academic**	adj	debate	n		
7		**academic**	adj	discipline	n		sci
8		**academic**	adj	discourse	n		
9		**academic**	adj	institution	n		
10		**academic**	adj	journal	n		sci
11		**academic**	adj	life	n		
12		**academic**	adj	performance	n		sci
13		**academic**	adj	research	n		sci
14		**academic**	adj	skills	n		
15		**academic**	adj	study	n		
16		**academic**	adj	success	n		
17		**academic**	adj	work	n		
18		**academic**	adj	world	n		
19		**academic**	adj	writing	n		
20		**academic**	adj	year	n		sci
21		**adverse**	adj	effect	n		sci
22		**adverse**	adj	reaction	n		sci
23		**alternative**	adj	approach	n		sci
24		**alternative**	adj	explanation	n		sci
25		**alternative**	adj	form	n		sci
26		**alternative**	adj	interpretation	n		sci
27		**alternative**	adj	means	n		sci
28		**alternative**	adj	method	n		sci
29		**alternative**	adj	model	n		sci
30		**alternative**	adj	solution	n		sci
31		**alternative**	adj	source	n		sci
32		**alternative**	adj	strategy	n		sci
33		**alternative**	adj	view	n		sci
34		**alternative**	adj	way	n		sci
35		**annual**	adj	conference	n		sci
36		**annual**	adj	meeting	n		sci
37		**annual**	adj	rate	n		sci
38		**annual**	adj	report	n		sci
39		**annual**	adj	review	n		
40		**appropriate**	adj	action	n		
41		**appropriate**	adj	behaviour	n		

#	AD	Component I	POS I	Component II	POS II	AD	Sci
42		**appropriate**	adj	conditions	n		sci
43		**appropriate**	adj	data	n		sci
44		**appropriate**	adj	form	n		
45		**appropriate**	adj	language	n		sci
46		**appropriate**	adj	level	n		sci
47		**appropriate**	adj	point	n		
48		**appropriate**	adj	response	n		
49		**appropriate**	adj	skills	n		
50		**appropriate**	adj	treatment	n		sci
51		**appropriate**	adj	way	n		sci
52		**assume**	v	responsibility	n		
53		**available**	adj	data	n		sci
54		**available**	adj	evidence	n		sci
55		**available**	adj	information	n		sci
56		**available**	adj	resources	n		sci
57		**beneficial**	adj	effect	n		sci
58		**biological**	adj	sex	n		
59		**briefly**	adv	describe	v		sci
60		**briefly**	adv	discuss	v		sci
61		**broad**	adj	definition	n		sci
62	(a)	**broad**	adj	range	n	(of)	sci
63		**broad**	adj	spectrum	n		sci
64		**chemical**	adj	reaction	n		sci
65		**classical**	adj	theory	n		sci
66		**climate**	n	change	n		sci
67		**complex**	adj	area	n		sci
68		**complex**	adj	interaction	n		sci
69		**complex**	adj	issue	n		sci
70		**complex**	adj	pattern	n		sci
71		**complex**	adj	problem	n		sci
72		**complex**	adj	process	n		sci
73		**complex**	adj	question	n		sci
74		**complex**	adj	relationship	n		sci
75		**complex**	adj	set	n		sci
76		**complex**	adj	situation	n		sci
77		**complex**	adj	structure	n		sci
78		**complex**	adj	system	n		sci
79		**comprehensive**	adj	account	n		sci
80		**comprehensive**	adj	approach	n		sci
81		**comprehensive**	adj	overview	n		sci
82		**comprehensive**	adj	review	n		sci
83		**comprehensive**	adj	system	n		
84	(a)	**considerable**	adj	amount	n	(of)	sci

#	AD	Component I	POS I	Component II	POS II	AD	Sci
85		**considerable**	adj	attention	n		sci
86		**considerable**	adj	debate	n		sci
87	(a)	**considerable**	adj	degree	n	(of)	
88	(in)	**considerable**	adj	detail	n		
89		**considerable**	adj	effort	n		sci
90		**considerable**	adj	evidence	n		sci
91	(to a)	**considerable**	adj	extent	n		sci
92	(be of)	**considerable**	adj	importance	n		sci
93		**considerable**	adj	influence	n		sci
94		**considerable**	adj	interest	n		sci
95		**considerable**	adj	research	n		sci
96		**considerable**	adj	support	n		
97		**considerable**	adj	variation	n		sci
98		**consistent**	adj	pattern	n		sci
99		**consistent**	adj	results	n		sci
100		**constant**	adj	rate	n		sci
101		**contribute**	v	significantly	adv		sci
102		**conventional**	adj	view	n		
103		**conventional**	adj	wisdom	n		
104		**core**	adj	area	n		sci
105		**core**	adj	element	n		sci
106		**core**	adj	issue	n		
107		**core**	adj	skills	n		
108		**core**	adj	value	n		
109		**crucial**	adj	difference	n		sci
110		**crucial**	adj	factor	n		sci
111		**crucial**	adj	importance	n		sci
112		**crucial**	adj	part	n		sci
113		**crucial**	adj	point	n		sci
114		**crucial**	adj	question	n		sci
115		**crucial**	adj	role	n		sci
116		**data**	n	gathering	n		
117		**data**	n	set	n		sci
118		**developmental**	adj	process	n		sci
119		**developmental**	adj	stage	n		sci
120		**digital**	adj	information	n		sci
121		**digital**	adj	media	n		
122		**digital**	adj	technology	n		sci
123		**distinct**	adj	group	n		sci
124		**distinct**	adj	type	n		sci
125		**dominant**	adj	culture	n		
126		**dominant**	adj	discourse	n		
127		**dominant**	adj	form	n		sci

#	AD	Component I	POS I	Component II	POS II	AD	Sci
128		**dominant**	adj	group	n		sci
129		**dominant**	adj	ideology	n		
130		**dominant**	adj	paradigm	n		
131		**dominant**	adj	position	n		sci
132		**dominant**	adj	role	n		sci
133		**economic**	adj	activity	n		
134		**economic**	adj	affairs	n		
135		**economic**	adj	analysis	n		sci
136		**economic**	adj	benefits	n		sci
137		**economic**	adj	change	n		
138		**economic**	adj	conditions	n		sci
139		**economic**	adj	consequences	n		sci
140		**economic**	adj	context	n		
141		**economic**	adj	crisis	n		
142		**economic**	adj	exploitation	n		
143		**economic**	adj	factors	n		sci
144		**economic**	adj	forces	n		
145		**economic**	adj	goal	n		
146		**economic**	adj	growth	n		sci
147		**economic**	adj	inequality	n		
148		**economic**	adj	integration	n		
149		**economic**	adj	interests	n		
150		**economic**	adj	policy	n		
151		**economic**	adj	power	n		
152		**economic**	adj	prosperity	n		
153		**economic**	adj	reform	n		
154		**economic**	adj	relations	n		
155		**economic**	adj	relationships	n		sci
156		**economic**	adj	resources	n		sci
157		**economic**	adj	sector	n		
158		**economic**	adj	stability	n		
159		**economic**	adj	status	n		
160		**economic**	adj	structure	n		
161		**economic**	adj	success	n		sci
162		**economic**	adj	system	n		sci
163		**economic**	adj	theory	n		
164		**economic**	adj	value	n		sci
165		**economic**	adj	welfare	n		
166		**empirical**	adj	data	n		sci
167		**empirical**	adj	evidence	n		sci
168		**empirical**	adj	investigation	n		sci
169		**empirical**	adj	research	n		sci
170		**empirical**	adj	study	n		sci

#	AD	Component I	POS I	Component II	POS II	AD	Sci
171		**empirical**	adj	support	n		sci
172		**empirical**	adj	work	n		sci
173		**environmental**	adj	changes	n		sci
174		**environmental**	adj	concern	n		sci
175		**environmental**	adj	consequences	n		sci
176		**environmental**	adj	damage	n		sci
177		**environmental**	adj	degradation	n		sci
178		**environmental**	adj	effects	n		sci
179		**environmental**	adj	factors	n		sci
180		**environmental**	adj	impact	n		sci
181		**environmental**	adj	issues	n		sci
182		**environmental**	adj	policy	n		
183		**environmental**	adj	pollution	n		sci
184		**environmental**	adj	protection	n		sci
185		**established**	adj	order	n		
186		**established**	adj	practice	n		sci
187		**established**	adj	principle	n		sci
188		**external**	adj	environment	n		sci
189		**external**	adj	factors	n		sci
190		**external**	adj	forces	n		sci
191		**external**	adj	influences	n		sci
192		**external**	adj	source	n		
193		**external**	adj	threat	n		sci
194		**external**	adj	world	n		sci
195		**facilitate (the)**	v	development	n		sci
196		**finite**	adj	number	n		sci
197		**flexible**	adj	approach	n		sci
198		**focus**	v	attention	n	(on)	sci
199		**fundamental**	adj	aspect	n		sci
200		**fundamental**	adj	assumption	n		sci
201		**fundamental**	adj	change	n		sci
202		**fundamental**	adj	component	n		sci
203		**fundamental**	adj	difference	n		sci
204		**fundamental**	adj	importance	n		sci
205		**fundamental**	adj	principle	n		sci
206		**fundamental**	adj	problem	n		sci
207		**fundamental**	adj	question	n		sci
208		**genetic**	adj	variation	n		sci
209		**geographic(al)**	adj	area	n		sci
210		**geographic(al)**	adj	distribution	n		sci
211		**geographic(al)**	adj	location	n		sci
212		**global**	adj	capitalism	n		sci
213		**global**	adj	context	n		sci

#	AD	Component I	POS I	Component II	POS II	AD	Sci
214		**global**	adj	culture	n		sci
215		**global**	adj	economy	n		sci
216		**global**	adj	issue	n		sci
217		**global**	adj	market	n		
218		**global**	adj	marketplace	n		
219		**global**	adj	media	n		
220		**global**	adj	network	n		sci
221		**global**	adj	perspective	n		sci
222		**global**	adj	shift	n		
223		**global**	adj	structure	n		sci
224		**global**	adj	trade	n		sci
225		**global**	adj	village	n		
226		**hierarchical**	adj	structure	n		sci
227		**high**	adj	expectations	n		
228		**high**	adj	incidence	n		sci
229		**high**	adj	intensity	n		sci
230		**high**	adj	level	n		sci
231		**high**	adj	order	n		sci
232		**high**	adj	percentage	n		sci
233		**high**	adj	priority	n		sci
234		**high**	adj	probability	n		sci
235		**high**	adj	profile	n		sci
236	(a)	**high**	adj	proportion	n	(of)	sci
237		**high**	adj	quality	n		sci
238		**high**	adj	rate	n		sci
239		**high**	adj	score	n		sci
240		**high**	adj	standard	n		sci
241		**high**	adj	status	n		
242		**high**	adj	turnover	n		sci
243		**high**	adj	unemployment	n		
244		**high**	adj	value	n		sci
245		**individual**	adj	behaviour	n		sci
246		**individual**	adj	case	n		
247		**individual**	adj	characteristics	n		sci
248		**individual**	adj	choice	n		sci
249		**individual**	adj	component	n		sci
250		**individual**	adj	differences	n		sci
251		**individual**	adj	element	n		sci
252		**individual**	adj	experience	n		sci
253		**individual**	adj	interests	n		
254		**individual**	adj	item	n		sci
255		**individual**	adj	needs	n		
256		**individual**	adj	response	n		sci

#	AD	Component I	POS I	Component II	POS II	AD	Sci
257		**individual**	adj	responsibility	n		
258		**individual**	adj	rights	n		
259		**individual**	adj	variable	n		
260		**individual**	adj	variation	n		sci
261		**initial**	adj	period	n		sci
262		**initial**	adj	phase	n		sci
263		**initial**	adj	position	n		sci
264		**initial**	adj	research	n		
265		**initial**	adj	stage	n		sci
266		**integral**	adj	part	n		sci
267		**integrated**	adj	approach	n		sci
268		**integrated**	adj	system	n		sci
269		**internal**	adj	affairs	n		
270		**internal**	adj	conflict	n		
271		**internal**	adj	control	n		sci
272		**internal**	adj	market	n		
273		**internal**	adj	organ	n		sci
274		**internal**	adj	structure	n		sci
275		**intrinsic**	adj	value	n		sci
276		**linear**	adj	relationship	n		sci
277		**longitudinal**	adj	study	n		sci
278		**major**	adj	advantage	n		sci
279		**major**	adj	area	n		
280		**major**	adj	cause	n		sci
281		**major**	adj	challenge	n		sci
282		**major**	adj	change	n		sci
283		**major**	adj	component	n		sci
284		**major**	adj	concern	n		sci
285		**major**	adj	contribution	n		sci
286		**major**	adj	decision	n		
287		**major**	adj	difference	n		sci
288		**major**	adj	factor	n		sci
289		**major**	adj	feature	n		sci
290		**major**	adj	focus	n		sci
291		**major**	adj	impact	n		sci
292		**major**	adj	implications	n		sci
293		**major**	adj	influence	n		sci
294		**major**	adj	issue	n		sci
295		**major**	adj	part	n		sci
296		**major**	adj	problem	n		sci
297		**major**	adj	reason	n		sci
298		**major**	adj	role	n		sci
299		**major**	adj	shift	n		

#	AD	Component I	POS I	Component II	POS II	AD	Sci
300		**major**	adj	source	n		sci
301		**major**	adj	theme	n		
302		**maximum**	adj	duration	n		
303		**media**	n	coverage	n		sci
304		**medical**	adj	assistance	n		
305		**medical**	adj	treatment	n		sci
306		**minimum**	adj	level	n		sci
307		**minimum**	adj	requirement	n		sci
308		**minimum**	adj	standard	n		sci
309		**minimum**	adj	value	n		sci
310		**minimum**	adj	wage	n		
311		**minor**	adj	change	n		sci
312		**minor**	adj	role	n		sci
313		**modified**	adj	form	n		sci
314		**modified**	adj	version	n		sci
315		**negative**	adj	aspect	n		sci
316		**negative**	adj	attitude	n		sci
317		**negative**	adj	connotation	n		sci
318		**negative**	adj	consequences	n		sci
319		**negative**	adj	correlation	n		sci
320		**negative**	adj	effect	n		sci
321		**negative**	adj	feedback	n		sci
322		**negative**	adj	impact	n		sci
323		**negative**	adj	outcome	n		sci
324		**negative**	adj	side	n		sci
325		**negative**	adj	stereotype	n		
326		**negative**	adj	value	n		sci
327		**negative**	adj	view	n		
328		**normal**	adj	conditions	n		sci
329		**normal**	adj	development	n		sci
330		**normal**	adj	distribution	n		sci
331		**normal**	adj	practice	n		
332		**nuclear**	adj	energy	n		sci
333		**nuclear**	adj	family	n		
334		**nuclear**	adj	power	n		sci
335		**nuclear**	adj	war	n		
336		**nuclear**	adj	weapon	n		sci
337		**objective**	adj	criteria	n		
338		**objective**	adj	reality	n		
339		**obvious**	adj	difference	n		sci
340		**obvious**	adj	example	n		sci
341		**obvious**	adj	point	n		
342		**obvious**	adj	reason	n		sci

#	AD	Component I	POS I	Component II	POS II	AD	Sci
343		**occur**	v	frequently	adv		sci
344		**occur**	v	naturally	adv		sci
345		**online**	adj	access	n		sci
346		**online**	adj	database	n		sci
347		**online**	adj	journal	n		
348		**online**	adj	version	n		sci
349		**optimal**	adj	solution	n		sci
350		**oral**	adj	history	n		
351		**oral**	adj	presentation	n		
352		**overall**	adj	aim	n		sci
353		**overall**	adj	effect	n		sci
354		**overall**	adj	level	n		sci
355		**overall**	adj	performance	n		sci
356		**overall**	adj	picture	n		sci
357		**overall**	adj	rate	n		sci
358		**overall**	adj	structure	n		sci
359		**physical**	adj	activity	n		sci
360		**physical**	adj	appearance	n		sci
361		**physical**	adj	characteristics	n		sci
362		**physical**	adj	contact	n		sci
363		**physical**	adj	development	n		
364		**physical**	adj	environment	n		sci
365		**physical**	adj	features	n		
366		**physical**	adj	health	n		
367		**physical**	adj	needs	n		
368		**physical**	adj	presence	n		sci
369		**physical**	adj	properties	n		sci
370		**physical**	adj	proximity	n		
371		**physical**	adj	science	n		sci
372		**physical**	adj	space	n		sci
373		**physical**	adj	symptom	n		
374		**physical**	adj	world	n		sci
375		**pose (a)**	v	challenge	n		
376		**pose (a)**	v	problem	n		sci
377		**pose (a)**	v	question	n		sci
378		**pose (a)**	v	threat	n	(to)	sci
379		**positive**	adj	action	n		sci
380		**positive**	adj	aspect	n		
381		**positive**	adj	attitude	n		sci
382		**positive**	adj	connotation	n		
383		**positive**	adj	correlation	n		sci
384		**positive**	adj	discrimination	n		
385		**positive**	adj	effect	n		sci

#	AD	Component I	POS I	Component II	POS II	AD	Sci
386		**positive**	adj	feature	n		
387		**positive**	adj	feedback	n		sci
388		**positive**	adj	image	n		
389		**positive**	adj	impact	n		sci
390		**positive**	adj	influence	n		sci
391		**positive**	adj	outcome	n		
392		**positive**	adj	relationship	n		sci
393		**positive**	adj	result	n		sci
394		**positive**	adj	value	n		sci
395		**positive**	adj	view	n		sci
396		**potential**	adj	benefits	n		sci
397		**potential**	adj	conflict	n		sci
398		**potential**	adj	customer	n		
399		**potential**	adj	harm	n		sci
400		**potential**	adj	impact	n		sci
401		**potential**	adj	problem	n		sci
402		**potential**	adj	risk	n		sci
403		**potential**	adj	source	n		sci
404		**potential**	adj	value	n		sci
405		**preceding**	adj	chapter	n		
406		**preceding**	adj	discussion	n		sci
407		**preceding**	adj	section	n		sci
408		**preliminary**	adj	data	n		sci
409		**preliminary**	adj	findings	n		sci
410		**preliminary**	adj	result	n		sci
411		**previous**	adj	chapter	n		
412		**previous**	adj	decade	n		sci
413		**previous**	adj	discussion	n		sci
414		**previous**	adj	experience	n		
415		**previous**	adj	generation	n		
416		**previous**	adj	knowledge	n		
417		**previous**	adj	paragraph	n		sci
418		**previous**	adj	part	n		
419		**previous**	adj	research	n		sci
420		**previous**	adj	section	n		sci
421		**previous**	adj	study	n		sci
422		**previous**	adj	work	n		sci
423		**primary**	adj	aim	n		
424		**primary**	adj	care	n		
425		**primary**	adj	concern	n		sci
426		**primary**	adj	data	n		sci
427		**primary**	adj	education	n		sci
428		**primary**	adj	focus	n		sci

#	AD	Component I	POS I	Component II	POS II	AD	Sci
429		**primary**	adj	function	n		sci
430		**primary**	adj	objective	n		sci
431		**primary**	adj	purpose	n		sci
432		**primary**	adj	reason	n		sci
433		**primary**	adj	research	n		sci
434		**primary**	adj	responsibility	n		
435		**primary**	adj	source	n		sci
436		**primary**	adj	task	n		
437		**prime**	adj	example	n		sci
438		**prime**	adj	time	n		sci
439		**principal**	adj	source	n		
440		**prior**	adj	experience	n		sci
441		**prior**	adj	knowledge	n		sci
442		**process**	v	data	n		sci
443		**process**	v	information	n		
444		**promote**	v	equality	n		
445		**published**	adj	literature	n		sci
446		**published**	adj	material	n		
447		**published**	adj	research	n		sci
448		**published**	adj	work	n		sci
449		**radical**	adj	change	n		sci
450		**radical**	adj	critique	n		
451		**radical**	adj	differences	n		
452		**radical**	adj	transformation	n		
453		**random**	adj	error	n		sci
454		**random**	adj	sample	n		sci
455		**random**	adj	variable	n		sci
456		**relevant**	adj	data	n		sci
457		**relevant**	adj	factors	n		sci
458		**relevant**	adj	information	n		sci
459		**relevant**	adj	issue	n		sci
460		**relevant**	adj	literature	n		sci
461		**relevant**	adj	material	n		sci
462		**reliable**	adj	data	n		sci
463		**reliable**	adj	information	n		sci
464		**research**	n	effort	n		sci
465		**research**	n	evidence	n		
466		**research**	n	findings	n		sci
467		**research**	n	methodology	n		
468	(for)	**research**	n	purposes	n		sci
469		**research**	n	topic	n		sci
470		**seek**	v	help	n		
471		**seek**	v	information	n		

#	AD	Component I	POS I	Component II	POS II	AD	Sci
472	(a)	**significant**	adj	amount	n	(of)	sci
473		**significant**	adj	change	n		sci
474		**significant**	adj	contribution	n		sci
475		**significant**	adj	correlation	n		sci
476	(a)	**significant**	adj	degree	n	(of)	sci
477		**significant**	adj	development	n		sci
478		**significant**	adj	difference	n		sci
479		**significant**	adj	effect	n		sci
480		**significant**	adj	factor	n		sci
481		**significant**	adj	feature	n		sci
482		**significant**	adj	figures	n		sci
483		**significant**	adj	growth	n		sci
484		**significant**	adj	impact	n		sci
485		**significant**	adj	improvement	n		sci
486		**significant**	adj	increase	n		sci
487		**significant**	adj	influence	n		sci
488		**significant**	adj	interaction	n		sci
489		**significant**	adj	number	n		sci
490		**significant**	adj	part	n		sci
491		**significant**	adj	portion	n		sci
492	(a)	**significant**	adj	proportion	n	(of)	sci
493		**significant**	adj	reduction	n		sci
494		**significant**	adj	relationship	n		sci
495		**significant**	adj	role	n		sci
496		**significant**	adj	shift	n		sci
497		**significant**	adj	variation	n		sci
498		**similar**	adj	approach	n		sci
499		**similar**	adj	argument	n		sci
500		**similar**	adj	characteristics	n		sci
501		**similar**	adj	effect	n		sci
502		**similar**	adj	issue	n		
503		**similar**	adj	pattern	n		sci
504		**similar**	adj	properties	n		sci
505		**similar**	adj	result	n		sci
506		**similar**	adj	situation	n		sci
507		**source**	n	material	n		sci
508		**specific**	adj	area	n		sci
509		**specific**	adj	aspect	n		sci
510		**specific**	adj	case	n		sci
511		**specific**	adj	characteristic	n		sci
512		**specific**	adj	context	n		
513		**specific**	adj	example	n		sci
514		**specific**	adj	factor	n		sci

#	AD	Component I	POS I	Component II	POS II	AD	Sci
515		**specific**	adj	feature	n		sci
516		**specific**	adj	focus	n		sci
517		**specific**	adj	form	n		sci
518		**specific**	adj	function	n		sci
519		**specific**	adj	information	n		sci
520		**specific**	adj	issue	n		
521		**specific**	adj	knowledge	n		sci
522		**specific**	adj	meaning	n		sci
523		**specific**	adj	needs	n		sci
524		**specific**	adj	problem	n		sci
525		**specific**	adj	purpose	n		
526		**specific**	adj	question	n		sci
527		**specific**	adj	reference	n		sci
528		**specific**	adj	sense	n		
529		**specific**	adj	type	n		sci
530		**statistical**	adj	analysis	n		sci
531		**statistical**	adj	data	n		sci
532		**statistical**	adj	information	n		sci
533		**statistical**	adj	method	n		sci
534		**statistical**	adj	significance	n		sci
535		**statistical**	adj	technique	n		sci
536		**statistical**	adj	test	n		sci
537		**stress**	n	level	n		sci
538		**subsequent**	adj	analysis	n		sci
539		**subsequent**	adj	chapter	n		
540		**subsequent**	adj	development	n		sci
541		**subsequent**	adj	study	n		sci
542		**subsequent**	adj	work	n		sci
543		**sufficient**	adj	condition	n		sci
544		**sufficient**	adj	detail	n		sci
545		**sufficient**	adj	evidence	n		sci
546		**sufficient**	adj	information	n		
547		**sufficient**	adj	resources	n		sci
548		**superior**	adj	performance	n		sci
549		**survey**	n	data	n		sci
550		**target**	n	audience	n		
551		**technical**	adj	aspect	n		sci
552		**technical**	adj	assistance	n		sci
553		**technical**	adj	detail	n		sci
554		**technical**	adj	expertise	n		sci
555		**technical**	adj	issue	n		
556		**technical**	adj	knowledge	n		
557		**technical**	adj	problem	n		sci

#	AD	Component I	POS I	Component II	POS II	AD	Sci
558		**technical**	adj	skill	n		
559		**technical**	adj	support	n		
560		**technical**	adj	term	n		sci
561		**traditional**	adj	approach	n		sci
562		**traditional**	adj	culture	n		sci
563		**traditional**	adj	form	n		
564		**traditional**	adj	method	n		sci
565		**traditional**	adj	practice	n		sci
566		**traditional**	adj	research	n		
567		**traditional**	adj	society	n		
568		**traditional**	adj	value	n		sci
569		**traditional**	adj	view	n		sci
570		**transport**	n	system	n		sci
571		**undergo**	adj	transformation	n		sci
572		**underlying**	adj	assumption	n		sci
573		**underlying**	adj	cause	n		sci
574		**underlying**	adj	principle	n		sci
575		**underlying**	adj	process	n		sci
576		**underlying**	adj	reason	n		sci
577		**underlying**	adj	structure	n		sci
578		**unique**	adj	individual	n		
579		**unique**	adj	opportunity	n		sci
580		**unique**	adj	position	n		
581		**virtually**	adv	impossible	adj		sci
582		**widespread**	adj	acceptance	n		
583		**widespread**	adj	belief	n		
584		**widespread**	adj	support	n		
585		**widespread**	adj	use	n		sci

2.2 Component II (?-) *Modified from Ackermann and Chen (2013)*

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
1		abstract	adj	**concept**	n		
2		allow	v	**access**	n		sci
3		deny	v	**access**	n		
4		direct	adj	**access**	n		sci
5		easy	adj	**access**	n		sci
6		electronic	adj	**access**	n		
7		equal	adj	**access**	n		sci
8		free	adj	**access**	n		sci
9		gain	v	**access**	n		sci
10		give	v	**access**	n		sci
11		have	v	**access**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
12		internet	n	**access**	n		sci
13		limited	adj	**access**	n		sci
14		online	adj	**access**	n		sci
15		open	adj	**access**	n		sci
16		provide	v	**access**	n		sci
17		public	adj	**access**	n		sci
18		ready	adj	**access**	n		
19		unlimited	adj	**access**	n		
20		great	adj	**accuracy**	n		
21		particularly	adv	**acute**	adj		sci
22		public	adj	**administration**	n		
23		adversely	adv	**affect**	v		sci
24		directly	adv	**affect**	v		sci
25		severely	adv	**affect**	v		sci
26		significantly	adv	**affect**	v		sci
27		provide (an)	v	**alternative**	n		sci
28		careful	adj	**analysis**	n		sci
29		comparative	adj	**analysis**	n		sci
30		conduct (an)	v	**analysis**	n		sci
31		critical	adj	**analysis**	n		sci
32		detailed	adj	**analysis**	n		sci
33		economic	adj	**analysis**	n		sci
34		final	adj	**analysis**	n		sci
35		full	adj	**analysis**	n		
36		further	adj	**analysis**	n		sci
37		historical	adj	**analysis**	n		
38		qualitative	adj	**analysis**	n		sci
39		quantitative	adj	**analysis**	n		sci
40		statistical	adj	**analysis**	n		sci
41		subsequent	adj	**analysis**	n		sci
42		systematic	adj	**analysis**	n		sci
43		textual	adj	**analysis**	n		
44		thematic	adj	**analysis**	n		
45		theoretical	adj	**analysis**	n		sci
46		use (the)	v	**analysis**	n		
47		become	v	**apparent**	adj		sci
48		immediately	adv	**apparent**	adj		sci
49		particularly	adv	**apparent**	adj		sci
50		adopt (an)	v	**approach**	n		sci
51		alternative	adj	**approach**	n		sci
52		analytical	adj	**approach**	n		sci
53		common	adj	**approach**	n		sci
54		comprehensive	adj	**approach**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
55		critical	adj	**approach**	n		
56		develop (an)	v	**approach**	n		sci
57		flexible	adj	**approach**	n		sci
58		general	adj	**approach**	n		sci
59		holistic	adj	**approach**	n		sci
60		integrated	adj	**approach**	n		sci
61		logical	adj	**approach**	n		
62		methodological	adj	**approach**	n		
63		qualitative	adj	**approach**	n		sci
64		quantitative	adj	**approach**	n		sci
65		similar	adj	**approach**	n		sci
66		standard	adj	**approach**	n		sci
67		systematic	adj	**approach**	n		sci
68		take (an)	v	**approach**	n		
69		theoretical	adj	**approach**	n		sci
70		traditional	adj	**approach**	n		sci
71		use (an)	v	**approach**	n		sci
72		consider	v	**appropriate**	adj		sci
73		deem	v	**appropriate**	adj		
74		particularly	adv	**appropriate**	adj		sci
75		seem	v	**appropriate**	adj		sci
76		complex	adj	**area**	n		sci
77		core	adj	**area**	n		sci
78		cover (an)	v	**area**	n		sci
79		geographic(al)	adj	**area**	n		sci
80		identify (an)	v	**area**	n		
81		key	adj	**area**	n		sci
82		local	adj	**area**	n		sci
83		main	adj	**area**	n		sci
84		major	adj	**area**	n		sci
85		metropolitan	adj	**area**	n		sci
86		particular	adj	**area**	n		sci
87		problem	n	**area**	n		sci
88		related	adj	**area**	n		sci
89		rural	adj	**area**	n		sci
90		specific	adj	**area**	n		sci
91		subject	adj	**area**	n		sci
92		urban	adj	**area**	n		sci
93		vast	adj	**area**	n		
94		whole	adj	**area**	n		sci
95		wide	adj	**area**	n		sci
96	(a)	vast	adj	**array**	n		sci
97	(a)	wide	adj	**array**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
98		various	adj	**aspects**	n		sci
99	(be)	randomly	adv	**assigned**	vpp		sci
100		change (an)	v	**attitude**	n		
101		negative	adj	**attitude**	n		sci
102		positive	adj	**attitude**	n		sci
103		become	v	**available**	adj		sci
104		currently	adv	**available**	adj		sci
105		freely	adv	**available**	adj		sci
106		make	v	**available**	adj		sci
107		publicly	adv	**available**	adj		sci
108		readily	adv	**available**	adj		sci
109		widely	adv	**available**	adj		sci
110		acutely	adv	**aware**	adj		sci
111		become	v	**aware**	adj		sci
112		fully	adv	**aware**	adj		sci
113		increasingly	adv	**aware**	adj		sci
114		keenly	adv	**aware**	adj		
115		make	v	**aware**	adj		sci
116		well	adv	**aware**	adj		sci
117		limited	adj	**capacity**	n		
118		face (a)	v	**challenge**	n		
119		major	adj	**challenge**	n		sci
120		pose (a)	v	**challenge**	n		
121		present (a)	v	**challenge**	n		sci
122		serious	adj	**challenge**	n		sci
123		current	adj	**climate**	n		
124		political	adj	**climate**	n		
125		direct	adj	**communication**	n		
126		effective	adj	**communication**	n		sci
127		electronic	adj	**communication**	n		sci
128		personal	adj	**communication**	n		sci
129		verbal	adj	**communication**	n		
130		written	adj	**communication**	n		sci
131		academic	adj	**community**	n		sci
132		ethnic	adj	**community**	n		
133		international	adj	**community**	n		sci
134		local	adj	**community**	n		sci
135		rural	adj	**community**	n		
136		scientific	adj	**community**	n		sci
137		virtual	adj	**community**	n		
138		wider	adj	**community**	n		
139		extremely	adv	**complex**	adj		sci
140		highly	adv	**complex**	adj		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
141		increasingly	adv	**complex**	adj		sci
142		abstract	adj	**concept**	n		sci
143		basic	adj	**concept**	n		sci
144		central	adj	**concept**	n		sci
145		defining	adj	**concept**	n		sci
146		key	adj	**concept**	n		sci
147		theoretical	adj	**concept**	n		sci
148		use (the)	v	**concept**	n		sci
149		draw (a)	v	**conclusion**	n		sci
150		general	adj	**conclusion**	n		sci
151		lead to (the)	v	**conclusion**	n		sci
152		logical	adj	**conclusion**	n		sci
153		annual	adj	**conference**	n		sci
154		attend (a)	v	**conference**	n		
155		hold (a)	v	**conference**	n		
156		international	adj	**conference**	n		sci
157		national	adj	**conference**	n		sci
158		relatively	adv	**constant**	adj		sci
159		remain	v	**constant**	adj		sci
160		impose	v	**constraints**	n		sci
161		social	adj	**construct**	n		
162		close	adj	**contact**	n		sci
163		come into	v	**contact**	n		sci
164		direct	adj	**contact**	n		sci
165		first	adj	**contact**	n		sci
166		maintain	v	**contact**	n		
167		make	v	**contact**	n		sci
168		personal	adj	**contact**	n		
169		physical	adj	**contact**	n		sci
170		sexual	adj	**contact**	n		sci
171		social	adj	**contact**	n		sci
172		broader	adj	**context**	n		sci
173		cultural	adj	**context**	n		
174		economic	adj	**context**	n		
175		global	adj	**context**	n		
176		historical	adj	**context**	n		sci
177		institutional	adj	**context**	n		
178		international	adj	**context**	n		sci
179		original	adj	**context**	n		
180		political	adj	**context**	n		
181		present	adj	**context**	n		sci
182		provide	v	**context**	n		sci
183		social	adj	**context**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
184		specific	adj	**context**	n		
185		wider	adj	**context**	n		sci
186		marked	adj	**contrast**	n		sci
187		sharp	adj	**contrast**	n		sci
188		stark	adj	**contrast**	n		sci
189		striking	adj	**contrast**	n		sci
190		central	adj	**core**	n		sci
191		high	adj	**correlation**	n		sci
192		negative	adj	**correlation**	n		sci
193		positive	adj	**correlation**	n		sci
194		significant	adj	**correlation**	n		sci
195		strong	adj	**correlation**	n		sci
196		newly	adv	**created**	vpp		
197		meet	v	**criteria**	n		sci
198		objective	adj	**criteria**	n		
199		use	v	**criteria**	n		sci
200		common	adj	**culture**	n		
201		dominant	adj	**culture**	n		
202		global	adj	**culture**	n		
203		local	adj	**culture**	n		
204		modern	adj	**culture**	n		
205		national	adj	**culture**	n		
206		political	adj	**culture**	n		
207		popular	adj	**culture**	n		
208		traditional	adj	**culture**	n		sci
209		appropriate	adj	**data**	n		sci
210		available	adj	**data**	n		sci
211		collect	v	**data**	n		sci
212		empirical	adj	**data**	n		sci
213		existing	adj	**data**	n		sci
214		experimental	adj	**data**	n		sci
215		extract	v	**data**	n		sci
216		gather	v	**data**	n		sci
217		historical	adj	**data**	n		sci
218		interpret	v	**data**	n		sci
219		missing	adj	**data**	n		sci
220		numerical	adj	**data**	n		sci
221		obtain	v	**data**	n		sci
222		original	adj	**data**	n		sci
223		preliminary	adj	**data**	n		sci
224		present	v	**data**	n		sci
225		primary	adj	**data**	n		sci
226		process	v	**data**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
227		provide	v	**data**	n		sci
228		qualitative	adj	**data**	n		sci
229		quantitative	adj	**data**	n		sci
230		raw	adj	**data**	n		sci
231		record	v	**data**	n		sci
232		relevant	adj	**data**	n		sci
233		reliable	adj	**data**	n		sci
234		report	v	**data**	n		sci
235		secondary	adj	**data**	n		
236		statistical	adj	**data**	n		sci
237		store	v	**data**	n		sci
238		survey	n	**data**	n		sci
239		transmit	v	**data**	n		sci
240		use (the)	v	**data**	n		sci
241		online	adj	**database**	n		sci
242		early	adj	**decades**	n		
243		recent	adj	**decades**	n		sci
244	(be)	broadly	adv	**defined**	vpp		sci
245	(be)	clearly	adv	**defined**	vpp		sci
246		environmental	adj	**degradation**	n		sci
247	(be)	clearly	adv	**demonstrated**	vpp		sci
248		experimental	adj	**design**	n		sci
249		urban	adj	**design**	n		
250		face	v	**discrimination**	n		
251		positive	adj	**discrimination**	n		
252		racial	adj	**discrimination**	n		
253		public	adj	**display**	n		
254		geographic(al)	adj	**distribution**	n		sci
255		normal	adj	**distribution**	n		sci
256		cultural	adj	**diversity**	n		
257		ethnic	adj	**diversity**	n		
258		great	adj	**diversity**	n		sci
259	(be)	well	adv	**documented**	vpp		sci
260		public	adj	**domain**	n		sci
261		change	v	**dramatically**	adv		sci
262		increase	v	**dramatically**	adv		sci
263		long	adj	**duration**	n		sci
264		maximum	adj	**duration**	n		
265		short	adj	**duration**	n		sci
266		constituent	adj	**elements**	n		sci
267	(be)	deeply	adv	**embedded**	vpp		sci
268		give	v	**emphasis**	n		
269		greater	adj	**emphasis**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
270		increasing	adj	**emphasis**	n		
271		particular	adj	**emphasis**	n		sci
272		place	v	**emphasis**	n		sci
273		shift	v	**emphasis**	n		
274		special	adj	**emphasis**	n		sci
275		strong	adj	**emphasis**	n		
276	(be)	commonly	adv	**encountered**	vpp		sci
277		atomic	adj	**energy**	n		sci
278		nuclear	adj	**energy**	n		sci
279		renewable	adj	**energy**	n		sci
280		solar	adj	**energy**	n		sci
281		greatly	adv	**enhance**	v		sci
282		roughly	adj	**equivalent**	n		
283		common	adj	**error**	n		
284		correct (an)	v	**error**	n		
285		random	adj	**error**	n		sci
286		standard	adj	**error**	n		sci
287		become	v	**established**	adj		sci
288	(be)	clearly	adv	**established**	vpp		sci
289	(be)	firmly	adv	**established**	vpp		sci
290		long	adv	**established**	adj		sci
291		newly	adv	**established**	adj		sci
292		once	adv	**established**	vpp		sci
293	(be)	well	adv	**established**	vpp		sci
294		ample	adj	**evidence**	n		sci
295		anecdotal	adj	**evidence**	n		sci
296		available	adj	**evidence**	n		sci
297		clear	adj	**evidence**	n		sci
298		compelling	adj	**evidence**	n		sci
299		considerable	adj	**evidence**	n		sci
300		convincing	adj	**evidence**	n		sci
301		direct	adj	**evidence**	n		sci
302		documentary	adj	**evidence**	n		
303		empirical	adj	**evidence**	n		sci
304		experimental	adj	**evidence**	n		sci
305		find	v	**evidence**	n		sci
306		further	adj	**evidence**	n		sci
307		give	v	**evidence**	n		sci
308		historical	adj	**evidence**	n		sci
309		little	adj	**evidence**	n		sci
310		present	v	**evidence**	n		sci
311		provide	v	**evidence**	n		sci
312		recent	adj	**evidence**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
313		research	n	**evidence**	n		
314		scientific	adj	**evidence**	n		sci
315		show	v	**evidence**	n		sci
316		strong	adj	**evidence**	n		sci
317		substantial	adj	**evidence**	n		sci
318		sufficient	adj	**evidence**	n		sci
319		supporting	adj	**evidence**	n		sci
320		biological	adj	**evolution**	n		sci
321		make	v	**explicit**	adj		sci
322		associated	adj	**factors**	n		sci
323		contextual	adj	**factors**	n		sci
324		cultural	adj	**factors**	n		
325		economic	adj	**factors**	n		sci
326		environmental	adj	**factors**	n		sci
327		external	adj	**factors**	n		sci
328		historical	adj	**factors**	n		sci
329		identify	v	**factors**	n		sci
330		political	adj	**factors**	n		
331		relevant	adj	**factors**	n		sci
332		social	adj	**factors**	n		sci
333		identify	v	**features**	n		sci
334		physical	adj	**features**	n		
335		give	v	**feedback**	n		sci
336		negative	adj	**feedback**	n		sci
337		positive	adj	**feedback**	n		sci
338		provide	v	**feedback**	n		sci
339		receive	v	**feedback**	n		sci
340		become (the)	v	**focus**	n		sci
341		central	adj	**focus**	n		
342		clear	adj	**focus**	n		
343		main	adj	**focus**	n		sci
344		major	adj	**focus**	n		sci
345		particular	adj	**focus**	n		sci
346		primary	adj	**focus**	n		sci
347		provide (a)	v	**focus**	n		sci
348		specific	adj	**focus**	n		sci
349		general	adj	**formula**	n		sci
350		small	adj	**fraction**	n		sci
351		conceptual	adj	**framework**	n		sci
352		institutional	adj	**framework**	n		
353		legal	adj	**framework**	n		
354		regulatory	adj	**framework**	n		sci
355		theoretical	adj	**framework**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
356		basic	adj	**function**	n		sci
357		essential	adj	**function**	n		sci
358		main	adj	**function**	n		sci
359		perform (a)	v	**function**	n		
360		primary	adj	**function**	n		sci
361		serve (a)	v	**function**	n		
362		social	adj	**function**	n		sci
363		specific	adj	**function**	n		sci
364		first	adj	**generation**	n		sci
365		next	adj	**generation**	n		sci
366		previous	adj	**generation**	n		
367		younger	adj	**generation**	n		
368		achieve (a)	v	**goal**	n		
369		common	adj	**goal**	n		
370		economic	adj	**goal**	n		
371		set (a)	v	**goal**	n		
372		ultimate	adj	**goal**	n		sci
373		almost	adv	**identical**	adj		sci
374	(be)	clearly	adv	**identified**	vpp		sci
375	(be)	easily	adv	**identified**	vpp		sci
376		positive	adj	**image**	n		
377		public	adj	**image**	n		
378		visual	adj	**image**	n		
379		assess (the)	v	**impact**	n		sci
380		consider (the)	v	**impact**	n		sci
381		direct	adj	**impact**	n		sci
382		emotional	adj	**impact**	n		
383		enormous	adj	**impact**	n		sci
384		environmental	adj	**impact**	n		sci
385		great	adj	**impact**	n		sci
386		likely	adj	**impact**	n		sci
387		little	adj	**impact**	n		sci
388		major	adj	**impact**	n		sci
389		make (an)	v	**impact**	n		
390		negative	adj	**impact**	n		sci
391		positive	adj	**impact**	n		sci
392		potential	adj	**impact**	n		sci
393		profound	adj	**impact**	n		sci
394		significant	adj	**impact**	n		sci
395		consider (the)	v	**implications**	n		sci
396		major	adj	**implications**	n		sci
397		political	adj	**implications**	n		sci
398		social	adj	**implications**	n		

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
399		wider	adj	**implications**	n		
400		high	adj	**incidence**	n		sci
401		particular	adj	**individual**	n		
402		private	adj	**individual**	n		
403		single	adj	**individual**	n		sci
404		unique	adj	**individual**	n		
405		technological	adj	**innovation**	n		
406		follow	v	**instructions**	n		
407	(be)	fully	adv	**integrated**	vpp		sci
408		high	adj	**intensity**	n		sci
409		low	adj	**intensity**	n		sci
410		alternative	adj	**interpretation**	n		sci
411		correct	adj	**interpretation**	n		sci
412		historical	adj	**interpretation**	n		
413		literal	adj	**interpretation**	n		
414		effective	adj	**intervention**	n		sci
415		government	n	**intervention**	n		
416	(be)	actively	adv	**involved**	vpp		sci
417		become	v	**involved**	adj		
418	(be)	directly	adv	**involved**	vpp		sci
419		get	v	**involved**	adj		sci
420	(be)	closely	adv	**linked**	vpp		sci
421	(be)	directly	adv	**linked**	vpp		sci
422	(be)	inextricably	adv	**linked**	vpp		sci
423	(be)	strongly	adv	**linked**	vpp		sci
424		geographic(al)	adj	**location**	n		sci
425		digital	adj	**media**	n		
426		electronic	adj	**media**	n		
427		global	adj	**media**	n		
428		national	adj	**media**	n		
429		popular	adj	**media**	n		
430		visual	adj	**media**	n		
431		alternative	adj	**method**	n		sci
432		apply (a)	v	**method**	n		sci
433		common	adj	**method**	n		sci
434		describe (a)	v	**method**	n		sci
435		develop (a)	v	**method**	n		sci
436		effective	adj	**method**	n		sci
437		employ (a)	v	**method**	n		sci
438		experimental	adj	**method**	n		sci
439		modern	adj	**method**	n		sci
440		qualitative	adj	**method**	n		
441		quantitative	adj	**method**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
442		scientific	adj	**method**	n		sci
443		standard	adj	**method**	n		sci
444		statistical	adj	**method**	n		sci
445		traditional	adj	**method**	n		sci
446		use (a)	v	**method**	n		sci
447		relatively	adv	**minor**	adj		sci
448	(be)	politically	adv	**motivated**	vpp		sci
449		global	adj	**network**	n		sci
450		cultural	adj	**norm**	n		
451		social	adj	**norm**	n		
452		achieve (an)	v	**objective**	n		sci
453		key	adj	**objective**	n		sci
454		learning	n	**objective**	n		
455		meet (an)	v	**objective**	n		
456		primary	adj	**objective**	n		sci
457		set (an)	v	**objective**	n		
458		strategic	adj	**objective**	n		
459		become	v	**obvious**	adj		sci
460		fairly	adv	**obvious**	adj		
461		immediately	adv	**obvious**	adj		sci
462		seem	v	**obvious**	adj		sci
463		sexual	adj	**orientation**	n		
464		brief	adj	**overview**	n		sci
465		comprehensive	adj	**overview**	n		sci
466		general	adj	**overview**	n		sci
467		give (an)	v	**overview**	n		sci
468		provide (an)	v	**overview**	n		sci
469		solar	adj	**panel**	n		
470		set (the)	v	**parameters**	n		
471		reach (a)	v	**peak**	n		sci
472		high	adj	**percentage**	n		sci
473		large	adj	**percentage**	n		sci
474		low	adj	**percentage**	n		sci
475		small	adj	**percentage**	n		sci
476		public	adj	**perception**	n		sci
477		visual	adj	**perception**	n		sci
478		brief	adj	**period**	n		sci
479		earlier	adj	**period**	n		
480		entire	adj	**period**	n		sci
481		extended	adj	**period**	n		sci
482		given	adj	**period**	n		sci
483		historical	adj	**period**	n		
484		initial	adj	**period**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
485		short	adj	**period**	n		sci
486		whole	adj	**period**	n		sci
487		critical	adj	**perspective**	n		
488		cultural	adj	**perspective**	n		sci
489		global	adj	**perspective**	n		sci
490		historical	adj	**perspective**	n		sci
491		new	adj	**perspective**	n		sci
492		theoretical	adj	**perspective**	n		sci
493		final	adj	**phase**	n		sci
494		first	adj	**phase**	n		sci
495		initial	adj	**phase**	n		sci
496		next	adj	**phase**	n		sci
497		cultural	adj	**phenomenon**	n		
498		social	adj	**phenomenon**	n		
499		large	adj	**portion**	n		sci
500		significant	adj	**portion**	n		sci
501		small	adj	**portion**	n		sci
502		central	adj	**position**	n		
503		dominant	adj	**position**	n		sci
504		final	adj	**position**	n		sci
505		initial	adj	**position**	n		sci
506		legal	adj	**position**	n		
507		original	adj	**position**	n		sci
508		privileged	adj	**position**	n		
509		unique	adj	**position**	n		
510		full	adj	**potential**	n		sci
511		great	adj	**potential**	n		sci
512		have	v	**potential**	n		sci
513		immediately	adv	**preceding**	adj		
514		basic	adj	**principle**	n		sci
515		established	adj	**principle**	n		sci
516		ethical	adj	**principle**	n		
517		fundamental	adj	**principle**	n		
518		general	adj	**principle**	n		sci
519		guiding	adj	**principle**	n		
520		key	adj	**principle**	n		
521		main	adj	**principle**	n		
522		moral	adj	**principle**	n		
523		organising	adj	**principle**	n		
524		underlying	adj	**principle**	n		sci
525		adopt (a)	v	**procedure**	n		sci
526		describe (a)	v	**procedure**	n		sci
527		follow (a)	v	**procedure**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
528		use (a)	v	**procedure**	n		sci
529		assessment	n	**process**	n		sci
530		begin (a)	v	**process**	n		sci
531		complex	adj	**process**	n		sci
532		continuous	adj	**process**	n		sci
533		creative	adj	**process**	n		
534		democratic	adj	**process**	n		sci
535		describe (a)	v	**process**	n		sci
536		developmental	adj	**process**	n		sci
537		due	adj	**process**	n		
538		dynamic	adj	**process**	n		sci
539		evolutionary	adj	**process**	n		sci
540		learning	n	**process**	n		sci
541		natural	adj	**process**	n		sci
542		ongoing	adj	**process**	n		sci
543		slow	adj	**process**	n		sci
544		start (a)	v	**process**	n		sci
545		thinking	n	**process**	n		sci
546		thought	n	**process**	n		sci
547		underlying	adj	**process**	n		sci
548		high	adj	**profile**	n		sci
549		low	adj	**profile**	n		sci
550	(a)	great	adj	**proportion**	n		sci
551	(a)	high	adj	**proportion**	n		sci
552	(a)	increasing	adj	**proportion**	n		sci
553	(a)	large	adj	**proportion**	n		sci
554	(a)	significant	adj	**proportion**	n		sci
555	(a)	small	adj	**proportion**	n		sci
556	(a)	broad	adj	**range**	n		sci
557		cover (a)	v	**range**	n		sci
558	(a)	diverse	adj	**range**	n		sci
559	(the)	entire	adj	**range**	n		sci
560	(a, the)	full	adj	**range**	n		sci
561	(a)	great	adj	**range**	n		sci
562	(a)	large	adj	**range**	n		sci
563	(a)	limited	adj	**range**	n		sci
564	(a)	narrow	adj	**range**	n		sci
565	(a)	vast	adj	**range**	n		sci
566	(a)	whole	adj	**range**	n		sci
567	(a)	wide	adj	**range**	n		sci
568		adverse	adj	**reaction**	n		
569		chemical	adj	**reaction**	n		sci
570		emotional	adj	**reaction**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
571		strong	adj	**reaction**	n		sci
572		specific	adj	**reference**	n		sci
573		consider	v	**relevant**	adj		sci
574		highly	adv	**relevant**	adj		sci
575		particularly	adv	**relevant**	adj		sci
576	(be)	far	adv	**removed**	vpp		sci
577		academic	adj	**research**	n		sci
578		basic	adj	**research**	n		sci
579		carry out	v	**research**	n		
580		comparative	adj	**research**	n		
581		conduct	v	**research**	n		sci
582		considerable	adj	**research**	n		sci
583		current	adj	**research**	n		sci
584		earlier	adj	**research**	n		sci
585		educational	adj	**research**	n		sci
586		empirical	adj	**research**	n		sci
587		existing	adj	**research**	n		sci
588		experimental	adj	**research**	n		sci
589		extensive	adj	**research**	n		sci
590		field	n	**research**	n		sci
591		further	adj	**research**	n		sci
592		future	adj	**research**	n		sci
593		initial	adj	**research**	n		
594		little	adj	**research**	n		
595		original	adj	**research**	n		sci
596		past	adj	**research**	n		sci
597		previous	adj	**research**	n		sci
598		primary	adj	**research**	n		sci
599		publish	v	**research**	n		sci
600		published	adj	**research**	n		sci
601		qualitative	adj	**research**	n		sci
602		quantitative	adj	**research**	n		sci
603		recent	adj	**research**	n		sci
604		scholarly	adj	**research**	n		sci
605		scientific	adj	**research**	n		sci
606		traditional	adj	**research**	n		
607		undertake	v	**research**	n		
608		conflict	n	**resolution**	n		
609		additional	adj	**resources**	n		
610		allocate	v	**resources**	n		sci
611		available	adj	**resources**	n		sci
612		economic	adj	**resources**	n		sci
613		electronic	adj	**resources**	n		

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
614		financial	adj	**resources**	n		sci
615		learning	n	**resources**	n		
616		limited	adj	**resources**	n		sci
617		natural	adj	**resources**	n		sci
618		provide	v	**resources**	n		sci
619		require	v	**resources**	n		sci
620		scarce	adj	**resources**	n		sci
621		sufficient	adj	**resources**	n		sci
622		use	v	**resources**	n		sci
623		valuable	adj	**resources**	n		sci
624		appropriate	adj	**response**	n		
625		emotional	adj	**response**	n		sci
626		individual	adj	**response**	n		sci
627		active	adj	**role**	n		sci
628		assume (the)	v	**role**	n		
629		central	adj	**role**	n		sci
630		consider (the)	v	**role**	n		sci
631		critical	adj	**role**	n		sci
632		crucial	adj	**role**	n		sci
633		direct	adj	**role**	n		sci
634		dominant	adj	**role**	n		sci
635		essential	adj	**role**	n		sci
636		examine (the)	v	**role**	n		sci
637		key	adj	**role**	n		sci
638		leading	adj	**role**	n		sci
639		major	adj	**role**	n		sci
640		minor	adj	**role**	n		sci
641		pivotal	adj	**role**	n		sci
642		play (a)	v	**role**	n		sci
643		prominent	adj	**role**	n		sci
644		significant	adj	**role**	n		sci
645		take (a)	v	**role**	n		sci
646		take on (the)	v	**role**	n		
647		take up (the)	v	**role**	n		
648		vital	adj	**role**	n		sci
649		average	adj	**score**	n		sci
650		high	adj	**score**	n		sci
651		mean	adj	**score**	n		sci
652		test	n	**score**	n		sci
653		concluding	adj	**section**	n		sci
654		final	adj	**section**	n		sci
655		introductory	adj	**section**	n		
656		opening	adj	**section**	n		

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
657		preceding	adj	**section**	n		sci
658		previous	adj	**section**	n		sci
659		business	n	**sector**	n		sci
660		economic	adj	**sector**	n		
661		manufacturing	adj	**sector**	n		sci
662		private	adj	**sector**	n		sci
663		public	adj	**sector**	n		sci
664		service	n	**sector**	n		
665		state	n	**sector**	n		
666	(be)	carefully	adv	**selected**	vpp		sci
667	(be)	randomly	adv	**selected**	vpp		sci
668		biological	adj	**sex**	n		
669		safe	adj	**sex**	n		
670		global	adj	**shift**	n		
671		major	adj	**shift**	n		sci
672		significant	adj	**shift**	n		sci
673		negative	adj	**side**	n		sci
674		highly	adv	**significant**	adj		sci
675		particularly	adv	**significant**	adj		sci
676		statistically	adv	**significant**	adj		sci
677		broadly	adv	**similar**	adj		sci
678		remarkably	adv	**similar**	adj		sci
679		relatively	adv	**simple**	adj		sci
680		alternative	adj	**source**	n		sci
681		become (a)	v	**source**	n		
682		common	adj	**source**	n		sci
683		external	adj	**source**	n		sci
684		key	adj	**source**	n		sci
685		main	adj	**source**	n		sci
686		major	adj	**source**	n		sci
687		original	adj	**source**	n		sci
688		possible	adj	**source**	n		sci
689		potential	adj	**source**	n		sci
690		primary	adj	**source**	n		sci
691		principal	adj	**source**	n		sci
692		provide (a)	v	**source**	n		sci
693		rich	adj	**source**	n		sci
694		secondary	adj	**source**	n		sci
695		single	adj	**source**	n		sci
696		use (a)	v	**source**	n		sci
697		useful	adj	**source**	n		sci
698		human	adj	**species**	n		sci
699		culturally	adv	**specific**	adj		

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
700		historically	adv	**specific**	adj		
701		domestic	adj	**sphere**	n		
702		private	adj	**sphere**	n		
703		public	adj	**sphere**	n		
704		economic	adj	**stability**	n		
705		political	adj	**stability**	n		
706		current	adj	**status**	n		sci
707		economic	adj	**status**	n		
708		equal	adj	**status**	n		
709		high	adj	**status**	n		
710		legal	adj	**status**	n		sci
711		low	adj	**status**	n		sci
712		political	adj	**status**	n		
713		professional	adj	**status**	n		
714		relative	adj	**status**	n		
715		social	adj	**status**	n		sci
716		socioeconomic	adj	**status**	n		sci
717		special	adj	**status**	n		
718		fairly	adv	**straightforward**	adj		sci
719		relatively	adv	**straightforward**	adj		sci
720		alternative	adj	**strategy**	n		sci
721		coping	adj	**strategy**	n		
722		develop (a)	v	**strategy**	n		sci
723		have (a)	v	**strategy**	n		sci
724		learning	n	**strategy**	n		
725		teaching	n	**strategy**	n		sci
726		use (a)	v	**strategy**	n		sci
727		cause	v	**stress**	n		sci
728		reduce	v	**stress**	n		sci
729		basic	adj	**structure**	n		sci
730		clear	adj	**structure**	n		sci
731		complex	adj	**structure**	n		sci
732		economic	adj	**structure**	n		
733		existing	adj	**structure**	n		
734		formal	adj	**structure**	n		sci
735		global	adj	**structure**	n		sci
736		hierarchical	adj	**structure**	n		sci
737		institutional	adj	**structure**	n		
738		internal	adj	**structure**	n		sci
739		organizational	adj	**structure**	n		sci
740		overall	adj	**structure**	n		sci
741		political	adj	**structure**	n		
742		social	adj	**structure**	n		sci

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
743		underlying	adj	**structure**	n		sci
744		brief	adj	**summary**	n		sci
745		present (a)	v	**summary**	n		sci
746		provide (a)	v	**summary**	n		sci
747		useful	adj	**summary**	n		sci
748		conduct (a)	v	**survey**	n		sci
749		national	adj	**survey**	n		sci
750		recent	adj	**survey**	n		sci
751		meet (a)	v	**target**	n		
752		set (a)	v	**target**	n		sci
753		carry out (the)	v	**task**	n		sci
754		complete (a)	v	**task**	n		sci
755		main	adj	**task**	n		sci
756		perform (a)	v	**task**	n		sci
757		primary	adj	**task**	n		
758		advanced	adj	**technology**	n		
759		current	adj	**technology**	n		sci
760		digital	adj	**technology**	n		
761		modern	adj	**technology**	n		sci
762		classic	adj	**text**	n		sci
763		introductory	adj	**text**	n		
764		key	adj	**text**	n		
765		literary	adj	**text**	n		
766		original	adj	**text**	n		
767		apply (the)	v	**theory**	n		sci
768		classical	adj	**theory**	n		sci
769		critical	adj	**theory**	n		sci
770		cultural	adj	**theory**	n		
771		develop (a)	v	**theory**	n		
772		economic	adj	**theory**	n		
773		evolutionary	adj	**theory**	n		sci
774		general	adj	**theory**	n		sci
775		scientific	adj	**theory**	n		
776		social	adj	**theory**	n		
777		test (a)	v	**theory**	n		sci
778		use (a)	v	**theory**	n		sci
779		cover (a)	v	**topic**	n		sci
780		discuss (a)	v	**topic**	n		sci
781		key	adj	**topic**	n		
782		related	adj	**topic**	n		
783		research	n	**topic**	n		sci
784		global	adj	**trade**	n		sci
785		radical	adj	**transformation**	n		

#	Ad	Component I	POS I	Component II	POS II	Ad	Sci
786		social	adj	**transformation**	n		
787		undergo	adj	**transformation**	n		sci
788		make (a)	v	**transition**	n		sci
789		public	adj	**transport**	n		
790		current	adj	**trend**	n		sci
791		general	adj	**trend**	n		sci
792		growing	adj	**trend**	n		sci
793		increasing	adj	**trend**	n		sci
794		show (a)	v	**trend**	n		sci
795		social	adj	**trend**	n		
796		remain	v	**unclear**	adj		sci
797		earlier	adj	**version**	n		sci
798		electronic	adj	**version**	n		
799		final	adj	**version**	n		sci
800		modified	adj	**version**	n		sci
801		online	adj	**version**	n		sci
802		original	adj	**version**	n		sci
803		revised	adj	**version**	n		
804		simplified	adj	**version**	n		sci
805		become	v	**widespread**	adj		sci

APPENDIX G
HEADWORDS OF THE POTENTIAL SAWL

Note: *Numbers in parenthesis = (Range / Frequency); * = appearing in AWL (Coxhead, 2000)*

001.absorb (11/1261)
002.abstract (11/975)*
003.abundant (11/1340)
004.academy (10/187)*
005.access (11/612)*
006.accompany (11/256)*
007.accumulate (11/891)*
008.accurate (11/1202)*
009.acetate (9/419)
010.acetic (8/267)
011.acetone (9/180)
012.achieve (11/1733)*
013.acid (9/7039)
014.acquire (11/823)*
015.activate (11/2816)
016.acute (11/255)
017.adapt (11/1083)*
018.additives (8/177)
019.adequate (11/291)*
020.adhesion (10/637)
021.adjacent (11/292)*
022.adjust (11/631)*
023.administrate (10/255)*
024.adult (9/745)*
025.adverse (10/188)
026.affect (11/2226)*
027.affinity (10/626)
028.agar (8/360)
029.agarose (8/230)
030.aggregate (10/610)*
031.aid (11/205)*
032.albumin (9/158)
033.alcohol (10/435)
034.algorithm (11/1536)
035.align (11/600)
036.alkaline (9/278)
037.allele (8/401)
038.allocate (10/166)*
039.alter (11/1005)*
040.alternative (11/1284)*
041.ambience (11/314)
042.amino (9/1684)
043.ammonia (9/550)
044.amplify (11/773)
045.amplitude (10/264)
046.anaerobic (9/265)
047.analogy (11/228)*
048.analyse (11/10162)*
049.anneal (10/681)
050.annual (10/297)*
051.antibiotic (8/420)
052.antibody (9/1386)
053.antimicrobial (8/173)
054.apoptosis (8/355)
055.apparent (11/549)*
056.append (10/284)*
057.approach (11/3622)*
058.appropriate (11/803)*
059.approximate (11/2406)*
060.aqueous (9/560)
061.arbitrary (11/275)*
062.architecture (10/241)
063.area (11/3347)*
064.aromatic (8/236)
065.array (11/780)
066.aspect (11/749)*
067.assay (10/2114)
068.assemble (11/701)*
069.assess (11/1726)*
070.assign (11/558)*
071.assist (11/409)*
072.assume (11/2441)*
073.atmosphere (11/392)
074.atom (11/947)
075.attach (11/508)*
076.attain (11/250)*
077.attitude (7/201)*
078.attribute (11/932)*
079.author (11/933)*
080.automate (11/376)*
081.available (11/2239)*
082.aware (11/178)*
083.axis (11/520)
084.bacillus (9/292)
085.bacterium (10/2419)
086.barrier (11/430)
087.basal (10/241)
088.baseline (11/240)
089.batch (10/414)
090.bead (10/424)
091.benefit (11/887)*
092.beta (11/376)
093.bias (11/814)*
094.binary (11/343)
095.bioactive (9/195)
096.biochemical (11/403)
097.biology (11/440)
098.biomass (10/1078)
099.biomedical (7/178)
100.biosynthetic (8/600)
101.biotechnology (8/206)
102.bond (11/855)*
103.bootstrap (8/303)
104.bovine (9/269)
105.breakdown (11/166)
106.breast (9/234)
107.brief (11/519)*
108.buffer (10/1350)
109.bulk (10/491)*
110.calcium (10/230)
111.calibrate (11/485)
112.cancer (11/1662)
113.candidate (11/494)
114.capable (11/654)*
115.capacity (11/1211)*
116.capillary (9/385)
117.capture (11/630)
118.carbohydrate (9/298)
119.carbon (10/2040)
120.cardiac (8/311)
121.cardiovascular (9/211)
122.cascade (10/252)
123.catalyse (9/780)
124.category (11/622)*
125.cavity (9/247)
126.cell (11/16056)
127.cellular (11/2189)
128.cellulose (9/472)
129.centrifuge (9/669)
130.cerevisiae (7/416)
131.challenge (11/983)*
132.chamber (9/438)
133.channel (11/611)*
134.chemical (11/2694)*
135.chemistry (11/313)
136.chip (9/394)
137.chloride (9/454)
138.cholesterol (6/296)

139.chromatography (9/693)
140.chromosome (9/387)
141.chronic (9/197)
142.clarify (11/216)*
143.classic (11/595)*
144.climate (10/370)
145.clinic (11/934)
146.clone (9/646)
147.cluster (11/2507)
148.coating (10/2440)
149.code (11/554)*
150.coefficient (11/1411)
151.coincide (11/186)*
152.colon (6/183)
153.column (11/2024)
154.comment (11/225)*
155.communicate (11/269)*
156.community (11/1152)*
157.compatible (11/300)*
158.compensate (11/192)*
159.complement (11/373)*
160.complex (11/2762)*
161.component (11/2512)*
162.composition (11/1915)
163.compound (11/2344)*
164.comprehensive (11/250)*
165.comprise (11/517)*
166.compute (11/1989)*
167.concentrate (11/3697)*
168.concept (11/1142)*
169.conclude (11/1581)*
170.conduct (11/1075)*
171.confer (11/221)*
172.configure (11/487)
173.confirm (11/1351)*
174.confocal (9/222)
175.conform (11/234)*
176.conjugate (11/700)
177.consequent (11/1102)*
178.conserve (11/616)
179.considerable (11/534)*
180.consist (11/2694)*
181.constant (11/1658)*
182.constitute (11/523)*
183.constrain (11/592)*
184.construct (11/2179)*
185.consume (11/1352)*
186.contact (11/1593)*
187.contaminate (10/555)
188.context (11/781)*
189.contrast (11/1456)*
190.contribute (11/1734)*
191.convene (11/849)*
192.converse (11/203)*
193.convert (11/958)*
194.convex (8/197)
195.cooperate (11/209)*
196.coordinate (11/638)*
197.core (11/1196)*
198.correlate (11/1957)
199.correspond (11/2705)*
200.couple (11/1055)*
201.covalent (7/164)
202.create (11/1097)*
203.criteria (11/908)*
204.crucial (11/375)*
205.crude (9/204)
206.crystal (10/968)
207.culture (11/3088)*
208.cumulative (9/183)
209.cycle (11/2225)*
210.cysteine (9/248)
211.dash (11/338)
212.data (11/7711)*
213.database (11/381)
214.dataset (10/572)
215.decade (11/401)*
216.decline (11/398)*
217.decompose (10/429)
218.defect (10/728)
219.deficiency (11/415)
220.define (11/3944)*
221.definite (11/197)*
222.degrade (11/1212)
223.dehydrogenased (8/262)
224.demonstrate (11/2525)*
225.denote (11/1187)*
226.dense (11/2910)
227.dependence (11/567)
228.depict (11/347)
229.deplete (10/270)
230.deposit (9/697)
231.deposition (9/803)
232.derive (11/2630)*
233.design (11/2389)*
234.despite (11/586)*
235.detect (11/4492)*
236.developmental (9/230)
237.deviate (11/713)*
238.device (11/1060)*
239.diagnosis (11/210)
240.diagnostic (10/268)
241.diagram (11/565)
242.diameter (10/955)
243.diet (7/1118)
244.differential (11/573)
245.differentiate (11/1413)*
246.diffract (7/325)
247.diffuse (11/1290)
248.digest (9/621)
249.digital (11/185)
250.dilute (10/889)
251.dimension (11/2615)*
252.discrepancy (11/241)
253.discrete (11/697)*
254.discriminate (9/239)*
255.disperse (11/818)
256.displace (10/252)*
257.display (11/1143)*
258.disrupt (10/420)
259.dissolution (8/194)
260.dissolve (9/455)
261.distil (10/256)
262.distinct (11/1014)*
263.distribute (11/4953)*
264.diverge (10/366)
265.diverse (11/1475)*
266.document (11/324)*
267.domain (11/2444)*
268.dominate (11/852)*
269.donor (9/424)
270.dose (11/732)
271.dotted (11/172)
272.downstream (9/278)
273.drama (11/272)*
274.droplet (8/847)
275.drug (11/1230)
276.dual (11/862)
277.duration (11/318)*
278.dye (9/474)
279.dynamic (11/1743)*
280.ecological (9/329)
281.economy (11/613)*
282.ecosystem (9/404)
283.edit (10/517)*
284.efficiently (11/223)
285.electrode (8/898)
286.electron (11/1566)
287.electrophoresis (8/170)
288.electrostatic (10/172)
289.element (11/1694)*
290.elemental (7/193)
291.elevate (11/582)
292.eliminate (11/539)*
293.elongate (10/261)
294.embed (11/285)

295.emerge (11/952)*
296.emit (9/667)
297.emphasis (11/318)*
298.empirical (11/525)*
299.enable (11/979)*
300.encode (10/711)
301.encounter (11/206)*
302.endogenous (9/485)
303.energy (11/3284)*
304.engineered (9/288)
305.enhance (11/1944)*
306.enrich (11/910)
307.ensure (11/648)*
308.environment (11/2622)*
309.enzymatic (9/409)
310.enzyme (9/2346)
311.epithelial (9/556)
312.equate (11/2396)*
313.equilibrium (11/608)
314.equip (11/394)*
315.equivalent (11/781)*
316.error (11/1776)*
317.establish (11/1318)*
318.ester (9/310)
319.estimate (11/4011)*
320.ethanol (9/785)
321.evaluate (11/2423)*
322.evaporate (10/295)
323.eventual (11/254)*
324.evident (11/1310)*
325.evolve (11/1461)*
326.exceed (11/295)*
327.excitation (10/397)
328.exclude (11/674)*
329.exhibit (11/1592)*
330.exogenous (9/202)
331.expand (11/829)*
332.expert (10/194)*
333.explicit (11/344)*
334.exploit (11/273)*
335.exponential (11/450)
336.expose (11/1783)*
337.external (11/465)*
338.extract (11/3493)*
339.fabricate (8/582)
340.facilitate (11/803)*
341.factor (11/4200)*
342.favorable (11/196)
343.favorably (8/17)
344.feasible (11/285)
345.feature (11/1434)*
346.feedback (10/167)
347.feeding (11/719)
348.ferment (8/1137)
349.fibre (10/1098)
350.file (11/387)*
351.filter (11/1143)
352.final (11/2178)*
353.finite (11/1462)*
354.flexible (11/354)*
355.fluctuate (10/380)*
356.fluid (11/523)
357.fluorescent (10/1803)
358.flux (9/610)
359.focus (11/1421)*
360.formula (11/1532)*
361.fraction (11/1443)
362.fragment (11/821)
363.framework (11/550)*
364.frequencies (11/264)
365.function (11/7973)*
366.functionalize (7/225)
367.fundamental (11/429)*
368.fungus (9/1706)
369.furthermore (11/1173)*
370.fuse (10/680)
371.gamma (11/185)
372.gel (10/832)
373.gene (11/5784)
374.genera (7/456)
375.generate (11/2394)*
376.generation (11/1035)*
377.genetic (10/1626)
378.genome (9/1614)
379.genotype (8/702)
380.genus (9/666)
381.geography (10/329)
382.geometry (10/826)
383.germ (8/181)
384.globe (11/871)*
385.glucose (10/1213)
386.glycerol (9/400)
387.goal (11/398)*
388.grade (10/534)*
389.gradient (10/702)
390.graph (11/745)
391.grid (11/570)
392.guarantee (11/173)*
393.gut (8/404)
394.height (11/450)
395.height (11/512)
396.hence (11/1232)*
397.heterogeneous (11/522)
398.hierarchy (11/246)*
399.highlight (11/750)*
400.homogeneous (11/401)
401.horizontal (11/279)
402.hybrid (11/956)
403.hydrogen (9/668)
404.hydrolysis (9/529)
405.hydrophilic (8/151)
406.hydrophobic (9/304)
407.hydroxyl (7/184)
408.hypothesis (11/1125)*
409.identical (11/487)*
410.identify (11/4441)*
411.illustrate (11/1157)*
412.image (11/8356)*
413.immobilize (9/368)
414.immune (11/836)
415.impact (11/1493)*
416.implement (11/832)*
417.implicate (11/432)*
418.imply (11/843)*
419.impose (11/184)*
420.incidence (11/416)*
421.incorporate (11/739)*
422.incubate (9/1537)
423.index (11/794)*
424.indicate (11/4784)*
425.individual (11/2357)*
426.induce (11/3027)*
427.infect (11/2686)
428.infer (11/540)*
429.inflame (9/949)
430.infrared (10/311)
431.inherent (11/169)*
432.inhibit (11/1974)*
433.initial (11/1891)*
434.initiate (11/650)*
435.inject (11/869)
436.injure (9/258)*
437.innovate (11/218)*
438.input (11/572)*
439.insert (11/532)*
440.inset (8/228)
441.insight (11/447)*
442.instance (11/706)*
443.instruct (11/478)*
444.intact (10/206)
445.intake (10/192)
446.integral (11/443)*
447.integrate (11/1102)*
448.intense (11/1448)*
449.interact (11/3762)*
450.interestingly (11/454)

451.interface (11/800)
452.interior (11/176)
453.intermediate (11/586)*
454.internal (11/700)*
455.interpret (11/688)*
456.interval (11/1142)*
457.intervene (11/231)*
458.intestine (8/545)
459.intrinsic (11/293)*
460.invasive (9/255)
461.inverse (11/296)
462.investigate (11/2543)*
463.involve (11/2370)*
464.ion (10/3098)
465.isolate (11/2475)*
466.item (10/254)*
467.kernel (10/347)
468.kidney (8/194)
469.kinase (7/979)
470.kinetic (10/722)
471.kit (9/403)
472.label (11/1681)*
473.laboratory (11/1343)
474.lactic (9/226)
475.laser (8/854)
476.latent (10/363)
477.lateral (9/298)
478.lattice (8/750)
479.layer (11/2541)*
480.linear (11/2739)
481.link (11/1277)*
482.lipid (9/1880)
483.liver (9/571)
484.localise (11/591)
485.locate (11/1798)*
486.locus (10/554)
487.longitudinal (10/164)
488.loop (11/668)
489.lysine (7/255)
490.magnesium (7/156)
491.magnet (10/1442)
492.magnify (9/244)
493.magnitude (11/424)
494.maintain (11/1370)*
495.major (11/1851)*
496.manipulate (11/331)*
497.manual (11/201)*
498.mapping (11/321)
499.margin (11/540)*
500.marine (9/486)
501.mathematics (11/3928)
502.matrix (11/3396)
503.mature (11/722)*
504.maximal (11/261)
505.maximise (11/2143)*
506.mechanism (11/2583)*
507.media (10/1070)*
508.median (11/243)
509.mediate (10/1026)*
510.medical (11/183)*
511.medium (11/1704)*
512.membrane (10/1461)
513.mesh (10/540)
514.metabolic (10/1187)
515.metabolism (9/830)
516.metabolite (8/1191)
517.methanol (9/361)
518.methionine (8/187)
519.method (11/8798)*
520.micro (10/340)
521.microbe (8/1079)
522.microorganism (9/560)
523.microscope (11/1289)
524.migrate (10/472)*
525.minimal (11/399)*
526.minimise (11/574)*
527.minimum (11/668)*
528.minor (11/270)*
529.mitochondria (9/645)
530.mobile (11/662)
531.mode (11/1398)*
532.modify (11/2323)*
533.molar (8/164)
534.molecule (11/3431)
535.monitor (11/884)*
536.morphology (10/1251)
537.mortal (9/661)
538.motive (11/233)*
539.mount (9/257)
540.muscle (10/1603)
541.mutant (8/1123)
542.mutate (9/1258)
543.nanoparticle (7/1342)
544.negate (11/1911)*
545.negligible (11/212)
546.network (11/1607)*
547.neural (9/259)
548.neutral (10/400)*
549.nevertheless (11/428)*
550.nitrogen (9/880)
551.node (10/1046)
552.norm (8/289)*
553.normal (11/2245)*
554.notion (10/211)*
555.novel (11/823)
556.nuclear (10/551)*
557.nucleotide (9/341)
558.nucleus (10/389)
559.null (11/514)
560.nutrient (10/987)
561.nutrition (8/340)
562.objective (11/341)*
563.obtain (11/5328)*
564.obvious (11/560)*
565.occupy (11/283)*
566.occur (11/2867)*
567.odd (7/238)*
568.online (11/500)
569.onset (9/163)
570.optic (10/1270)
571.optimal (11/1034)
572.optimise (11/1057)
573.optimum (11/253)
574.option (11/431)*
575.oral (8/219)
576.organic (9/1438)
577.organism (11/777)
578.orient (11/469)*
579.orthogonal (11/270)
580.outcome (11/415)*
581.output (11/414)*
582.oven (8/154)
583.overall (11/1029)*
584.overlap (11/380)*
585.overnight (10/295)
586.overview (11/290)
587.oxidant (9/759)
588.oxide (11/1019)
589.oxidise (9/313)
590.oxygen (9/1014)
591.pairwise (9/161)
592.panel (11/704)*
593.parallel (11/590)*
594.parameter (11/3590)*
595.participate (11/469)*
596.partner (10/178)*
597.passive (11/199)*
598.pathogen (9/1584)
599.pathogenic (10/241)
600.pathway (11/2227)
601.patients (10/871)
602.peak (11/2175)
603.penetrate (10/297)
604.peptide (8/1433)
605.perceive (10/274)*
606.percent (11/1156)*

607.period (11/1800)*
608.periphery (11/250)
609.peroxide (8/162)
610.persist (11/364)*
611.perspective (11/362)*
612.pharmaceutical (11/194)
613.phase (11/3749)*
614.phenomenon (11/581)*
615.phenotypic (7/198)
616.phosphate (9/805)
617.phylogenetic (8/816)
618.physical (11/855)*
619.physiological (10/568)
620.plasma (11/1585)
621.plasmid (8/659)
622.plastic (10/388)
623.platform (11/450)
624.plot (11/1208)
625.plus (11/297)*
626.polar (10/708)
627.poly (9/324)
628.polymer (8/792)
629.polymerase (8/186)
630.polynomial (10/482)
631.pooled (9/194)
632.pore (9/514)
633.portion (11/288)*
634.pose (11/158)*
635.positive (11/2365)*
636.posterior (8/562)
637.potassium (9/291)
638.potent (10/182)
639.potential (11/3503)*
640.precede (11/171)*
641.precipitate (10/446)
642.precise (11/723)*
643.precursor (11/438)
644.predict (11/2011)*
645.predominant (11/354)*
646.preliminary (11/187)*
647.presume (11/161)*
648.prevalent (11/398)
649.previous (11/3281)*
650.primary (11/1438)*
651.prime (11/241)*
652.primer (9/823)
653.principal (11/261)*
654.principle (11/533)*
655.prior (11/911)*
656.probe (11/1043)
657.proceed (11/2041)*
658.process (11/7045)*
659.profile (11/1155)
660.progression (9/246)
661.project (11/887)*
662.proliferate (9/560)
663.proline (8/176)
664.promote (11/1667)*
665.propagate (11/505)
666.proportion (11/1145)*
667.protease (8/229)
668.protein (11/9058)
669.protocol (11/712)*
670.publication (11/247)*
671.publish (11/844)*
672.pulse (9/555)
673.purchase (11/344)*
674.purify (10/809)
675.purity (9/183)
676.putative (9/302)
677.qualitative (11/234)*
678.quantify (11/1056)
679.quantitative (11/833)
680.radiate (10/437)
681.radical (11/354)*
682.radius (11/442)
683.random (11/1754)*
684.range (11/3616)*
685.ratio (11/2434)*
686.rational (11/290)*
687.react (11/3790)*
688.reagent (9/525)
689.receptor (9/1369)
690.recombination (8/180)
691.recover (11/1121)*
692.redox (8/280)
693.reference (11/1315)
694.refine (11/267)*
695.regime (11/377)*
696.region (11/3710)*
697.regress (10/873)
698.regulate (11/2776)*
699.reject (11/225)*
700.relax (11/333)*
701.release (11/1688)*
702.relevant (11/870)*
703.rely (11/801)*
704.remove (11/1812)*
705.replicate (11/1147)
706.require (11/3579)*
707.research (11/2450)*
708.reservoir (9/415)
709.reside (11/278)*
710.residue (11/1735)
711.resistance (11/2679)
712.resolve (11/1374)*
713.resonance (9/367)
714.resource (11/494)*
715.respond (11/3429)*
716.restore (11/297)*
717.restrict (11/833)*
718.resuspend (8/245)
719.retain (11/876)*
720.reveal (11/1568)*
721.reverse (11/864)*
722.revise (11/175)*
723.robust (11/645)
724.role (11/2504)*
725.rotate (11/621)
726.route (11/362)*
727.routine (11/235)
728.saline (9/427)
729.saturate (11/498)
730.scaling (11/578)
731.scan (11/937)
732.scenario (11/532)*
733.scheme (11/1229)*
734.score (11/815)
735.seasonal (8/294)
736.secrete (8/561)
737.section (11/3321)*
738.sector (11/175)*
739.seek (11/173)*
740.segment (11/710)
741.select (11/3564)*
742.sensing (10/417)
743.sensor (11/884)
744.sequence (11/4572)*
745.series (11/927)*
746.serine (7/158)
747.serum (9/818)
748.setup (11/297)
749.sex (10/589)*
750.shift (11/1035)*
751.sigma (10/587)
752.significant (11/5486)*
753.silica (8/384)
754.silicon (9/321)
755.similar (11/4666)*
756.simulate (11/1484)*
757.simultaneous (11/706)
758.site (11/3350)*
759.situ (10/284)
760.skeletal (9/227)
761.sodium (9/738)
762.software (11/768)

763.sole (11/171)*
764.soluble (9/823)
765.solvent (9/892)
766.somewhat (11/168)*
767.source (11/1928)*
768.spatial (11/1544)
769.species (11/8798)
770.specific (11/4715)*
771.specify (11/366)*
772.spectra (10/1568)
773.spectre (11/838)
774.spectrometer (9/231)
775.spectrometry (10/315)
776.spectroscopy (10/515)
777.spectrum (11/825)
778.sphere (11/567)*
779.spontaneous (9/264)
780.stable (11/2952)*
781.static (10/217)
782.stationary (11/439)
783.statistic (11/2092)*
784.status (11/346)*
785.stimulate (11/704)
786.storage (11/970)
787.straightforward (11/177)*
788.strain (10/3401)
789.strand (9/253)
790.strategy (11/2330)*
791.stress (11/1509)*
792.structure (11/5954)*
793.subsequent (11/1269)*
794.subset (10/390)
795.substitute (11/423)*
796.substrate (9/2580)
797.subunit (10/354)
798.successor (11/170)*
799.sufficient (11/1110)*
800.sum (11/918)*
801.summary (11/984)*
802.superior (11/283)
803.supernatants (9/397)
804.supplement (11/1394)*
805.suppress (11/973)
806.survey (11/786)*
807.survive (11/1184)*
808.susceptible (11/823)
809.suspend (10/718)*
810.sustain (11/337)*
811.switch (11/514)
812.symbol (11/339)*
813.symmetry (11/1174)
814.symptom (10/433)
815.synergistic (9/174)
816.synthesis (11/1914)
817.synthetic (10/550)
818.tank (9/422)
819.target (11/3156)*
820.task (11/431)*
821.taxonomy (9/401)
822.team (11/155)*
823.technical (11/266)*
824.technique (11/2098)*
825.technology (11/1025)*
826.template (11/261)
827.temporal (11/467)
828.terminate (11/836)*
829.tertiary (9/172)
830.text (11/198)*
831.theory (11/1879)*
832.therapeutic (11/467)
833.therapy (10/358)
834.thereby (11/474)*
835.thermal (10/1003)
836.threshold (11/658)
837.tissue (11/2492)
838.tolerance (11/506)
839.topic (11/333)*
840.topology (11/602)
841.toxic (9/522)
842.toxin (8/324)
843.trace (11/377)*
844.tract (10/212)
845.tradition (11/583)*
846.transcript (10/1326)
847.transfer (11/1804)*
848.transform (11/1734)*
849.transient (11/418)
850.transit (11/1553)*
851.transmit (11/842)*
852.transport (11/1232)*
853.trend (11/614)*
854.triangle (11/929)
855.trigger (11/453)*
856.triple (11/333)
857.triplicate (7/201)
858.tumour (9/976)
859.tyrosine (6/260)
860.ultimate (11/265)*
861.unclear (11/200)
862.undergo (11/377)*
863.underlie (11/481)*
864.uniform (11/709)*
865.unique (11/903)*
866.untreated (10/238)
867.uptake (10/474)
868.urea (9/194)
869.utilise (11/1111)*
870.vacuum (10/331)
871.valid (11/966)*
872.vary (11/7367)*
873.vascular (9/200)
874.vector (11/2134)
875.velocity (11/457)
876.verify (11/475)
877.version (11/1097)*
878.versus (11/1082)
879.vertical (11/392)
880.via (11/1742)*
881.viable (11/623)
882.video (10/308)
883.virus (9/1304)
884.visible (11/496)*
885.visual (11/888)*
886.vitamin (8/310)
887.vitro (10/1016)
888.vivo (10/1024)
889.volatile (11/479)
890.volt (10/482)
891.volume (11/2586)*
892.wavelength (10/585)
893.weighted (11/320)
894.whereas (11/1487)*
895.widespread (11/178)*
896.worldwide (10/222)
897.yeast (8/1059)
898.zinc (8/471)
899.zone (11/773)

BIOGRAPHY

Name	Mr. Todsaporn It-ngam
Date of Birth	April 25, 1983
Education Attainment	2001-2004: Bachelor of Arts, English (Hons) Srinakharinwirot University 2005-2007: Master of Arts, English Srinakharinwirot University 2012: Certificate of Teaching Thai to Foreigners School of Langauge and Culture, Technology Promotion Association (Thailand-Japan)
Work Position	Full-time English Instructor Language Institute Burapha University
Scholarship	2015-2018: Ph.D. Scholarship, Language Institute, Burapha University

Publications

It-ngam, T., & Phoocharoensil, S. (2019). The development of science academic word list. *Indonesian Journal of Applied Linguistics, 8*(3), 657-667. doi:http://dx.doi.org/10.17509/ijal.v8i3.15269

It-ngam, T., & It-ngam, S. (2013). Foreign languages in workplaces: A reflection from human resource officers. *HRD Journal*. *4*(1), 100-105.

Work Experiences	Dec 2011 – present: Full-time English Instructor, Language Institute, Burapha University
	Nov 2008 – Nov 2011: Full-time English Instructor, The Department of Western Languages, Faculty of Humanities and Social Sciences, Burapha University

May 2005 – Sep 2008: English Course Specialist, School of Language and Culture, Technology Promotion Association (Thailand-Japan)