Developing of a High Frequency Word List in Social Sciences

Sorawut Chanasattru
Supong Tangkiengsirisin
Language Institute, Thammasat University

Abstract

This study aimed to develop a high frequency content word list discovered in social science research articles, henceforth referred to as the Social Science Word List (SSWL) from the Social Science Corpus (SSC). The SSC was compiled from 64 open-access English social science research papers from 11 journals in the General Category, published during 2013-2015 on the ScienceDirect website. AntWordProfiler 1.4.0 and AntConc 3.4.3 were employed to calculate the ranges and frequencies of words occurring in the social science corpus, in comparison with the New General Service List (NGSL) and the Academic Word List (AWL). By using Coxhead's range and word frequency criteria, the results revealed that 394 high frequency content headwords and 1,120 word members were obtained. The validation results corroborated that the SSWL can assist teachers in selecting appropriate words. Also, the SSWL is worth introducing to students to familiarize them with essential words for the reading and writing of social science research papers in vocabulary pedagogy, as it exhibits twice the coverage of the AWL in the validating corpora.

Keywords: Corpus, corpus-based lexical study, high frequency social science word list, word list

บทคัดย่อ

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนารายการคำแสดงเนื้อหาที่พบบ่อยในบทความวิจัยด้านสังคมศาสตร์จากคลังข้อมูลบทความวิจัยด้านสังคมศาสตร์ที่สร้างจากบทความจำนวน 64 บทความ จากวารสารทางวิชาการแบบเสรี 11 ฉบับ ที่ตีพิมพ์ระหว่างปี 2556-2558 และถูกจัดอยู่ในประเภททั่วไปในฐาน ScienceDirect โปรแกรม AntWordProfiler 1.4.0 และ AntConc 3.4.3 ใช้คำนวณหาค่าความถี่และพิสัยของคำที่ปรากฏในคลังข้อมูลบทความวิจัย โดยเปรียบเทียบกับรายการคำที่พบบ่อยในภาษาอังกฤษ (New General Service List) และรายการคำที่พบบ่อยในงานทางวิชาการ (Academic Word List) โดยใช้เกณฑ์พิสัยของ Coxhead และความถี่การปรากฏของคำ ผลการวิจัยพบคำแสดงเนื้อหาที่พบบ่อยในงานวิจัยด้านสังคมศาสตร์ จำนวน 394 คำหลัก และสมาชิกของคำหลัก 1,120 คำ นอกจากนี้ ผลการตรวจสอบพบว่าคำแสดงเนื้อหาที่พบส่วนใหญ่นั้นเป็นคำที่ปรากฏในเอกสารทางวิชาการ และรายการคำที่ได้นี้ครอบคลุมคำในคลังข้อมูลทดสอบบทความวิจัยด้านสังคมศาสตร์โดยมากกว่ารายการคำที่พบบ่อยในงานทางวิชาการประมาณสองเท่า จึงเห็นได้ว่ารายการคำแสดงเนื้อหาที่พบบ่อยในบทความวิจัยด้านสังคมศาสตร์นี้ เหมาะสมที่จะแนะนำให้นักเรียนได้คุ้นเคยกับคำศัพท์ที่จำเป็นในการอ่านและเขียนงานวิจัยทางด้านสังคมศาสตร์

คำสำคัญ: คลังข้อมูล, การศึกษาคำศัพท์โดยใช้คลังข้อมูล, รายการคำที่พบบ่อยทางสังคมศาสตร์, รายการคำ

Introduction

Vocabulary knowledge is a key aspect in reading proficiency (Nation, 2001; Chanier & Selva, 1998; Groot, 2000). The importance of vocabulary is affirmed in a famous quotation from David Wilkins, who wrote that "…while without grammar, very little can be conveyed. Without vocabulary, nothing can be conveyed" (Wilkins, 1972, pp. 111–112). Like Wilkins, Lewis (1993, p. 89) declared that "Lexis is the core or heart of language". These quotations exemplify that without sufficient knowledge of vocabulary, English learners can neither understand a text nor convey their own ideas. Vocabulary acquisition is regarded as a vital factor in learning a foreign language since vocabulary knowledge has a direct effect on the reading and writing proficiency of English learners (Nation, 2001). For academic purposes, learning vocabulary aids students' academic achievement because vocabulary greatly contributes to reading comprehension and proficiency (Chanier & Selva, 1998; Tozcu & Coady, 2004). Students who have better vocabulary knowledge generally perform better in reading comprehension. Previous vocabulary knowledge also has a considerable effect on the reading comprehension of university students (Constantinescu, 2007). Students who have high levels of previous vocabulary knowledge can read faster than those who have a limited vocabulary (Calvo, Estevez, & Dowens, 2003). Additionally, vocabulary knowledge is very closely and positively related to reading comprehension (Stahl, 1990; Salah, 2008; Anjomshoa & Zamanian, 2014). Vocabulary knowledge can assist reading comprehension, and reading can in turn contribute to an increase in vocabulary size (Chall, 1987). Similarly, a substantial vocabulary size can contribute to reading comprehension (Groot, 2000). It is impossible to understand texts either in one's own language or in a foreign language without understanding essential vocabulary (Laufer, 1997). Moreover, not only does general vocabulary knowledge affect students' reading comprehension, but specific

vocabulary also assists students to comprehend a specific text effectively (Mehrpour & Rahimi, 2010). Readers' familiarity with the key vocabulary of a particular field combined with reading strategies leads to better reading comprehension (Mehrpour & Rahimi, 2010). From the aforesaid studies, teaching specific vocabulary not only leads to greater familiarization with general vocabulary, but also facilitates students to comprehend a text more effectively.

Regarding vocabulary pedagogy in Thailand, inadequate vocabulary knowledge is one of the predicaments that Thai students encounter when reading (Aegpongpaow, 2008). Many educational institutions introduce new vocabulary to their students to prepare them for reading and writing assignments for academic texts in specific fields. However, one problem is that teachers cannot decide which vocabulary should be introduced to prepare students to read and write research articles. This is because they do not know which words frequently appear and are truly representative of the vocabulary in specific fields. To this end, this study attempted to develop a Social Science Word List (SSWL), a list of content words representing words frequently utilized in social science research articles. The aim is for the SSWL to be used in vocabulary pedagogy in both undergraduate and graduate studies in the field of social science. Teachers, instructors and course designers of English for Academic Purposes (EAP) and English for Specific Purposes (ESP) may select appropriate vocabulary to introduce to students studying in the social science field. This word list can assist undergraduate and graduate students in familiarizing themselves with essential words useful for the reading and writing of social science papers.

Research Question

This study intended to develop a high frequency content word list in social science research papers. The research question of this study is as follows:

"What content words are found most frequently in social science research papers?"

Literature Review

Vocabulary classification Nation (2001, p. 11) classified vocabulary into four groups as follows:

High Frequency Words

High frequency words refer to words that appear most often in printed materials. High frequency words include both function words: *the, a, in, for,* etc. and content words: *social, forests, student, research,* etc. A well-known list of high frequency words is Michael West's (West, 1953) General Service List (GSL), which contains approximately 2,000 word families. Another high frequency word list, created by Charles Browne, Brent Culligan and Joseph Phillips, is known as the New General Service List (NGSL). The New General Service List (NGSL) is a new vocabulary list obtained from a major update of West's GSL. The NGSL contains 2,801 headwords divided into three levels, which represent the most significant high frequency words of the English language required by L2. The NGSL covers more than 90 percent of most general English texts (Browne, 2014).

Academic Words

Academic words are derived from academic textbooks or academic materials in different registers, for instance: linguistic, politics, business, etc. These words cover about nine percent of the running words in the texts. The most well-known academic word list is Coxhead's Academic Word List (AWL). The AWL includes 570 word families covering around 10 percent of academic texts (Coxhead, 2000; Coxhead & Byrd, 2007). However, it covers only 1.4 percent in fiction, which is the same size as academic texts (Coxhead, 2000).This group of words is most suitable for learners who want to read or write academic texts.

Technical Words

Technical words are closely associated with topics and subject areas, and they vary depending on subject area. These words cover around five percent of the running words in a text. This word group can be identified by systematically limiting the range of topics or language investigated or by using a technical dictionary. In technical dictionaries, approximately 1,000 entries in each dictionary are technical words. Both academic and technical words are sometimes known as specialized words.

Low-frequency Words

Low-frequency words are a very large group of words that rarely occur. They account for only a small proportion of any text or only about five percent of the words in an average academic text. These words consist of words that are not high frequency words, not academic words, and not technical words in a particular subject.

Low-frequency words can be proper names, or words that are infrequently used.

Word Classes

The traditional approaches to language analysis divide words into nine classes: nouns, verbs, adjectives, pronouns, prepositions, conjunctions, adverbs, determiners, and interjections (Nunan, 2013, p. 48). Many single words may be classified into different classes according to their function and meaning in a sentence. In English, the vocabulary system consists of closed class and open class categories.

Closed Class

Closed class refers to a class to which no new words are added. This class is sometimes also called "function words" or "grammar words" (Nunan, 2013, p. 49). Members of function words include adverbial particles, auxiliary verbs, conjunctions, determiners, modal verbs, numbers, prepositions, pronouns, and quantifiers (Gilner & Morales, 2006; Nation, 2001; Nunan, 2013). Although function words are small, they make up a significant amount of corpus running words. Johansson and Hofland (1989) as cited in Nation (2001, p. 206) points out that function words cover around 43-44 percent of the running words in most texts.

Open Class

Open class denotes a class to which new words can be added easily. Nouns, verbs, adjectives and adverbs are associated with this class. This class may sometimes be called "content words". Content

words play a significant role in allowing English users to communicate about different events or states of affairs. New content words, which are added to the English lexicon, prevent fossilization in the language (Nunan, 2013).

Previous Studies on Word List Developments

Word lists have been studied in diverse registers. For example, Wang, Liang, and Ge (2008), created a 623 word family Medical Academic Word List (MAWL) from a one million word medical corpus. Words included in the MAWL must appear at least 30 times in the medical corpus. Vongpumivitch, Huang, and Chang (2009) created a 1.5 million word corpus from applied linguistics research articles called the ALC from 200 applied linguistics research articles from five international journals: Applied Linguistics, Language Learning, The Modern Language Journal, Language Research, and TESOL Quarterly. The 585 applied linguistics word list was obtained from words that appeared at least 50 times in the ALC. Moreover, word lists in education (Mozaffari & Moini, 2014) and nursing (Yang, 2015) registers have been developed. Mozaffari and Moini (2014) created a 356 educational word list from a 1.7 million word corpus compiled from 239 research articles in the education field. Words found at least 50 times in the educational research article corpus and appearing more than five times in all the educational journals were included. Similarly, Yang (2015) unveiled a 676 nursing word list that was created from around a one million word Nursing Research Articles Corpus (NRAC) from 252 nursing research papers. Words discovered more than 33 times in the NRAC and occurring more than 11 times in 21 subject areas were included in the nursing word list. Also, Hsu (2013) developed a list of 595 frequently occurring word families in medical science called the Medical Word List (MWL) obtained from

a corpus of 155 textbooks in 31 medical areas from 15 million running words of e-book databases by employing range and frequency of words that were outside the most frequent 3,000 word families of the British National Corpus. The MWL accounted for 10.72 percent of the medical textbooks in this study. Medical instructors who use the MWL can raise students' awareness and foster vocabulary knowledge of the commonly used medical words in medical textbooks. Conversely, some word lists have been generated from smaller size corpora. Shabani and Tazik (2014), for example, generated the Revised Academic Word List (RAWL) from around a 300,000 token corpus obtained from 120 English research articles from two journals ' ESP and Asian EFL journals. Words occurring at least nine times in the corpus and at least twice in each journal were included in the RAWL. Likewise, Liu and Han (2015) created a total of 458 words of the Environmental Academic Word List (EAWL) based on around an 800,000 running word corpus from environmental research papers.

From the aforementioned studies, most researchers selected words based on range as the first criteria followed by word frequencies, similar to Coxhead's AWL word selection (Coxhead, 2000). Furthermore, West's GSL was criticized for being based on an outdated and relatively small corpus (approximately 2.5 million running words) compared to current corpus standards. GSL was also disparaged for not clearly defining "word" (Browne, Culligan, & Phillips, 2013). For this reason, the New General Service List (NGSL), which was based on a more modern and ten times larger corpus -- around 273 million words -- than the GSL, was employed in this study. In addition, corpus have usually been compiled from textbooks; research articles in diverse fields such as medical, agricultural, linguistics, financial, or education; or from politics, economics and business news. Nevertheless, social science research articles have seldom been the focus of studies.

Methodology

Journal and Paper Selection

The Social Science Corpus (SSC) was generated from open−access English academic articles in the Social Sciences, General Category of the ScienceDirect website (http://www.sciencedirect.com) published during the period 2013−2015 in order to avoid outdated words. These articles were statistically selected by the simple random sampling method. Some journal sections, such as forewords, editorials, and book reviews, were excluded.

Data Standardization

Graphs, charts, diagrams, equations, bibliographies, text headers, footnotes, author's name and affiliates, or other parts of the texts that cannot be processed by concordance programs were removed from the research paper files as suggested by Chen and Ge (2007) by using PDF editing software. Standardized PDF paper files were then converted to UTF−8 plain text format in order to avoid the wrong character conversion from non−English language characters. These text files were checked for typographical errors, and all hyphenated words were joined together as a single word. Subsequently, the running words of each text file were counted before combining them to create a journal sub−corpus. The running words of each journal were limited to between 36,000 to 39,000 running words. This was done to balance the number of running words in order to avoid word selection bias arising from long texts. The SSC consisted of 11 sub−corpora from 11 journals in different areas in the social science field. The journal information, subject areas used in Social Science Corpus (SSC), as well as the number of articles, types and tokens of each journal are displayed in Table 1.

Table 1 Summary of journals used in the social science corpus (SSC)

ID	Journal Title	Subject Areas	No. of Articles Used	No. of Types	No. of Tokens
S_1	Energy Research & Social Science	Energy Technologies, Fuels, Resources, and Energy Production Affecting People	4	4,501	38,131
S_2	Environmental Innovation and Societal Transitions	Innovations and Socio-Economic Transitions, Environmental Problems, and Environmental Sustainable Economy	4	4,194	38,261
S_3	Procedia – Social and Behavioral Sciences	Social Behavioral Sciences concerning Arts and Humanities	11	4,904	37,878
S_4	Public Relations Review	Public Relations, Mass Communications, Organizational Communications, Marketing Management and Public Policy Formation	6	5,038	37,732
S_5	Sport Management Review	Sport Management and Marketing	4	4,504	38,799
S_6	Studies in Communica-tion Sciences	Public Communication	5	4,489	35,985
S_7	The Journal of Social Studies Research	Social Science Studies	5	4,384	38,074
S_8	The Social Science Journal	Social Sciences, Humanities, and Natural Sciences	7	4,709	38,253
S_9	Travel Behaviour and Society	Travel Behavior, Transportation and Environmental, Transportation Geographic Information Systems (TGIS)	6	4,067	38,869
S_10	Wine Economics and Policy	Wine Business and Economics Management	6	4,880	36,481
S_11	Urban Climate	Urban Climatic Conditions and Change concerning Geography Demographic, and Socioeconomic	6	4,316	36,082
		Total	64	49,986	414,545

Research Tools

Reference Word Lists

New General Service List (NGSL)

The New General Service List (NGSL) version 1.01 was downloaded from the website "A New General Service List (1.01)" (http://www.newgeneralservicelist.org). The list contains words with three different levels of occurrence in Microsoft Excel format. Level one represents the 1,000 words that appeared most frequently in general English. The second and third levels present the next most frequent 1,000 words and 801 elements of vocabulary in general English, respectively. At each level, headwords and their word members were reformatted in the same pattern as in the Academic Word List (AWL) so that the AntWordProfiler could identify headwords and their word family members as suggested by Paul Nation's Range program and Laurence Anthony's AntWordProfiler help file. The list is comprised of headwords in the left column and family members characterized by Bauer and Nation's Level 6 scale: inflectional suffixes of Laurie Bauer and Paul Nation's word family taxonomy (Bauer & Nation, 1993), by right indenting. The NGSL includes the headwords in all parts of speech and all inflected forms but excludes headwords with a prefix and headwords with non-inflectional suffixes (derivational suffixes) as a different headword. Discrepancies between American English and British English spelling words are grouped within the same headword, but the NGSL presents an American English spelling word as a headword and assigns a British English spelling word as a headword member (Browne, 2014). High frequency words in general English such as numbers, days of the week, and months of the year were excluded from the NGSL and from this study. The New General Service List (NGSL) is comprised of 2,801 headwords and 8,452 family members.

Academic Word List (AWL)

The Academic Word List (AWL) contains headwords and family members characterized by Bauer and Nation's Level 6 scale: frequent but irregular affixes. This level of affixes contains all inflections and the most frequent and regular prefixes and suffixes, such as *−able, −ee, −ic, −ify, −ion, −ist, −ition, −ive, −th, −y, pre−, re* (Bauer & Nation, 1993, p. 261). The AWL classifies British English spelling words as a headword and sets American English spelling words as a word member. The Academic Word List (AWL) consists of 570 headwords and 3,107 family members. The AWL can be downloaded from the Victoria University of Wellington, New Zealand at http://www.victoria.ac.nz/lals/resources/academicwordlist/awl−headwords.

Concordancers

Anthony's AntWordProfiler (1.4.0) and AntConc (3.4.3) (Anthony, 2014a, 2014b) were used in this study. In this study, the General Vocabulary Profiling Tool of the AntWordProfiler (1.4.0) was the primary tool used to generate statistical values, the range and the frequency of the appearance of words across sub−corpora in the social science corpus, in comparison with the New General Service List (NGSL) and the Academic Word List (AWL). The AntConc (3.4.3) was only used to exclude function words out of content words.

Social Science Word List (SSWL) Generation Process

Stage 1: word selection criteria

Word selection for the SSWL was also based on the method by which Coxhead (2000) selected words for the Academic Word List (AWL). The range and frequency criteria as utilized by Coxhead's Academic Word List (AWL) were employed in this study. The word range was the first criterion to consider word selection to avoid the domination of long texts or extremely high frequency words in some subject areas. This was followed by word family frequency and uniformity of frequency' minimum frequency of a word family occurring in each sub−corpus. To be precise, a word family range had to appear in at least five from 11 sub−corpora with a word family frequency of at least 10 times in each sub−corpus, and the word families had to occur at least 100 times in 414,545 running words to be counted in the SSWL, which is the same criteria used by Coxhead. The method of comparing the corpus running word size to Coxhead's calculated range and word frequency ratios was not employed in this study. This was because the Social Science Corpus (SSC) is more than seven times smaller than Coxhead's corpus. If the idea of proportion were applied, the range and the frequency of word sizes would be only two and 14 times, respectively, in the whole SSC, which is extremely low, thus resulting in an abundance of representative words in the Social Science Word List (SSWL). In terms of uniformity of frequency, word family members had to be present at least ten times in each corpus. Further−more, words selected for the SSWL could be those appearing in the NGSL or in the AWL. This concept was dissimilar to Coxhead's word selection in that the words in the AWL must be the word families that are not listed in West's General Service List (GSL) (1953), a 2,000−frequency word list. This was because the present study aimed to create a list of words that are frequently used in social science research

articles, which can be either high frequency words found in general English texts or academic words found mainly in academic papers and research articles. The word output obtained from this stage was amalgamated with function and content words.

Stage 2: functional word removal

The Stop List function of the AntConc (3.4.3) program (Anthony, 2014a) was applied to separate function words from content words. In this study, the list of function words was modified using Paul Nation's list of 320 function words (Nation, 2001) as a primary function word list. In addition, some function words in Gilner and Morales' list (Gilner & Morales, 2006) were added for function words that were not listed in Paul Nation's list, for example, *across, against, because, whether,* etc. to filter only the content words. Only content words were retained to be employed in the creation of the Social Science Word List (SSWL).

Stage 3: word list examination

Proper names, prefixes, numbers, or words from other languages than English were removed from the list of content words. Finally, the words remaining in this stage were the Social Science Word List (SSWL).

Establishing the Validity of the Social Science Word List (SSWL)

Headword rank comparison and coverage were employed for validation of the SSWL. Firstly, the rank of the SSWL headwords were compared to those in the Corpus of Contemporary American English (COCA) Core Academic Vocabulary List (AVL) (Davies & Gardner, n.d.,−a), and the COCA Most Frequent Word List in Academic English (Davies & Gardner, n.d.,−b). Secondly, the coverage of the SSWL was checked with two small validating corpora containing new social science research papers in accordance with the method suggested by Liu and Han (2015); both corpora were generated using the same paper selection criteria and the method used in the SSC creation. The first validating corpus was compiled from 20 research papers selected from different social science academic journals used in the SSC compilation, while the other corpus was created from 20 research papers taken from the same social science academic journals used in the SSC compilation.

Results and Discussion

This study aimed to investigate and develop the Social Science Word List (SSWL), a high frequency content word list from social science research papers, in order to answer the following research question: “What content words are found most frequently in social science research papers?”

Identification of the Social Science Word List (SSWL)

The Social Science Word List (SSWL) differed from most word lists in previous studies that presented academic words in diverse registers. The SSWL contained both high frequency general words and academic words, since, in fact, social science academic papers might contain high frequency general words, academic words, technical words, as well as low frequency words as classified by Nation (2001, p. 11). In order to familiarize students with high frequency words used in social science research papers, they are required to learn high frequency words found in general English texts in addition to academic words used mainly in academic papers and research articles. This notion is also in accordance with the study of Mehrpour and Rahimi (2010), which found that general vocabulary and specific vocabulary both play a crucial role in effective reading comprehension.

The first 1,000 headwords of the NGSL accounted for 62.18 percent (245 words), whilst the next 1,000 headwords in the second level and the following 801 word families in the third level of NGSL covered 5.33 percent (21 words) and 0.25 percent (1 word), respectively. In total, the SSWL contained 267 headwords from the NGSL covering 67.77 percent. Additionally, content words accounted for 32.24 percent (127 words). One content word –– *better,* which was not shown in either the NGSL or the AWL, was also included in the SSWL by grouping it as a word family member under the headword *good,* which represents the closest meaning of the word *better.* Word family members of each head word in the SSWL were drawn from those in the AWL and the NGSL that occurred in the SSC. Any word family members that were not present in the SSC were removed. The SSWL consisted of 394 content words and 1,120 word family members in all three levels of the NGSL and AWL as presented in Table 2.

Table 2 Summary of the Social Science Word List (SSWL)

Word Type	Reference Word List				Total
	NGSL-Level 1	NGSL-Level 2	NGSL-Level 3	AWL	
Content Words (SSWL)	245 (62.18%)	21 (5.33%)	1 (0.25%)	127 (32.24%)	394 (100.00%)

A complete list of the 394 headwords and 1,120 word family members of the SSWL is presented alphabetically in Appendix A.

Examination of the validity of the Social Science Word List (SSWL)

The SSWL was examined for validity by two methods: headword rank comparison and SSWL coverage.

Headword Rank Comparison

The first method was comparing the rank of the SSWL headwords to those in the Corpus of Contemporary American English (COCA) Core Academic Vocabulary List (AVL) (Davies & Gardner, n.d., a) sorted by word frequency in the social science genre and the COCA Most Frequent Word List in Academic English (Davies & Gardner, n.d., b). First frequency-sorted headwords in the SSWL were compared to the 3,015 headwords of The COCA Core AVL. If the headwords in the SSWL were not found in the COCA Core AVL, the COCA 20,000 Most Frequent Word List in Academic English was used to check the SSWL's word rank instead. The results revealed that, as can be seen in table 3, among the 394 SSWL headwords, 232 head-words (58.88 percent) were academic words found in the COCA Core AVL in the social science category. Meanwhile, the rest (162 headwords

or 41.12 percent) were high frequency words in the COCA Most Frequent Word List.

Table 3 Summary of the number of SSWL headwords found in the COCA word lists

COCA Word Lists	No. of Headwords	Percent
Core Academic Vocabulary List	232	58.88%
Most Frequent Word List in Academic English	162	41.12%
Total	394	100.00%

It is interesting to note that 17 out of the 162 SSWL headwords found in the COCA Most Frequent Word List in Academic English were domain-specific words. These words occurred with a frequency at least three times higher in one or two among the nine genres in the COCA corpus. As demonstrated in table 4, these words were mostly found in the genres of Education, Education and Social, Medical, Finance, Law, and Religion. For example, the words *community, program, utilize, learn, implement, diverse, skill,* and *ability* were discovered mainly in the field of Education; the words *behavior* and *statistic* primarily occurred in Education and Social, while the words *invest* and *company* were found chiefly in Finance.

Table 4 List of headwords found in the COCA most frequent word list in Academic English classified by other domains

No.	Reference Word List	Head word	Range	Freq.	Rank in COCA Word List	Domain
1	AWL	community	8	680	134	Education
2	NGSL–1	program	8	350	89	Education
3	NGSL–1	behavior	6	333	211	Education and Social
4	NGSL–1	risk	7	331	399	Medical
5	AWL	invest	8	301	1835	Finance
6	NGSL–1	health	5	269	174	Medical
7	AWL	utilise	7	264	2204	Education
8	NGSL–1	company	6	260	187	Finance
9	NGSL–1	learn	5	236	243	Education
10	AWL	implement	7	180	1124	Education
11	AWL	regulate	6	160	2363	Law
12	AWL	diverse	5	152	1579	Education
13	NGSL–1	skill	6	152	365	Education
14	AWL	statistic	6	118	8417	Education and Social
15	NGSL–3	frequency	5	117	1330	Medical
16	NGSL–2	self	6	176	1297	Religious
17	NGSL–1	ability	5	100	434	Education

The Social Science Word List (SSWL) Coverage

The second method was an examination of the coverage of the SSWL with two validating corpora. To this end, two small-sized corpora containing new social science papers were compiled according to the method of Liu and Han (2015). The headwords and their word family members in the SSWL file used in the coverage examination of both validating corpora were formatted in the same pattern as the AWL. Both validating corpora were created using the identical paper selection criteria and method that were applied in the creation of the SSC. Each corpus, however, had discrepancies in the following details:

The Validating Corpus 1 (VC01)

The twenty research papers in this corpus were selected from ten social science academic journals in the General Category, which was different from those used in the SSC. Two research texts were statistically selected from each journal. The VC01 contained 113,472 tokens.

The Validating Corpus 2 (VC02)

This validating corpus was also comprised of twenty academic papers derived from the same social science academic journals as those used in the compilation of the SSC, but the research papers used in the VC02 differed from those used in the SSC. Also, two research texts were statistically chosen from each journal. The VC02 comprised 140,411 tokens.

The results revealed that on average, the SSWL accounted for 21.52 percent and 24.16 percent in validating corpus 1 (VC01) and validating corpus 2 (VC02), respectively. Usually, the AWL accounts

for around 10 percent in academic texts (Coxhead, 2000; Coxhead & Byrd, 2007). The coverage of the SSWL in the two corpora was higher than those of the AWL, occurring in 5.5 to 17.6 percent of all papers of both validating corpora. The SSWL, however, covered higher than 20 percent in all social science academic papers taken from the same journals as those in the SSWL in the second validating corpus (VC02). To be precise, the SSWL covered from 21 to 29 percent in all papers (VC02_A1−VC02_J2). On the contrary, in the first validating corpus (VC01) collected from papers from different social science journals, 14 out of 20 social science papers presented SSWL coverage of more than 20 percent, whilst the coverage of the rest ranged from 14.0 to 18.0 percent. The coverage details of the SSWL and the AWL in each paper of both validating corpora are shown in table 5.

Table 5 Summary of the coverage of SSWL in each paper of validating corpora

Papers compiled in VC01	Coverage		Papers compiled in VC02	Coverage	
	AWL	SSWL		AWL	SSWL
VC01_A1	10.5	21.4	VC02_A1	12.3	22.0
VC01_A2	9.6	25.8	VC02_A2	12.2	21.1
VC01_B1	13	21.5	VC02_B1	12.9	24.4
VC01_B2	9.5	15.3	VC02_B2	12.1	23.5
VC01_C1	17.6	17.3	VC02_C1	16.4	29.4
VC01_C2	9.5	14	VC02_C2	17.6	26.8
VC01_D1	12.4	24.2	VC02_D1	10.1	22.7
VC01_D2	13.8	21.8	VC02_D2	7.7	22.1
VC01_E1	10.6	24	VC02_E1	12.3	28.3
VC01_E2	10.5	24.8	VC02_E2	14.1	24.0
VC01_F1	6.7	17	VC02_F1	14.8	27.0
VC01_F2	5.5	15.8	VC02_F2	11.4	25.4
VC01_G1	10.1	18	VC02_G1	10.5	23.8
VC01_G2	11.1	21.7	VC02_G2	10.7	23.4
VC01_H1	13.8	27.5	VC02_H1	12.3	24.5
VC01_H2	13.2	26.4	VC02_H2	14.3	22.5
VC01_I1	15.4	25.5	VC02_I1	13.0	23.2
VC01_I2	13.8	20.8	VC02_I2	11.4	23.8
VC01_J1	10.4	20	VC02_J1	10.7	23.6
VC01_J2	16.5	27.5	VC02_J2	11.4	21.7
Average Coverage	11.68	21.52	Average Coverage	12.41	24.16

Pedagogical Implications of the Social Science Word List (SSWL)

The Social Science Word List (SSWL) is aimed at being a ready-to-use word list for teachers, instructors and course designers of English for Academic Purposes (EAP) and English for Specific Purposes (ESP) in social sciences for undergraduate and graduate students. As presented in Appendix A, the SSWL consists of 394 headwords and their family members. As illustrated in Figure 1, the word on the left-hand side represents the headword and words indented to the right are their members. The number at the end of each word indicates its frequency in the SSC.

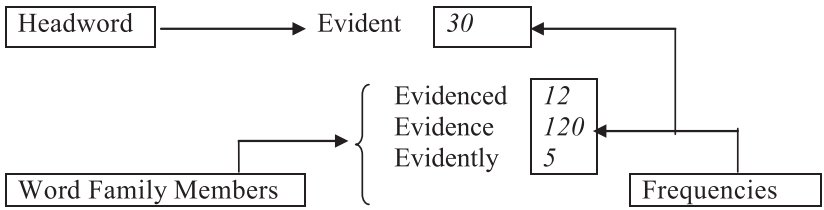

Figure 1 Structure of headwords and word family members of the Social Science Word List (SSWL)

Instructors should begin by introducing word families and their members found in the NGSL level 1 since the NGSL level 1 consists of the most frequent words among level 2 and 3. After students are familiar with the word families in the NGSL level 1, instructors may shift to teach them word families from NGSL levels 2 and 3 respectively. Introducing the word families found in the NGSL is consistent with the findings of Hirsh and Nation (1992) and Na and Nation (1985),

which suggested that L2 learners require a minimum of 3,000 high frequency English words so as to increase their understanding to at least 95 percent. Subsequently, after students have learnt the words from the NGSL, instructors should then introduce words from the AWL, since the coverage of the AWL was found to be far lower than the NGSL at just around 10 percent in academic texts (Coxhead, 2000; Coxhead & Byrd, 2007). This agrees with the findings of Nation and Kyongho (1995), and Mehrpour and Rahimi (2010), which suggested that both high frequency English words in the GSL and academic words in the AWL should be introduced to students studying in academic reading and writing classes. To be clearer, instructors may not need to teach all the word families and their members appearing in the SSWL. Instead, they may opt to introduce high frequency words. For instance, regarding the word family *evident* in Figure 1, in case of time constraints, teachers may introduce words with high frequency first because these words were discovered most often in social science research articles. For the word family *evident,* instructors may teach it only as a noun *evidence* and an adjective *evident,* which were the first and second ranks of frequency at 120 times and 30 times in the SSC, respectively. Moreover, information on word frequencies can also assist teachers to teach students how these words are frequently used in real academic prose contexts.

The SSWL can be applied in myriad ways. Flashcard learning is an effective direct vocabulary learning method for words from the SSWL regarding the return on time and effort spent for students who have limited knowledge of vocabulary. Learners and teachers can control vocabulary learning repletion and deliberately focus on some necessary words that cannot be easily gained from context or dictionary use methods (Nation, 2001; Nation & Meara, 2002). Teachers can create flashcards using headwords or word family members from the SSWL on one side and put word definitions on the other side. Other

lexical information, for instance, pronunciation, part of speech, collocations, grammatical patterns, and contexts in use, can also be added to flashcards to enhance learners' understanding (Nation & Meara, 2002). Nowadays, computer technology can aid teachers in producing their own flashcards without cutting paper. Many free and paid websites offer services for teachers that facilitate the creation of their own sets of flashcards effortlessly. To name a few, Quizlet (https://quizlet.com) is a free online vocabulary flashcard website generator allowing teachers or instructors to create their own flashcards by uploading the SSWL. The website not only creates online flashcards but also generates other useful online learning materials such as exercises and Scatter –– an interactive game. Meanwhile, Memrise (http://www.memrise.com) contains an outstanding feature that allows users to add Mempty, a picture of things or words associated with learning words, which can assist learners in memorizing target words.

The Limitations of the Social Science Word List (SSWL)

Word Frequency Counting

The major limitation of this corpus-based study derived from the limitations of the AntWordProfiler. The program automatically counts word frequency and groups of words with the same word forms that the program can recognize together without the classification of word class or word meaning in context. For instance, AntWord Profiler counts the adjective or adverb *fine* and the singular noun *fine*, or the proper noun *Fine* together under the same word family *fine* regardless of its meaning in context. Moreover, the word frequencies in SSWL are not tied to their meanings. The SSWL only presents frequencies of words discovered in social science research papers. The list does not distinguish how many times the word *fine* is in the

adjective class, in which it means “good”, from its meaning in the noun class of “money paid for punishment”.

Multiword Units

Multiword units are exceptions in the SSWL due to the occurrence of a space between them. AntWordProfiler counts and identifies words by noticing a space between each word. For this reason, multiword units are counted separately; for example, the word work out is counted as two separate words: one content word work and one function word out. Each word is stored under separate word families. This case differs from the counting of the word workout, with AntWordProfiler counting this word as one word.

American English and British English Spelling variations

The SSWL does not focus on the occurrence of American English and British English spelling variations in social science research articles. Consequently, the differences of the word frequency in American English and British English spelling variation was not taken into account in this study. The SSWL groups words with American English and British English spelling variations with their frequency together as one word as demonstrated in Appendix A.

Conclusion

This study created the Social Science Word List (SSWL), a list of 394 high frequency content words in social science research papers from 64 open-access English research papers in the social science, General Category in the ScienceDirect website. A total of 267

content words in the SSWL (67.77 percent) divided into three sub-levels at 62.18 percent, 5.33 percent, and 0.25 percent, respectively, were primarily derived from high frequency English words in the NGSL, whilst 127 academic words in the AWL (32.24 percent) were also included in the SSWL (Table 2). Although this result revealed that the SSWL mainly contained high frequency words rather than academic words, the first validity result corroborated that 232 words in the SSWL (58.88 percent) represented academic words in social science research papers in COCA Core AVL, while the rest (162 words) were general English words excluded from academic texts (Table 3). In terms of coverage, the second validity result demonstrated that on average the SSWL covered around 20 percent of academic words in social science papers, higher than the coverage of the AWL. The SSWL is worth introducing to students in order to familiarize them with the words occurring in social science research paper rather than relying on words from the NGSL or the AWL for the following reasons: The majority of words in the SSWL are academic words in the social science field compared with the words from the COCA Core AVL and the COCA Most Frequent Word List in Academic English, which were created from an enormous corpus. The SSWL can help learners to save time because the SSWL contains fewer words than the NGSL and the AWL, and the words in the SSWL are more genre-specific to the social science field than those in the AWL, which was compiled from academic texts in four disciplines: Law, Art, Commerce, and Science (Coxhead, 2000). Additionally, on average, the coverage of the SSWL is twice as high as the AWL, ensuring that learners using the SSWL have a greater chance to encounter words in academic texts in the social science field.

The limitations of this study are as follows. The SSWL was created from the SSC, a corpus compiled from academic papers only in the social science, General Category. In fact, social science in the

ScienceDirect website includes 12 other subject areas. Furthermore, the SSC contained only around 400,000 running words, and was therefore comparatively smaller than other corpora. These drawbacks may lessen the chances of discovering vocabulary that frequently appears in social science research articles. Apropos the limitations of the Social Science Word List (SSWL), three major restrictions, namely, word frequency counting, multiword units, and American English and British English spelling variations, were also highlighted as a reminder to teachers and instructors using the SSWL.

References

Aegpongpaow, O. (2008). *A qualitative investigation of metacognitive strategies in Thai students' English academic reading.* (Unpublished master's thesis), Srinakharinwirot University, Thailand.

Anjomshoa, L. & Zamanian, M. (2014). The Effect of Vocabulary Knowledge on Reading Comprehension of Iranian EFL *Learners in Kerman Azad University. International Journal on Studies in English Language and Literature (IJSELL), 2*(5), 90–95.

Anthony, L. (2014a). AntConc (Version 3.4.3) [Computer software]. Tokyo, Japan: Waseda University. Retrieved from http://www.laurenceanthony.net

Anthony, L. (2014b). Ant Word Profiler (Version 1.4.0) [Computer software]. Tokyo, Japan: Waseda University. Retrieved from http://www.laurenceanthony.net

Browne, C. (2014). A New General Service List: The Better Mousetrap We've Been Looking for?. *Vocabulary Learning and Instruction, 3*(1), 1–10. doi: 10.7820/vli.v03.1.browne

Browne, C., Culligan, B. & Phillips, J. (2013). *The New General Service List.* Retrieved from http://www.newgeneralservicelist.org

Calvo, M. G., Estevez, A., & Dowens, M. G. (2003). Time course of elaborative inferences in reading as a function of prior vocabulary knowledge. *Learning and Instruction, 13*(6), 611–631.

Chall, J. (1987). Two vocabularies for reading: Recognition and meaning. In M.G. McKeown, & M.E. Curtis (Eds), *The nature of vocabulary acquisition* (pp. 7–17). Hillsdale, NJ: Erlbaum.

Chanier, T., & Selva, T. (1998). The ALEXIA system: the use of visual representations to enhance vocabulary learning. *Computer Assisted Language Learning, 11*(5), 489–521.

Chen, Q., & Ge, G. (2007). A corpus–based lexical study on frequency and distribution of Coxhead's AWL word families in medical research articles (RAs). *English for Specific Purposes,* 26(4), 502–514.

Constantinescu, A. I. (2007). Using technology to assist in vocabulary acquisition and reading comprehension. *The Internet TESL Journal, 13*(2). Retrieved from http://iteslj.org/Articles/Constantinescu–Vocabulary.html

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

Coxhead, A., & Byrd, P., (2007). Preparing writing teachers to teach the vocabulary and grammar of academic prose. *J. Second Lang. Writing, 16*(3), 129–147.

Gilner, L., & Morales, F. (2006). *Academic Resources.* Retrieved from http://www.sequencepublishing.com/academic.html

Davies, M., & Gardner, D. (n.d.,–a). *Academic Vocabulary Lists (Corpus–based; 120 million words).* Retrieved from http://www.academicwords.info/download.asp

Davies, M., & Gardner, D. (n.d.,–b). *Word frequency: based on 450 million word COCA corpus.* Retrieved from http://www.word–frequency.info/5k_lemmas_download.asp

Groot, P. J. (2000). Computer assisted second language vocabulary acquisition. *Language Learning & Technology, 4*(1), 60–81.

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

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

Johansson, S., & Hofland, K. (1989). *Frequency analysis of English vocabulary and grammar: Based on the LOB Corpus.* Oxford, England: Clarendon.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 20–34). Cambridge: Cambridge University Press.

Lewis, M. (1993). *The lexical approach* (Vol. 1, p. 993). Hove: Language Teaching Publications.

Liu, J., & Han, L. (2015). A corpus–based environmental academic word list building and its validity test. *English for Specific Purposes,* 39, 1–11.

Mehrpour, S., & Rahimi, M. (2010). The impact of general and specific vocabulary knowledge on reading and listening comprehension: A case of Iranian EFL learners. *System, 38*(2), 292–300.

Mozaffari, A., & Moini, R. (2014). Academic Words in Education Research Articles: A Corpus Study. *Procedia–Social and Behavioral Sciences,* 98, 1290–1296.

Na, L., & Nation, I. S. P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal, 16*(1), 33–42.

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

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

Nation, I.S.P., & Meara, P. (2002). Vocabulary. In N. Schmitt (Ed.), *An introduction to applied linguistics* (2nd ed., pp. 34–52). London: Arnold.

Nunan, D. (2013). *What is this thing called language?* (2nd ed.). Basingstoke: Palgrave Macmillan.

Salah, S. M. (2008). *The relationship between vocabulary knowledge and reading comprehension of authentic Arabic texts.* (Unpublished master's thesis). Brigham Young University, Provo, UT.

Shabani, M. B., & Tazik, K. (2014). Coxhead' s AWL Across ESP and Asian EFL Journal Research Articles (RAs): A Corpus–based Lexical Study. *Procedia–Social and Behavioral Sciences, 98,* 1722–1728.

Stahl, S. A. (1990). Beyond the instrumentalist hypothesis: some relations between word meanings and comprehension. (Technical Report No. 505). University of Illinois at Urbana- Champaign: Center for the Study of Reading. (ERIC Document Reproduction Service No. ED321242.).

Tozcu, A., & Coady, J. (2004) Successful Learning of Frequent Vocabulary through CALL also Benefits Reading Comprehension and Speed. *Computer Assisted Language Learning,* 17 (5), 473-495.

Vongpumivitch, V., Huang, J., & Chang, Y. (2009). Frequency analysis of the words in the Academic Word List (AWL) and non-AWL content words in applied linguistics research papers. *English for Specific Purposes, 28*(1), 33-41. doi:10.1016/j.esp.2008.08.003

Wang, J., Liang, S. L., & Ge, G. C. (2008). Establishment of a medical academic word list. *English for Specific Purposes, 27*(4), 442-458.

West, M. (1953). *A General Service List of English words.* London, England: Longman, Green & Co.

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

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

Appendix A

The Complete List of 394 Word Families and 1,120 Word Family Members of the Social Science Word List (SSWL)

Below is the complete alphabetical list of the 394 word families of the Social Science Word List (SSWL) and their word family members. Words in italics indicate words that were found in the Social Science Corpus (SSC). The end of each word is marked with an asterisk (*) or a caret (^) symbol to indicate the word list that the word family was derived from: * for words from NGSL Level 1, ** for words from NGSL Level 2, *** for words from NGSL Level 3, and ^ for words for words from AWL. The number at the end of the symbol indicates the frequency of that word in the SSC. Words with American English and British English spelling variations are presented together and counted as one word.

Ability 88*
Abilities 12
Able 105*
Access^ 197
Accessed 5
Accesses 1
Accessibility 39
Accessible 26
Accessing 16
Inaccessible 2
Accord 1*
Accorded 1
According 210
Account 106*
Accounts 16
Accounted 13
Accounting 14
Achieve^ 53
Achieved 31
Achievement 13
Achievements 13
Achieves 1
Achieving 16
Activity 117*
Activities 242
Addition 150*
Additions 1
*Additional** 111*
Address 85*
Addresses 15
Addressed 45
Addressing 41
Advantage 88*
Advantages 41
Advantaged 1
Affect^ 63
Affected 59
Affecting 16
Affective 1
Affects 21
Unaffected 1
Age 128*
Ages 9
Aged 9
Aging 21
Ageing 9
Aim 56*
Aims 53
Aimed 61
Aiming 12
Allow 72*
Allows 63
Allowed 43
Allowing 38
Also 934*
Analyze/Analyse^ 51
Analysed 13
Analyses 71
Analysing 11
Analysis 487
Analyst 2
Analysts 2
Analytic 6
Analytical 28
Analytically 2
Analyzed 55
Analyzes 13
Analyzing 36
Answer 47*
Answers 34
Answered 11
Answering 9
Appear 57*
Appears 48
Appeared 16
Appearing 6
Application 64*
Applications 65
Apply 24*
Applies 10
Applied 75
Applying 30
Approach^ 251
Approachable 1
Approached 3
Approaches 76
Approaching 6
Area^ 156
Areas 161
Argue 58*
Argues 45
Argued 61
Arguing 13
Around 155*
Article 140*
Articles 42
Ask 20*
Asks 6
Asked 91
Asking 14
Aspect^ 42
Aspects 128
Assess^ 45
Assessed 22
Assesses 1
Assessing 17
Assessment 67
Assessments 23
Associate 6*
Associates 3
Associated 207
Associating 1
*Association** 79*
Associations 63
Assume^ 37
Assumed 42
Assumes 19
Assuming 22
Assumption 54
Assumptions 52
Attitude^ 65
Attitudes 80

Available^ 55
Availability 114
Unavailable 2
Average 146*
Averages 1
Averaged 5
Averaging 2
Aware^ 44
Awareness 64
Unaware 7
Base 50*
Bases 5
Based 461
Become 125*
Becomes 28
Became 52
Becoming 36
Begin 29*
Begins 12
Began 47
Beginning 37
Begun 16
Beginnings 3
Behavior/Behaviour 305*
Behaviors 25
Behaviours 3
*Belief** 73*
Beliefs 88
Benefit^ 60
Beneficial 21
Beneficiaries 2
Benefited 3
Benefiting 4
Benefits 173
Bring 63*
Brings 16
Brought 54
Bringing 15
*Broad** 44*
Broader 54
Broadest 2
Build 64*
Builds 10
Building 118
Buildings 20
Built 42
Business 290*
Businesses 114
Call 30*
Calls 9
Called 69
Calling 2
Capital 136*
Capitals 2
Case 352*
Cases 247
Cause 43*
Causes 39
Caused 44
Causing 12
Challenge^ 68
Challenged 11
Challenges 103
Challenging 30
Change 504*
Changes 257
Changed 46
Changing 37
*Characteristic** 26*
Characteristics 221
Choice 264*
Choices 93
Choose 52*
Chooses 6
Chose 20
Choosing 27
Chosen 68
Class 79*
Classes 22
Classed 3
Clear 97*
Cleared 1
Clearer 6
*Climate** 413*
Climates 2
Close 71*
Closes 2
Closed 16
Closing 7
Closer 26
Closest 9
Come 77*
Comes 43
Came 43
Coming 25
Common 145*
Communicate^ 12
Communicated 5
Communication 449
Communications 29
Communicative 24
Community^ 541
Communities 139
Company 75*
Companies 185
Compare 26*
Compares 4
Compared 134
Comparing 19
*Comparison** 78*
Comparisons 24
Complex^ 87
Complexities 10
Complexity 19
Concept^ 131
Conception 12
Concepts 120

Conceptual 49
Conceptualisation 1
Conceptualise 4
Conceptualised 2
Conceptualising 2
Conceptually 8
Concern 56*
Concerns 94
Concerned 56
Concerning 21
Conclude^ 28
Concluded 20
Concludes 9
Concluding 9
Conclusion 44
Conclusions 53
Conclusive 1
Condition 30*
Conditions 1 56
Conditioned 2
Conditioning 22
Conduct^ 18
Conducted 1 25
Conducting 18
Conducts 4
Consequent ^ 5
Consequence 43
Consequences 98
Consequently 44
Consider 143*
Considers 18
Considered 172
Considering 70
Consist^ 13
Consisted 14
Consistency 19
Consistent 96
Consistently 23
Consisting 6
Consists 24
Inconsistencies 7
Inconsistent 14
Construct^ 27
Constructed 52
Constructing 13
Construction 72
Constructions 14
Constructive 9
Constructs 17
Reconstruct 5
Reconstructed 4
Reconstructing 3
Reconstruction 14
Reconstructs 1
Consume^ 9
Consumed 11
Consumer 129
Consumers 336
Consumes 1
Consuming 5
Consumption 92
Context^ 285
Contexts 63
Contextual 12
Contextualise 1
Contextualised 1
Contextualize 3
Contextualized 3
Continue 54*
Continues 16
Continued 31
Continuing 16
Contribute^ 69
Contributed 28
Contributes 23
Contributing 36
Contribution 48
Contributions 53
Contributor 4
Contributors 5
Control 97*
Controls 10
Controlled 19
Controlling 16
Cost 112*
Costs 167
Country 110*
Countries 233
Create^ 116
Created 100
Creates 31
Creating 64
Creation 36
Creative 11
Creatively 2
Creativity 4
Creator 1
Creators 1
Recreate 2
Recreating 1
*Critical** 189
Culture^ 67
Cultural 177
Culturally 20
Cultures 16
Current 194*
Currents 1
Data^ 500
Day 127*
Days 63
Decision 291*
Decisions 139
Define^ 37
Defined 84
Defines 8
Defining 24
Definition 58
Definitions 27
Redefine 3
Redefined 1

- *Redefines 1*
- *Redefining 2*
- *Undefined 1*

Demand 151*
- *Demands 39*
- *Demanded 4*
- *Demanding 7*

Demonstrate^ 34
- *Demonstrably 1*
- *Demonstrated 42*
- *Demonstrates 20*
- *Demonstrating 8*
- *Demonstration 4*
- *Demonstrative 1*

Describe 53*
- *Describes 30*
- *Described 84*
- *Describing 13*

Design^ 163
- *Designed 60*
- *Designers 10*
- *Designing 15*
- *Designs 13*

Determine 59*
- *Determines 6*
- *Determined 29*
- *Determining 36*

Develop 139*
- *Develops 3*
- *Developed 142*
- *Developing 106*

Development 626*
- *Developments 35*

Difference 97*
- *Differences 133*

Different 461*

Direct 146*
- *Directs 3*
- *Directed 22*
- *Directing 1*

Discuss 57*
- *Discusses 7*
- *Discussed 94*
- *Discussing 23*

Discussion 107*
- *Discussions 37*

Distribute^ 8
- *Distributed 31*
- *Distributing 5*
- *Distribution 108*
- *Distributional 1*
- *Distributions 3*
- *Distributive 1*
- *Distributor 3*
- *Distributors 2*
- *Redistribution 4*

Diverse^ 59
- *Diversification 32*
- *Diversified 8*
- *Diversify 2*
- *Diversifying 3*
- *Diversity 43*

Dominate^ 19
- *Dominance 24*
- *Dominant 64*
- *Dominated 24*
- *Dominates 3*
- *Dominating 5*
- *Domination 7*

Draw 32*
- *Draws 18*
- *Drew 14*
- *Drawing 29*
- *Drawn 30*

Due 201*

During 297*

Early 150*
- *Earlier 71*
- *Earliest 6*

Economy^ 78
- *Economic 358*
- *Economical 7*
- *Economically 9*
- *Economics 162*
- *Economies 17*
- *Economist 1*
- *Economists 5*

Effect 201*
- *Effects 221*
- *Effected 1*

Effort 51*
- *Efforts 95*

Element^ 36
- *Elements 89*

Emerge^ 31
- *Emerged 40*
- *Emergence 39*
- *Emergent 8*
- *Emerges 16*
- *Emerging 52*

Empirical^ 111
- *Empirically 21*

End 115*
- *Ends 13*
- *Ended 19*
- *Ending 1*

*Engage** 56*
- *Engages 5*
- *Engaged 51*
- *Engaging 23*

Environment^ 113
- *Environmental 127*
- *Environmentalists 4*
- *Environmentally 8*
- *Environments 17*

Establish^ 43
- *Established 84*
- *Establishes 5*
- *Establishing 21*

Establishment 35
Establishments 3
Estimate^ 62
Estimated 46
Estimates 39
Estimating 14
Estimation 25
Estimations 3
Overestimate 5
Overestimates 4
Underestimate 4
Underestimated 2
Underestimates 1
Evaluate^ 40
Evaluated 24
Evaluates 1
Evaluating 20
Evaluation 94
Evaluations 35
Evaluative 25
Event 89*
Events 202
Evident^ 30
Evidenced 12
Evidence 120
Evidently 5
*Examine** 62*
Examines 20
Examined 47
Examining 37
Example 351*
Examples 63
Exist 37*
Exists 19
Existed 14
Existing 157
Expand^ 34
Expanded 14
Expanding 16
Expands 7
Expansion 38
Expansive 3
Expect 22*
Expects 4
Expected 142
Expecting 4
Experience 211*
Experiences 172
Experienced 61
Experiencing 24
Explain 51*
Explains 15
Explained 64
Explaining 32
*Explore** 59*
Explores 22
Explored 16
Exploring 28
Face 95*
Faces 7
Faced 33
Facing 22
Fact^ 167
Facts 20
*Factor*120*
Factored 1
Factoring 1
Factors 249
Feel 54*
Feels 3
Felt 37
Feeling 18
Feelings 22
Field^ 173
Fields 22
Figure 26*
Figures 35
Figuring 2
Final^ 65
Finalized 3
Finally 107
*Finance*153*
Financed 17
Finances 2
Financial 208
Financially 8
Financiers 10
Financing 18
Find 111*
Finds 20
Finding 53
Findings 202
Found 233
Focus 194*
Focuses 48
Focused 107
Focusing 43
Focussed 8
Focussing 1
Follow 50*
Follows 50
Followed 51
Following 139
Form 161*
Forms 70
Formed 27
Forming 17
*Frame** 93*
Frames 47
Framed 12
Framing 51
Framework^ 207
Frameworks 49
*Frequency*** 114*
Frequencies 3
Function^ 115
Functional 21
Functionally 1
Functioned 16
Functioning 13
Functions 77

Further 229*
Furthers 1
Furthered 1
Furthering 2
Future 246*
Futures 5
General 221*
Generate^ 35
Generated 37
Generates 9
Generating 23
Get 82*
Gets 9
Got 10
Getting 29
Gotten 2
Give 100*
Gives 22
Gave 16
Giving 47
Given 224
Globe^ 4
Global 145
Globally 10
Globalization 7
Go 73*
Goes 18
Went 21
Going 67
Gone 7
Gonna 1
Goal^ 76
Goals 120
Good 137*
Better 182
Government 272*
Governments 58
Great 60*
Greater 119
Greatest 25
Group 340*
Groups 321
Grouped 3
Grouping 1
Grow 23*
Grows 6
Grew 9
Growing 98
Grown 18
Growth 134*
Health 269*
Help 171*
Helps 21
Helped 29
Helping 25
High 375*
Higher 286
Highest 51
Highlight^ 43
Highlighted 34
Highlighting 14
Highlights 36
Hold 25*
Holds 11
Held 51
Holding 18
Holdings 3
*Household** 183*
Households 102
However 481*
Human 150*
Humans 5
Idea 90*
Ideas 99
Identify^ 127
Identifiable 4
Identification 29
Identified 214
Identifies 15
Identifying 60
Identities 23
Identity 76
Impact^ 272
Impacted 20
Impacting 6
Impacts 131
Implement^ 31
Implementation 86
Implemented 47
Implementing 16
*Implicate^*
Implicated 1
Implication 12
Implications 118
*Importance** 139*
Important 466*
Improve 93*
Improves 2
Improved 74
Improving 34
Include 169*
Includes 68
Included 148
Including 247
Income^ 122
Incomes 11
Increase 224*
Increases 48
Increased 158
Increasing 115
Indicate^ 51
Indicated 61
Indicates 58
Indicating 16
Indication 14
Indications 10
Indicative 9
Indicator 14
Indicators 56

Individual^ 314
Individualised 1
Individuality 2
Individually 5
Individuals 283
Industry 25*
Industries 222
Influence 221*
Influences 44
Influenced 50
Influencing 17
Information 438*
Innovate^ 1
Innovation 151
Innovated 1
Innovates 4
Innovating 4
Innovations 66
Innovative 41
Innovator 5
Innovators 4
Insight^ 30
Insightful 4
Insights 75
Institute^ 46
Instituted 1
Institutes 1
Institution 18
Institutional 131
Institutionalized 7
Institutionally 1
Institutions 115
Interact^ 12
Interacted 4
Interacting 3
Interaction 67
Interactions 57
Interactive 9
Interacts 3
Interest 132*
Interests 66
Interested 43
Interesting 53
Interpret^ 10
Interpretation 43
Interpretations 11
Interpretative 2
Interpreted 34
Interpreting 4
Interpretive 22
Interprets 4
Misinterpreted 2
Reinterpret 2
Reinterpretations 1
Interview 89*
Interviews 144
Interviewed 69
Interviewing 5
*Introduction** 111*
Introductions 4
Invest^ 18
Invested 10
Investing 9
Investment 178
Investments 48
Investor 4
Investors 33
Reinvestment 1
Investigate^ 33
Investigated 39
Investigates 13
Investigating 17
Investigation 37
Investigations 7
Investigators 1
Involve^ 34
Involved 159
Involvement 67
Involves 43
Involving 27
Uninvolved 2
Issue^ 155
Issued 6
Issues 240
Just 144*
Key 214*
Keys 4
Know 92*
Knows 3
Knew 5
Knowing 11
Known 65
Knowledge 327*
Lack 208*
Lacks 9
Lacked 15
Lacking 13
Large 228*
Larger 87
Largest 25
Late 59*
Later 87
Latest 6
Lead 83*
Leads 40
Led 74
Leading 51
Learn 62*
Learns 1
Learned 13
Learning 159
Learnt 1
Less 277*
Level 454*
Levels 186
Leveled /Levelled 1
Leveling /Levelling 2
Like 215*
Likes 1
Liked 7
Liking 2

Likely 221*
Limit 27*
Limits 16
Limited 108
Limiting 9
Link^ 35
Linkage 3
Linkages 20
Linked 65
Linking 15
Links 33
*Literature** 204*
Literatures 1
Live 53*
Lives 21
Lived 15
Living 56
Local 409*
Locals 1
Long 189*
Longer 52
Longest 3
Look 52*
Looks 7
Looked 10
Looking 37
Low 247*
Lows 2
Lower 151
Lowest 18
Lowers 3
Lowered 4
Lowering 5
Main 179*
Maintain^ 37
Maintained 18
Maintaining 20
Maintains 6
Maintenance 32
Major^ 152
Majority 89
Make 212*
Makes 63
Made 190
Making 226
Management 211*
Market 365*
Markets 102
Marketed 3
Marketing 92
Mean 137*
Means 170
Meaning 57
Meant 14
Meanings 27
Measure 76*
Measures 114
Measured 58
Measuring 30
Media^ 201
Meet 73*
Meets 9
Meeting 37
Met 26
Meetings 16
Member 45*
Members 152
Method^ 78
Methodological 26
Methodologies 14
Methodology 51
Methods 121
Model 436*
Models 238
Modeled / Modelled 19
Modeling / Modelling 40
National 277*
Nationals 2
Nature 111*
Need 313*
Needs 183
Needed 108
Needing 7
*Negate^*
Negative 143
Negatively 39
Network^ 118
Networked 1
Networking 31
Networks 113
New 573*
Note 59*
Notes 28
Noted 90
Number 357*
Numbers 29
Numbered 1
Occur^ 55
Occurred 40
Occurrence 30
Occurrences 7
Occurring 19
Occurs 29
Reoccurring 2
Offer^ 60
Offers 42
Offered 55
Offering 20
Offerings 7
Open 89*
Opens 9
Opened 4
Opening 20
Opportunity 50*
Opportunities 72
Order 234*
Orders 3
Ordered 5
Ordering 1

Organization/Organisation 121*
Organizations 129
Organisations 83
Outcome^ 96
Outcomes 246
Overall^ 162
Own 183*
Owns 2
Owned 25
Owning 8
Paper 182*
Papers 21
Part 212*
Parts 24
Participate^ 41
Participant 55
Participants 406
Participated 19
Participates 4
Participating 26
Participation 154
Participatory 2
Particular 232*
Particulars 1
Particularly 145*
Partner^ 45
Partners 43
Partnership 26
Partnerships 34
Past 125*
Pattern 48*
Patterns 81
Patterned 2
People 528*
Peoples 2
Per 130*
Perceive^20
Perceived 106
Perceives 1
Perceiving 2
Perception 62
Perceptions 43
Period^ 155
Periodic 1
Periodically 1
Periods 55
Personal 118*
Personals 166
Perspective^ 132
Perspectives 62
Place 137*
Places 33
Placed 11
Placing 11
Plan 64*
Plans 37
Planned 43
Planning 169
Play 77*
Plays 30
Played 23
Playing 16
Point 177*
Points 65
Pointed 27
Pointing 31
Policy^ 648
Policies 157
Political 228*
Population 224*
Populations 31
Positive^ 173
Positively 37
Possible 172*
Potential^ 236
Potentially 45
Power 189*
Powers 5
Practice/ Practise 176*
Practices 134
Practiced 6
Practicing 27
Practises 1
Practising 1
Predict^ 18
Predictability 3
Predictable 5
Predictably 2
Predicted 31
Predicting 8
Prediction 14
Predictions 11
Predicts 8
Unpredictable 3
Present 150*
Presents 42
Presented 86
Presenting 12
Previous^ 94
Previously 60
Primary^ 74
Primarily 56
Private 122*
Problem 180*
Problems 247
Process^ 242
Processed 12
Processes 141
Processing 51
Produce 62*
Produces 16
Produced 74
Producing 56
Product 205*
Products 99
Production 294*
Productions 2
Professional^ 234

Professionally 1
Professionals 51
Professionalism 5
Program/Programme 235*
Programs 78
Programmes 35
Programing / Programming 2
Project^ 224
Projected 16
Projection 3
Projections 17
Projects 373
Promote^ 60
Promoted 20
Promoter 2
Promoters 1
Promotes 7
Promoting 42
Promotion 47
Promotions 5
Provide 231*
Provides 103
Provided 153
Providing 81
Public 942*
Publics 29
Purpose 98*
Purposes 53
Purposing 1
Quality 428*
Qualities 15
Question 167*
Questions 210
Questioned 17
Questioning 29
Range^105
Ranged 17
Ranges 14
Ranging 14
Rate 107*
Rates 70
Rated 14
Rating 19
Ratings 15
Reason 60*
Reasons 74
Reasoned 2
Reasoning 10
Recent 145*
Reduce 86*
Reduces 17
Reduced 60
Reducing 39
*Reduction** 78*
Reductions 28
Refer 58*
Refers 33
Referred 39
Referring 15
Reflect 49*
Reflects 34
Reflected 28
Reflecting 20
Regard 50*
Regards 12
Regarded 29
Regarding 108
Region^ 97
Regional 103
Regionally 2
Regions 111
Regulate^ 8
Deregulation 3
Regulated 10
Regulates 1
Regulating 6
Regulation 43
Regulations 67
Regulator 1
Regulators 4
Regulatory 17
Relate 29*
Relates 19
Related 385
Relating 20
Relation 87*
Relations 569
Relationship 278*
Relationships 180
Relevant^133
Irrelevance 1
Irrelevant 14
Relevance 38
Rely^ 36
Reliability 38
Reliable 12
Reliably 7
Reliance 21
Reliant 4
Relied 8
Relies 6
Relying 8
Unreliable 1
Report 77*
Reports 51
Reported 119
Reporting 24
Represent 80*
Represents 56
Represented 63
Representing 43
Require^ 70
Required 117
Requirement 17
Requirements 44
Requires 55
Requiring 15
Research^ 774
Researched 4
Researcher 24
Researchers 82

Researches 5
Researching 1
Resource^56
Resourced 4
Resources 272
Resourcing 5
Respond^40
Responded 16
Respondent 93
Respondents 249
Responding 8
Responds 6
Response 107
Responses 106
Responsive 8
Responsiveness 4
Result 146*
Results 363
Resulted 29
Resulting 41
Reveal^ 20
Revealed 67
Revealing 12
Reveals 25
Revelation 2
Review 169*
Reviews 23
Reviewed 24
Reviewing 10
Risk 259*
Risks 71
Risking 1
Role^ 325
Roles 30
Same 347*
*Sample**160*
Samples 19
Sampled 6
Sampling 34
Say 60*
Says 5
Said 65
Saying 12
*Scale** 189*
Scales 27
Scaled 2
Scaling 7
Science 61*
Sciences 42
Section^ 172
Sections 23
Sector^ 272
Sectors 65
See 255*
Saw 14
Seeing 13
Seen 102
Seek^ 31
Seeking 41
Seeks 19
Sought 46
Seem 34*
Seems 56
Seemed 10
Seeming 2
Select^27
Selected 51
Selecting 8
Selection 61
Selections 2
Selective 6
Selectively 2
Selects 1
*Self**176
Service 125*
Services 184
Serviced 2
Servicing 1
Set 212*
Sets 22
Setting 51
Settings 38
Several 212*
Share 80*
Shares 9
Shared 64
Sharing 43
Show 141*
Shows 128
Showed 74
Showing 15
Shown 91
Significant^ 335
Insignificant 6
Significance 144
Significantly 90
Signified 1
Signify 2
Signifying 2
Similar^ 162
Dissimilar 1
Similarities 12
Similarity 11
Similarly 72
Situation 82*
Situations 61
Skill 14*
Skills 162
Skilled 12
Small 173*
Smaller 51
Social 968*
Socials 5
Society 221*
Societies 42
Source^ 139
Sources 138
Sourcing 1

Specific^ 257
Specifically 102
Specification 37
Specifications 7
Specificity 2
Specifics 3
Stage 117*
Stages 54
Staged 1
Staging 1
Standard 63*
Standards 187
Start 53*
Starts 12
Started 26
Starting 31
State 369*
States 157
Stated 83
Stating 11
Statistic^ 3
Statistical 42
Statistically 23
Statistics 50
Strategy^ 124
Strategic 84
Strategies 209
Strategically 13
Strong 133*
Stronger 30
Strongest 8
Structure^112
Restructured 1
Restructuring 5
Structural 63
Structured 47
Structures 50
Structuring 6
Study 749*
Studies 704
Studied 39
Studying 15
*Success*154*
Successes 3
Suggest 141*
Suggests 126
Suggested 83
Suggesting 30
Support 511*
Supports 21
Supported 55
Supporting 60
Survey^ 285
Surveyed 18
Surveying 2
Surveys 49
Sustain^ 8
Sustainable 77
Sustainability 141
Sustained 12
Sustaining 6
Unsustainable 4
System 323*
Systems 285
Table 226*
Tables 9
Take 137*
Takes 35
Took 36
Taking 62
Taken 84
Target^ 55
Targeted 27
Targeting 17
Targets 45
Technology^ 135
Technological 81
Technologically 5
Tend 79*
Tends 18
Tended 17
Tending 1
Term 235*
Terms 222
Termed 5
Test 115*
Tests 19
Tested 31
Testing 18
Theory^ 293
Theoretical 69
Theoretically 23
Theories 88
Theorists 4
Therefore 256*
Think 95*
Thinks 1
Thought 26
Thinking 98
Thoughts 15
Time 655*
Times 185
Timing 23
Total 103*
Totals 1
Tradition^ 18
Traditional 147
Traditionalist 8
Traditionally 17
Traditions 10
Transport^149
Transportation 130
Transported 3
Transporting 1
Turn 73*
Turns 10
Turned 8
Turning 13
Type 132*
Types 168
Typed 1

Understand 108*
Understands 1
Understood 36
Understanding 209
Understandings 31
University 177*
Universities 33
Use 478*
Uses 28
Used 452
Using 325
*User** 57*
Users 191
Utilize/ Utilise^ 11
Utilisation 19
Utilised 6
Utilising 2
Utility 147
Utilities 51
Utilization 2
Utilized 17
Utilizes 2
Utilizing 7
Value 251*
Values 171
Valued 13
Valuing 6
Various 135*
Vary^ 27
Invariably 2
Variability 78
Variable 56
Variables 171
Variance 35
Variant 1
Variants 2
Variation 33
Variations 24
Varied 20
Varies 25
Varying 12
View 95*
Views 76
Viewed 41
Viewing 5
Water 272*
Waters 1
Watering 1
Way 219*
Ways 162
Well 400*
Work 340*
Works 59
Worked 25
Working 109
Workings 5
World 247*
Worlds 2
Year 139*
Years 26