Journal of Teaching Language Skills (JTLS)
35(3), Fall 2016, ISSN: 2008-8191
pp. 65-90

Do We Need Discipline-Specific Academic Word Lists? Linguistics Academic Word List (LAWL)

Raouf Moini *
Assistant Professor
University of Kashan
raoufmoini@yahoo.com

Zahra Islamizadeh
M.A., TEFL
University of Kashan
zahraeslamizade@yahoo.com

Abstract

This corpus-based study aimed at exploring the most frequently-used academic words in linguistics and compare the wordlist with the distribution of high frequency words in Coxhead's Academic Word List (AWL) and West's General Service List (GSL) to examine their coverage within the linguistics corpus. To this end, a corpus of 700 linguistics research articles (LRAC), consisting of approximately 4 million words from four main linguistics sub-disciplines (phonology, morphology, semantics and syntax) was compiled and analyzed based on two criteria; frequency and range. Based on the analysis, a list consisting of 1263 academic word families was produced to provide a useful linguistics academic word list for native and non- native English speakers. Results showed that AWL words account for 10.18 % of the entire LRAC, and GSL words account for 72.48% of the entire LRAC. The findings suggested that of 570 word families in Coxhead's AWL, 381 (66.84%) word families correspond with the word selections criteria which provide 29.88% of the word families in Linguistics Academic Word List (LAWL). Furthermore, 224 word families that were frequently used in linguistic research article corpus (LRAC) were not listed in GSL and AWL. They accounted for 18.51% of the word families in LAWL with coverage of 5.07% over LRAC, and compared with the 2000 GSL, 658 word families were identified. The results have pedagogical implications for linguistics practitioners and EAP practitioners, researchers, and material designers.

Keywords: academic word list, general service list, research articles, English for academic purposes, corpus linguistics, academic writing

Received: 30/07/2016 Accepted: 20/12/2016
* Corresponding author

One predominant language problem of non-native and native learners in academic reading and writing is learning vocabulary (Shaw, 1991), especially if the goal is to achieve a high literacy level in the second language (Cobb & Horst, 2004). Unfamiliarity with a set of special words, i.e., academic vocabularies, is one of the greatest problems in learning the target language. Within this area many researchers have examined academic texts (e.g., Campion & Elley, 1971; Coxhead, 2000, Xue & Nation, 1984) to find their specific features; in most of these studies, academic corpora have been used as an important source of empirical information to study specialized vocabulary in academic texts.

Nation (2001) divided words in academic writing into four categories: High frequency general words, technical words, academic words, and low frequency words. High frequency general words are core words and used very frequently in most language domains (Nation & Hwang, 1995). The 2000 West's GSL represents these low frequency words. The list has been suggested to account for about 80% of the words in any texts but Engels (1968) criticized the 2nd1000 words as General Service List words because of their low coverage and frequency. They have been also criticized for being old as the list was developed more than half a century ago. Technical words, on the other hand, are words used in specialized field and they differ from one discipline to another. They are basically technical terms that are learned in the course of content courses and classes rather than language classes. For identification of these technical words, Chung and Nation (2004) considered the rating scale approach as the "most accurate" for identifying these words.

Farrell (1990) defined academic vocabulary as "formal, context-independent words with a high frequency and/or a wide range of occurrence across scientific disciplines, not usually found in basic general English Academic words, not usually found in basic general English texts, refer to words that account for a relatively high proportion of running words in all academic texts, ccourses, words with high- frequency across scientific disciplines"(p.11). Some researchers pointed out that these words were the most challenging and problematic than technical words for learners, (Shaw, 1991; Thurstun & Candlin, 1998). They argued that these words were not so frequent i.e., they had middle frequency and they occurred across different disciplines (Li & Pemberton, 1994). The

Academic Word List (Coxhead, 2000), is the most widely cited academic word list in the literature that is recommended as a rich and useful source of vocabulary teaching and learning in academic contexts.

The fourth category which is the largest group of words in any field is low frequency words. This group includes rarely used words such as proper names, words that learners rarely need and are technical words for other subject areas (Nation, 2001) which occur in low frequency.

Literature Review

Within EAP, several studies have been conducted to develop AWLs, either general or specialized one, Campion and Elley (1971) and Praninskas (1972), developed word lists based on different texts from different disciplines, Lynn (1973) and Ghaddesy (1979) provided word lists by counting students' annotations. Xue and Nation (1984) published the University Word List (UWL) by combining the four lists published by Campion and Elley (1971), Praninskas (1972), Lynn (1973) and Ghadessy (1979). The UWL contains 808 high frequent non-GSL words that were common in academic texts. This UWL accounted for approximately 8% of the words in typical academic texts (Nation, 1990).

Another academic word list was developed by Coxhead (2000); known as Academic Word List (AWL), which contained 570 word families. In her study, Coxhead conducted a corpus of 3.5 million words from different academic journals and university textbooks in 28 sub-disciplines of 4 main areas: The Arts, Commerce, Law, and Natural science. Each of them contained approximately 875,000 words. She applied three criteria of frequency (occurrence of 100 times in the entire corpus), range (occurrence of at least 10 times or more in each of four main areas and in at least 15 of the 28 sub-disciplines), and specialized occurrence (to be outside the 2000 GSL words). The AWL is divided into 10 sublists based on their frequency of occurrence. Each sublist consists of sixty words, except the 10thsublist which contains thirty words. Coxhead (2000) claimed that 570 word families in the AWL accounted for approximately 10% of the words in her corpus. However, its coverage over the 4 areas was not the same; 9.3% in the Arts, 12% in Commerce, which is the highest, 9.4% in Law and 9.1 % in Science, which is the lowest. By considering a similar size of fiction, Coxhead found that the AWL covered

only 1.47% of it. She argued that AWL was more relevant for learners with an academic purpose. However, Coxhead (2000) argued that AWL was essential for students in different fields, specifically for those in higher education.

Many researchers have recently developed specialized academic words that frequently occur in specialized disciplines. Ward (1999) developed a specific engineering word list. He argued that this word list had coverage of about 95% of engineering texts. In another study, Ward (2009) conducted a wordlist for engineering students at the same university, but they were very poor in English proficiency. He concluded that previously developed wordlist was too long. Lam (2001), developed a computer science wordlist, she found that academic words were semantically different from the same words when they appeared in general texts. In engineering field, Mudraya (2006) built SEEC (Students Engineering English Corpus). Also, she compared the 50 most frequent words in her list with BNC and COBUILT bank of corpus.

Due to the importance of AWL, some researchers have explored the coverage of AWL word in specialized field. Wang, Liang and Ge (2008) provided a medical academic wordlist and they found that of 570 AWL families only 342 word families in MAWL overlapped with Coxhead's (2000) AWL. In applied linguistics, Vongpumivitich, Huang & Chung (2009), provided an applied linguistics wordlist and they found that 475 AWL words, which accounted for 11.17% of their applied linguistics research articles corpus (ALC), coincided with their wordlist. They indicated that although AWL played an important role in applied linguistics academic word list, it was unable to represent field specific words within this discipline. Chen and Ge (2007) by developing a medical academic word list found that AWL words accounted for 10.07% of their medical research articles corpus. They also indicated that there was a marked difference between frequency rank of AWL words in Coxhead's study and their corpus, some of the most frequent words in medical research articles were listed in the 8th and 9^{h}sublists of Coxhead's AWL. Martinez, Beck & Panza (2009) developed agriculture academic word list. They found AWL words covered only 9.06% of their corpus, which was lower than what was found by Coxhead (2000) and Hyland and Tse (2007).

In another study, Li and Qian (2010) explored the existence of AWL in a 6.5 million words corpus of financial texts. They found that only 162 (28.42%) of AWL word families occurred frequently in their corpus. In a corpus of 3.3 million words of different academic disciplines and genres, Hyland and Tse (2007) found that of 570 AWL word family 541 families occurred in all sub-corpora which covered 10.6 % of the words in the corpus. Their finding also suggested that AWL word families did not appeare with the same frequency and meaning across different disciplines, for instance, AWL words was most useful to computer science with a coverage of 16%, and least useful to biology students because of its low coverage (6.2%). Mungra and Ganziani (2013) conducted clinical case histories wordlist. They found 498 word families of Coxhead's AWL occurred in their corpus which accounted for 13.6% of the corpus.

Yang (2014) created nursing academic word list, the word list consisted of 676 word families, and he found 398 word families coincided with 570 AWL word families, which had coverage of 8.93% of the corpus. Recently Liu and Han (2015) compiled and analyzed a corpus of 200 texts containing 862,242 tokens and 26,612 types in the field of environmental science, their finding suggested that of 570 AWL word families, 569 appeared in the corpus, which covered 12.82% of the running words. However, they argued that due to two reasons of the narrow coverage of some word families and the shortage of frequently used environmental academic words, AWL was not entirely useful for students in this specialized field.

Some researchers (Billuroglu & Neufeld, 2005, 2007; Martinez et al., 2009; Paquot, 2007; Ward, 2009) have criticized the exclusion of 2000 GSL word families from the list because these words sometimes have different collocations, uses and meanings in specialized corpus. For example, in Chemistry Valipoori and Nassaji (2013) developed Chemistry academic wordlist. In their study, they found 1400 word families, out of them 327 word families overlapped with Coxhead's AWL, which coved 9.60% of the Chemistry corpus.

All the studies mentioned above argued that general academic word lists like Coxhead's AWL, was not suitable for all students across different fields and majors and advised developing discipline-specific word lists and each discipline to establish its own academic wordlist. An academic

wordlist exclusively for linguistics students can be taught and directly studied in the same way as the words from GSL. Therefore, this study aimed to establish a Linguistics Academic Word List (LAWL) of the most frequently-used English academic words across different sub-disciplines in linguistics. To guide the present study, a linguistics research articles corpus (LRAC), a collection of journals in the field of linguistics was established to address the following questions:

1. What are the most frequent English academic words used in linguistics research articles?
2. Does the new word list -as compared with AWL and GSL- have a better coverage of linguistics research articles corpus (LARC)?

Method

Data Collection

To this end, a linguistics research articles corpus (LRAC) containing 4 million words in four main sub-disciplines: phonology, morphology, syntax, and semantic, was specifically compiled. The relevant research articles were adopted from11 international journals in the field of linguistics, namely, *Lingua, Natural languages and Linguistics Theory*, *Natural Language Semantics, Language, Phonetics, Syntax, Morphology*, *Journal of Asian Linguistics, Stadia Linguistica, Linguistics inquiry* and *The Modern Language Journal* published between 2010-2015. Developing and compiling the corpus included two steps. Firstly, following Swales (1990), only articles focusing on empirical studies and those which were written in identifiable introduction, method, results, and discussion were included in LRAC. Secondly, the research articles had to be balanced in terms of their length. Articles falling out of 6000-10000 words range were excluded. Also, the total number of words from longer and shorter articles were equal. Finally, 700 research articles were randomly selected. For standardization of the corpus all the tables, bibliographies, acknowledgments, figures, charts, footnotes, references and appendices were removed to exclude the irrelevant factors in the final analysis run by the Range software.

Computer Programs. Two computer software programs, namely, ABBYY FineReader and RANGE were used. ABBYY Finereader is an Optimal Character Recognition (OCR) document processing software. It was used for converting pdf files, documents, and image files into editable formats. Since the first step for the study after gathering the corpus was inverting all the pdfs into plain text in order to be readable by another computer program, namely, RANGE, which is downloadable at *http://www.vuw.ac.nz/lals/staff/Paul_Nation*. RANGE can count all the words in the corpus and sort them out into 4 categories: 1) those found in the first 1000 most frequent GSL, 2) those found in the second 1000 most frequent GSL 3), those found in Coxhead's AWL, and finally 4) those were not found in 2000 GSL and 570 AWL. The software can produce the total frequency of each word family, also frequency of each word in each family. It also represents the frequency of each word and word family in each sub-discipline (phonology, morphology, syntax, and semantics).

Data Analysis and Processing

This study aimed at developing a list of linguistics English academic words that were most frequently used in linguistics RAs. To this aim the corpus was quantitatively analyzed, using RANGE program. Regarding word selections criteria as mentioned earlier, this study followed Coxhead, and the two criteria of frequency, range were used. Coxhead (2000) ranked range as the first criterion and frequency as the second. So, following Coxhead's (2000) study, she selected word families which occurred at least100 times or above it in her 3.5 million corpus of academic texts, that is 28.5 times per 1 million. The corpus used in this study consisted of 4 million words which is 0.5 million words larger than Coxhead's corpus. So, in terms of frequency, for a word to be selected as frequent in LRAC, it had to occur at least 114 times or more in the entire corpus [(4×28.5) = 114]. Regarding range, only word families which appeared in all sub-disciplines and at least 10 times or more in each sub-discipline, were selected. Most studies on academic words used word families, so words in LRAC were identified as word family which was defined by Bauer and Nation (1993) as the basic word plus its inflected forms and transparent derivations, for example:

Consist*: consisted, consistency, consistent, consistently, consisting, consists, inconsistencies, inconsistency, inconsistent.*

As a learner's knowledge of affixation develops, the size of the word family increases (Bauer & Nation; 1993). The important principle is that, once the base word or even the derived word is known, the recognition of other members of the family requires little or no extra effort (Coxhead, 2000).

In this study, only content words were selected, consequently, all the function words such as pronouns, auxiliaries, articles, numbers, conjunctions, determiners, prepositions, quantifiers were excluded as they were too general. All abbreviations such as (L1, PP, NP, CF, etc.), which occurred very frequently in LRAC were excluded from the study. For identification of technical words, Chung and Nation's (2003) rating scale which is very practical and frequently used by other researchers was adapted. Accordingly, words considered to be unknown or not used frequently in other sub-disciplines, were regarded as too technical. They were all excluded. The final list – Linguistics English Academic Wordlist- was compared with the Coxhead Academic Word List (AWL) and General Service List (GSL).

Results

One of the mains aim of this study was finding the most frequent words in the field of linguistics. Following Coxhead's study, as it was mentioned earlier only word families which had a frequency of at least 114 in the entire corpus also occurring at least 10 times in each sub-discipline, were selected. Based on the word selection criteria, a total of 1263-word families formed the LAWL (See Appendix A which presents the alphabetic list of 1263 linguistic academic words -LAWL). Out of 570 AWL word families, 568 AWL word families occurred in the entire corpus, the total frequency of these words counts for 411,445, which cover 10.18% of the whole LRAC, which is 0.18% higher than Coxhead's AWL. Having applied the two criteria of frequency and range out of 568 word families, 381 (66.84%) word families met the set word selection criteria and were considered as frequently used AWL items in LRAC. (See Appendix A which presents the alphabetic list of 1263 linguistic academic words -LAWL).

Within the first 1000 GSL words 995 word families occurred in LRAC which covered 68.21% of the corpus but out of them 504 word families met the set criteria. Within the second 1000 GSL word families, 918 word families with a coverage of 4.27% appeared in LRAC and 82 word families did not occur in the LRAC at all. However, only 154 word families met the set frequency criteria and 764 word families occurred infrequently, these word families occurred less than 114 in the corpus or less than 10 times in each sub-discipline. Totally, 658 out of 2000 GSL word families were used frequently in LRAC. The total number of GSL word families occurred in the corpus was 1913 and 87 GSL word families did not appear in LRAC at all. Altogether, the first and second words of GSL with AWL words accounted for 82.66 % of the LRAC; Table 1 presents the coverage and number of word families of 1st, 2nd and AWL word lists over the entire LRAC.

Table 1

Results of Coverage of First and Second GSL and AWL in LRAC Corpus

Word list	Coverage of LRAC (%)	Number of word families
First 1000 words of GSL	68.21%	995
Second 1000 words of GSL	4.27%	918
AWL	10.18%	568
Total	**82.66**	**2,481**

As shown in table 1, the first 1000 words (68.21%), the second 1000 words (4.27%), and AWL (10.18%) covered 82.66% of LRAC. This leaves 17.34% of the words that belong to non-GSL words. The result also showed the coverage of AWL over the corpus is 2.38% times as big as the second 1000 words of GSL, showing that many AWL words occurred with a higher frequency than the second 1000 GSL, which is thought to include more frequent words than AWL.

The first 100 most frequent 1stGSL, 2ndGSL and AWL words that occurred frequently are shown in Table 2.

Table 2

Result of Frequency Comparison of First 100most Frequent Words of 1st and 2nd 2000 GSL and AWL.

No.	Word family	No.	Word family	No.	Word family	No.	Word family
1	Case	26	**Clause**	51.	Movement	*76*	*Model*
2	**structure**	27	Effect	52	Agreement	**77**	**Complex**
3	See	28.	Set	53	Reading	78	Based
4	Same	29	Fact	54	Way	*79*	*Plural*
5	Language	30.	Possible	55	**Tense**	**80**	**Role**
6	Different	31.	Result	56	Relative	81	Relation
7	**Section**	32.	Account	57	Show	82	Person
8	Like	**33.**	**Contrast**	**58**	**Stress**	83	General
9	Speak	34	Head	**59**	**Scope**	**84**	**Significant**
10	*Verb*	35	Type	*60*	*Noun*	*85*	*Root*
11	**Data**	**36**	**Feature**	**61**	**Duration**	**86**	**Hypothesis**
12	Position	37	Found	**62**	**Theory**	**87**	**Instance**
13	Word	38	Object	63	Difference	**88**	**Domain**
14	Example	39	Speaker	64	Main	89	System
15	**Context**	40	Shown	65	Particular	**90**	**Final**
16	Time	41	Forms	66	Event	91	Terms
17	**Focus**	42	Study	**67**	**Target**	**92**	**Assume**
18	*Vowel*	43	Question	68	Paper	**93**	**Specific**
19	Order	44	Even	*69*	*Pattern*	**94**	**Initial**
20	Present	45	Number	70	Past	95	Situation
21	Given	**46.**	**Similar**	71	Point	96	True
22	Following	*47*	*Sentence*	72	Part	97	Consider
23	Speech	48.	Properties	**73**	**Relevant**	98	Experiment
24	Used	49.	Meaning	**74**	**Approach**	**99**	**Occur**
25	Argument	50.	Difference	**75**	**Function**	*100*	*Grammar*

Note: AWL words are bold, 1st GSL are in normal font and the 2nd GSL are italicized.

As table 2 shows, among the 100 most frequent words, 62 word families (62%) are among the first 2000 of GSL 29 word families (29%) are AWL word families, and 9 (9%) are among the second 2000 GSL word families. The number of AWL in this table shows that among the most frequent 100 words, the AWL words were 3 times as big as the second 2000 GSL. As it was mentioned before, the results suggest that 2000 GSL words and AWL words produced the coverage of 82.66% in LRAC. In

other words, it failed to account for 17.34% of the tokens in LRAC. The findings suggest that there might be other words outside GSL and AWL which occurred with high frequency in LRAC.

To determine these high frequency words, the same criteria of frequency (occurring 114 or more than it) and range (occurring at least 10 times in each four sub-disciplines) used for identification of frequent GSL/ AWL words in the LRAC were used. Table 3 presents coverage and frequency of these non-GSL non-AWL in **the** LRAC.

Table3

Frequency Analysis of Non- AWL and Non GSL Words in LAWL

Times of occurrence	Number of words	Frequency count	%LRAC
≥2000	11	32,822	0.81%
1999- 1000	30	56,781	1.40%
999- 700	20	28,210	0.7%
699- 400	43	36,323	0.9%
399- 114	120	50,909	1.26%
Total	**224**	**205,045**	**5.07%**

After applying the criterion of the study, a list of 300 words families was obtained. Based on experts' views- two experienced linguistic professors- 76 word families out of 300 word families were evaluated as too technical and were excluded from the list. As shown in table 3, frequency count of these non-GSL/ non-AWL showed that these words were counted for 701,148 tokens in LRAC which counted for 17.34% of the entire corpus. The first 2 groups of non-GSL non-AWL word, included 41 word families and a frequency count of 89,603 account for 2.21 % of the corpus, the last three group with 183 word families and a frequency count of 115,442 cover 2.86% of entire corpus. Knowing this non-GSL/non-AWL is essential for linguistics students. Overall, the results produced a total of 224 word families which account for 205,045 words in the corpus with a coverage of 5.07 % of LRAC.

In another phase of data analysis, the corpus was also analyzed for identifying the non-GSL non-AWL that frequently occurred in the LARC. Table 4 shows the 40 most frequent non-GSL non-AWL words in **the** LARC.

Table 4

The First 40 Most Frequent Non- AWL/ Non- GSL Words in LRAC

No.	Words	No.	Words
1	SYNTACTIC	21	NOMINAL
2	LEXICAL	22	CLITIC
3	SEMANTIC	23	LINGUISTIC
4	PHONOLOGICAL	24	SECT
5	PROSODIC	25	PITCH
6	MORPHOLOGY	26	MORPHEME
7	PREDICATE	27	MODAL
8	PHRASE	28	GRAMMATICAL
9	ACOUSTIC	29	CUES
10	SUFFIX	30	STIMULI
11	SYLLABLE	31	ONSET
12	ADJECTIVES	32	OVERT
13	TEMPORAL	33	UTTERANCE
14	DISCOURSE	34	TONE
15	VERBAL	35	DEFAULT
16	SYNTAX	36	PHRASE
17	EMBEDDED	37	ACCENT
18	PHONETIC	38	PRONOUN
19	NULL	39	SINGULAR
20	PARTICIPLES	40	REFERENCE

It is worth mentioning that, these 40 non-GSL non-AWL words accounted for 73,989 of the tokens with coverage of 1.83% of the entire corpus.

To find out how AWL words are distributed in LAWL, their coverage and distribution was further analyzed. Table 5 presents the frequency and distribution of non-GSL non-AWL words in LAWL.

Table 5

Coverage of the 381 AWL Families That Occurred in LAWL

AWL sublists	1	2	3	4	5	6	7	8	9	10	**Total**
1-100	26	17	9	12	10	7	6	7	5	0	**99**
101-200	13	11	14	19	9	11	5	8	4	6	**100**
201-300	8	14	13	10	10	8	14	9	12	4	**102**
301-381	3	4	11	5	12	10	11	15	5	4	**80**
Total	**50**	**46**	**47**	**46**	**41**	**36**	**36**	**39**	**26**	**14**	**381**

As shown in table 5, only 381 (66.84%) of Coxhead's academic word families appeared in LARC. Ten word families of sublist1 and 14 words from sublist2 did not appear LAWL. More details of this phase of the analysis as indicated in table 5 revealed that most of words common to both LAWL and LAWL occurred in sublists 1, 2, 3, 4 and 5. They included 230 word families with coverage of (60.37%). The remaining 151 word families appeared in sublists 6, 7, 8, 9 and 10 that accounted for (39.63%) of the words in LAWL. As table 5 shows, among the first 100 word families, 43 (11.28%) word families belong to Coxhead's sublist 1 and 2, and no word families from sublist 10 of AWL appeared within the first 100 words of LAWL. The word families from sublists 1,2,3,4 and 5 greatly exceeded those of sublists 6 to 10.

It is interesting to note that the low frequency words in Coxhead's AWL, which are mostly included in sublist 9 and sublist 10, are also among the least frequent items in LAWL. However, some word families such as *tense*, *duration*, *complement*, and *passive*, which are in sublists of 8 and 9 of Coxhead's AWL were among the most frequent word families in LAWL.

Table 6 presents the first 60 most frequent words in LAWL along with their rank in AWL.

Table 6

The First 60 Most Frequent AWL Word Families in LAWL Compared With Those in AWL

No.	Headword	Sub	No.	Headword	Sub
1	**STRUCTURE**	**1**	31	WHEREAS	5
2	**SECTION**	1	32	CATEGORY	2
3	**DATA**	**1**	**33**	**PREVIOUS**	2
4	**CONTEXT**	**1**	**34**	**RESEARCH**	**1**
5	FOCUS	*2*	**35**	**PROCESS**	**1**
6	CLAUSE	5	**36**	**AVAILABLE**	**1**
7	CONTRAST	4	**37**	**INDIVIDUAL**	**1**
8	FEATURE	*2*	38	ELEMENT	2
9	**SIMILAR**	**1**	39	TOPIC	7
10	TENSE	8	40	HENCE	4
11	STRESS	**4**	41	TASK	3
12	SCOPE	6	*42*	ALTERNATIVE	3
13	DURATION	9	43	RANGE	2

No.	Headword	Sub	No.	Headword	Sub
14	**THEORY**	**1**	44	DISTINCT	2
15	TARGET	5	**45**	**PRINCIPLE**	**1**
16	RELEVANT	2	46	GENDER	6
17	**APPROACH**	**1**	47	ASPECT	2
18	**FUNCTION**	**1**	48	INTERNAL	4
19	COMPLEX	2	**49**	ABSTRACT	6
20	**ROLE**	**1**	**50**	**ISSUE**	**1**
21	**SIGNIFICANT**	**1**	51	STATUS	4
22	HYPOTHESIS	*4*	52	FURTHERMORE	6
23	INSTANCE	**3**	53	EXTERNAL	5
24	DOMAIN	6	**54**	**INVOLVE**	**1**
25	**FINAL**	**2**	**55**	EMPIRICAL	7
26	**ASSUME**	**1**	**56**	NOTION	5
27	**SPECIFIC**	**1**	**57**	VIA	8
28	INITIAL	3	58	PHASE	4
29	**OCCUR**	**1**	59	INPUT	6
30	COMPLEMENT	8	60	PHENOMENON	7

Note: Word families common to LAWL and Coxhead's sublist1 of AWL sublist1are in bold face.

As shown in table 6, what is very clear about these first 60 most frequent words in LAWL is that they belong to almost all AWL sublists. They are not limited to one or two sublists which indicate the necessity of developing discipline-specific word lists. Based on such findings one can argue for different saliency, rank, and frequency of academic words across disciplines. For example, the words *clause* and *contrast* belong to sublists 5 and 4 in AWL while they are among the first ten most frequent words in LAWL. Or in the case of the word *duration*, it is among the least frequent words in AWL (sublist 9), while in LAWL it is one of the most frequent words.

Discussion and Conclusions

In this study, a 4 million words corpus of linguistics research articles (LRAC) was compiled and compared with Coxhead's (2000) AWL and West's (1953) GSL word list. The study specifically focused on the academic words used in one specific field, that is, linguistics. The LRAC established for this study is representative, balanced, and genre-specific

(only research articles were included). LAWL targets only linguistics students.

Results showed that AWL accounted for 10.18% of the LRAC. Out of 570 word families in AWL, only 381 (66.84%) were found to be frequently used in the linguistics research articles. This AWL's coverage over LRAC was 1.08% higher than Coxhead's (2000) science sub-corpus, and about 0.88% higher than Hyland and Tse's (2007) science sub-corpus, Coxhead's (2000) Art sub-corpus Coxhead's (2000) Law corpus and 1.12% higher than Martinez et al. (2009) AgroCorpus. The coverage of AWL in LRAC was slightly lower than in engineering and social science sub-corpora in Hyland and Tse (2007) study, suggesting that specific fields have different coverage of AWL.As for GSL words only 658 were used frequently in LARC. In general, 2000 GSL words accounted for 72.48% of the entire LRAC which was higher than Coxhead's (2000) science sub-corpus, Hyland and Tse' s 2007) science sub-corpus, Martinez et al. (2009) AgroCorpus, Valipoori and Nassaji's (2013) chemistry corpus, and slightly lower than Coxhead's (2000) Art sub-corpus, Coxhead's (2000) Commerce corpus, Hyland and Tse's (2007) Engineering sub-corpus and Hyland and Tse's (2007) social Sciences sub-corpus. The less difference in GSL coverage is between LRAC and Hyland and Tse's (2007) engineering sub-corpus, which is 0.82 % is higher in Hyland and Tse's (2007) Engineering sub-corpus.

Collectively, AWL and GSL words accounted for 82.66% of the entire LRAC which was 7.34% less than 90% coverage of 2000 GSL and AWL words suggested by Coxhead and Nation (2001). In addition, AWL coverage (10.18%) was 2.38% times higher than the 2nd 2000 GSL (4.27%). This finding shows that it is not necessary to learn the 2nd 2000 GSL words before AWL words.

As mentioned by Nation (2016)**,** word lists can be used for multiple pedagogical purposes, ranging from course design, and material development for language learning and teaching to designing graded reading programs, analysis of vocabulary load of texts, and language vocabulary test development. One main challenge of course designers, material developers and test developers is vocabulary. With such a list, course designers can decide on long term and mid-term vocabulary learning goals. In the light of findings of this study, a discipline-specific

word lists such as LAWL course designers can plan long term goals and develop graded listening and reading material most appropriate to the needs and proficiency levels of linguistics students. Drawing on the same material, test developers can decide on realistic goals for test content and develop tests that are realistic to the needs and interests of test takers and other practitioners of the filed.

As noted earlier, this study led to the development of an academic word list LAWL that can be useful for linguistics students and teachers. The findings questioned the existence of a substantial, cross disciplinary academic vocabulary as indicated by earlier research across different disciplines (Hyland & Tse, 2007; Martinez et al., 2009; Valipoori & Nassaji, 2013). LAWL, can serve as a reference for developing teaching materials, especially for designing textbooks for EAP and ELP. This list can also help material designers to design relevant teaching materials that reflect the target language needs of students. Since the list provides information about the frequency of occurrence of these words in RAs, the material designer can use this list for sequencing of teaching content. This study resulted in developing a linguistics-specific word list which represents the specific vocabulary needs of the linguistics students by providing them with a list of high frequency words in this discipline, because these students have limited time for learning English, and have limited exposure to academic language, it would help these students to make such a listing a focus of explicit learning. With such tools, all practitioners of linguistics will have more confidence in the material they teach, learn, and test. It is worth mentioning that, in the context of the present study, word was defined as word family, which is useful for students because as Baur and Nation (1993) acknowledged, knowledge of a base word can facilitate understanding of its derived or inflected forms of words.

The linguistic academic word list (LAWL) developed in this study was not concerned with meaning and function of the words. Words serve different meanings and functions in different disciplines. It is advisable that teachers consider these two dimensions in their practice. Another consideration for teachers and researchers might be looking at the collocation of these words and find out how these words collocate with other words. The LAWL was based on two main criteria; frequency and

range which might not be the only important criteria. Further research may need to consider ease or difficulty of learning, necessity, usefulness in the development of using word lists. LAWL was based on research articles. Research articles are one genre in the domain of linguistics and academia. The question remains whether research articles are the best representative genre of the domain or not. Inclusions of other academic generes may shed some light on the findings. In the last decade, some more disciplines were the focus of word list studies; medicine (Wang et al., 2008), agriculture (Martinez, et al., 2009), applied linguistics (Khani & Tazik, 2013), chemistry (Valipoor & Nassaji, 2013) environmental sciences (Liu & Han, 2015), and nursing (Yang, 2015). Further research is needed to cover other disciplines and develop more discipline-specific word lists.

References

Baur, L. & Nation, I. S. P. (1993). Word families. *International journal of lexicography, 6*(3),1-27.

Billuroglu, A., & Neufeld, S. (2005). The bare necessities in lexis: A new perspective in vocabulary profiling. Retrieved April 20, 2015, from http://lextutor.ca/vp/BNL_Rational.doc on December 13, 2008.

Billuroglu, A., & Neufeld, S. (2007). BNL 2709: *The essence of English (4th ed).* Nicosia: Rustem Kitabevi.

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

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*, 502-514.

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

Chung, T., & Nation, I. S. P. (2004). Identifying technical vocabulary, *System, 32,* 251-263.

Cobb, T., & Horst, M. (2004). Is there room for an Academic Word List in French? In P. Boggards, & B. Laufer (Eds.), *Vocabulary in a second language* (pp.15-38). Amsterdam: John Benjamins.

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

Engels, L., K. (1968). The fallacy of word counts. *International Review of Applied Linguistics, 6,* 213-231.

Farrell, P. (1990). A lexical analysis of the English of electronics and a study of semi- technical vocabulary. *CLCS Occasional Paper No.25 Trinity College.*

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

Heatly,A., Nation, I. S. P., Coxhead, A. (2002). Range and Frequency Programs.
Retrieved from: http://www.vuw.ac.nz/lals/staff/Paul_Nation.

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

Khani, R., & Tazik, K. (2013). Towards the development of an academic word list for applied linguistics research articles. *RELC journal*. *44*(2), 195-214.

Lam, J (2001). A study of semi-technical vocabulary in computer science texts, with special reference to ESP teaching and lexicography. *Research reports, Vol.3.* Language Center. Hong Kong University of Science & Technology.

Li, Y., Qian, D.D. (2010). Profiling the academic word list (AWL) in a financial corpus. *System 38,* 402-411.

Li, S.-L., & Pemberton, R. (1994). An investigation of students' knowledge of academic and Sub-technical vocabulary. In L. Flowerdew & A. K. K.Tong (Eds.), *Entering text* (pp. 183-196).
Hong Kong: The Hong Kong University of Science and Technology.

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

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

Martinez, I. A. Beck, S., & Panza, C.B (2009). Academic vocabulary in Agriculture: A corpus-based study. *English for specific purposes, 28,* 183-198.

Mudraya, O. (2006). Engineering English. A lexical frequency instruction model. *English for Specific Purposes, 25,* 235-256.

Mungra, P. & Canziani, T. (2013). Lexicographic studies in medicine: Academic Word List for clinical case. *Lberica*, *25,* 39-62.

Nation, I., S., P. (1990). *Teaching and learning vocabulary*. Boston, MA: Heinle & Heinle Publishers.

Nation, P. & Hwang, K. (1995). Where would general service vocabulary stop and special purposes vocabulary begin*? System 23*(1), 35-41.

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

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

Paquot, M. (2007). Towards a productively-oriented academic wordlist. In J. Walinski, K. Kredens, & S. Gozdz-Roszkowski (Eds.), *PALC Proceedings* (pp. 127- 140). Frankfurt: Peter Lang.

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

Shaw, P. (1991). Science research students' composing process. *English for Specific Purposes, 10,* 189-206.

Swales,J. M. (1990). *Genre analysis*. Cambridge: Cambridge University Press.

Thurstun, J., & Candlin, C. N. (1998). Concording and the teaching of the vocabulary academic English. *English for Specific Purposes, 17*(3), 267-280.

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

Vongpumivitich, V., Huang, J., & Chung, Y. (2008). Frequency analysis of the words in the Academic Word List (AWL) and non-AWL content words in applied linguistics papers. *English for Specific Purposes*, *28*(1), 33-41.

Wang, J., Liang, S., & Ge, G. (2008). Establishment of a medical Academic wordlist. *Englishfor Specific Purposes,* 27(4), 442–458.

Ward, J. (1999). How large a vocabulary do EAP engineering students need? *Reading in a foreign language, 12*(2), 309-324.

Ward, J. & Chuenjundaeng, J. (2009). Suffix knowledge: Acquisition and applications. *System, 37*(3), 461-469.

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

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.

Appendix A. An alphabetical ordering of 1263 Linguistics Academic Word list (LAWL)

Letter	Word family
A	ABLE, *ABSENCE*, *ABSENT*, *ABSOLUTE*, ABSOLUTIVE, **ABSTRACT** , ACCENT, ACCEPT, **ACCESS,** *ACCIDENT*, **ACCOMMODATE** , **ACCOMPANY,** ACCORD, ACCOUNT , **ACCURATE,** ACCUSATIVE, **ACHIEVE, ACKNOLEDGE,** ACOUSTIC, **ACQUIRE,** ACT, ACTIVATION, ACTIVE, ACTUAL, **ADAPT,** ADDITIONALLY, ADRESS, **ADEQUATE, ADJACENT,** ADJECTIVES, ADMIT, ADOPT, **ADULT,** ADVANCE, ADVANTAGE, ADVERBIAL, **AFFECT,** AFFIX, AGAIN, AGE, AGENT, GREE, *AIM*, ALGORITHM, ALIGNMENT, **ALTER**, ALTERNATION, **ALTERNATIVE**, **AMBIGUOUS**, **ANALOGY, ANALYSE**, ANSWER, *APART*, **APPARENT**,APPEAR, **APPEND**, **APPROACH**, **APPROPRIATE**, **APPROXIMATE**, **ARBITRARY** ARCHITECTURE, **AREA**, ARGUABLY, *ARGUE*, ARISE, ARRIVE, ARTICLE, ARTICULATION, *ASIDE*, ASK, **ASPECT, ASSESS, SSIGN,** ASSOCIATE, **ASSUME**, ASYMMETRY, **ATTACH,** ATTEMPT, *ATTEND*, *ATTENTION*, ATTESTED, **ATTITUDE**, *ATTRACT*, **ATTRIBUTE**, AUDITORY, **AUTHOR AUTOMATE, AVAILABLE**, AVERAGE, *AVOID*, **AWARE**.
B	BACKGROUND, *BAKE*, BAR, BASE, BASELINE, BEAR, BECAUSE, BECOME, BEGIN, *BEHAVE*, *BEHAVIOUR*, BELIEVE, BELONG, **BENEFIT,** BEST, **BIAS,** BIG, BILINGUAL, BINARY, *BIT*, BLACK, *BLOCK*, BLUE, BOOK, *BORROW*, *BOTTOM*, *BOUND*, *BOUNDARY*, BOX, BOY, BRNCH, BREAK, BRIGIN, **BRIEF.** BROAD, BUILD, BUY.
C	*CACUATE*, CALL, CANDIDATE, CANONICAL, **CAPABLE, CAPACITY,** CAPTURE, CARE, CARRY, CASE, CATEGORICAL, **CATEGORY,** CAUSATIVE, CAUSE, CENTRE, *CENTURY*, CERTAIN, *CHAIN*, **CHALLENGE,** CHANGE, **CHAPTER,** *CHECK*, CHOOSE, CHARACTER, **CIRCUMSTANCE, CITE,** CLAIM, **CLARIFY,** CLASS, **CLASSIC, CLAUSE,** CLEAR, CLITIC, CLOSE, CLUSTERS, **CODE,** COGNITIVE, **COHERENT, COINCIDE,** *COLLECT*, COLOUR, COMBINE, COME, **COMMENT,** COMMON, **COMMUNICATE, COMMUNITY,** *COMPARE*, **COMPATIBLE,** *COMPETITION*, **COMPLEMENT**, COMPLETE, **COMPLEX**, *COMPLICATE*, **COMPONENT,** *COMPOSE*, **COMPOUND**, COMPREHENSION, **COMPREHENSIVE, COMPRISE, COMPUTE, CONCENTRATE, CONCEPT,** CONCERN, **CONCLUDE ,** CONCRETE, CONDITION, CONDITIONED, **CONDUCT,** CONFIGURATION, **CONFIRM, CONFLICT,**

CONFORM, *CONFUSE*, CONJUGATION, CONJUNCTION, *CONNECT*, **CONSEQUENT,** CONSIDER, **CONSIDERABLE, CONSIST, CONSTANT, CONSTITUTE, CONSTRAIN, CONSTRUCT, CONSULT, CONTACT,** CONTAIN, **CONTEMPORARY,** CONTENT, **CONTEXT,** CONTEXTUALLY, CONTINUE, CONTOUR, **CONTRADICT, CONTRARY, CONTRAST, CONTRIBUTE,** CONTROL, **CONTROVERSY,** *CONVERSATION*, CONVERSATIONAL, **CONVERSE, CONVERT, CONVEVE,** CONVEY, **CONVINCE,** CO-OCCUR, **COORDINATE,** *COPMETE*, **CORE,** CORPORA, *CORRECT*, CORRELATION, **CORRESPOND,** COUNT, COUNTERPART, **COUPLE, CREATE, CRITERIA ,** *CRITIC*, CROSS, CROSS-LINGUISTIC, **CRUCIAL,** CUES, **CULTURE,** CURRENT, CUT, **CYCLE.**

D **DATA** DATE, DAY, DECIDE, **DEBATE,** DECLARATIVE, DECLARATIVES, *DECREASE*, DEEP, DEFAULT, **DEFINE,** DEGREE, *DELAY*, DELETION, DEMAND, **DEMONSTRATE, DENOTE, DENY,** DEPEND, DEPENDENCY, DERIVATIONAL, **DERIVE,** DESCRIBE, **DESIGN,** DESIRE, **DESPITE,** DETAIL, **DETECT,** DETERMINE, DEVELOP, **DEVIATE, DEVICE,** DEVIDE, DIACHRONIC, DIAGNOSTIC, DIALECTS, DIFFERENCE, DIFFERENTIAL, **DIFFERENTIATE,** DIFFICULT, **DIMENSION,** DIRECT, DISCOURSE, DISCOVER, *DISCUSS*, **DISPLACE, DISPLAY,** DISTANCE, **DISTINCT,** DISTINGUISH, **DISTRIBUTE, DIVERSE, DOCUMENT, DOMAIN, DOMINATE,** DOOR, *DOUBLE*, DOUBT, DRAW, DRESS, DRIVE, DROP DUAL, DUE, **DURATION,** *DURING*, **DYNAMIC.**

E EARLY, EAST, EASY, EAT, **ECONOMY,** *EDGE*, EFFECT, **ELEMENT,** ELICITED, **ELIMINATE,** ELSE, EMBEDDED, **EMERGE, EMPHASIS, EMPIRICAL,** *EMPTY*, **ENABLE,** ENCODED, **ENCOUNTER,** END, ENOUGH, **ENSURE,** ENTAILS, ENTER, *ENTIRE*, **ENVIRONMENT**, EQUAL, **EQUIVALENT**, ERGATIVE, **ERROR,** *ESPECIAL*, *ESSENTIAL*, **ESTABLISH, ESTIMATE, EVALUATE,** EVEN, EVENT, **EVENTUAL,** EVER, EVIDENT, **EVOLVE.** *EXACT*, *EXAMINING*, EXAMPLE, EXCEPT, **EXCLUDE,** EXEMPLIFIED, **EXHIBIT, EXPAND,** EXPERIMENT, EXPLAIN, **EXPLICIT, EXPLOIT,** EXPLORE, **EXPOSE,** EXTEND, **EXTERNAL,** *EXTRA*, **EXTRACT,** *EXTREME*, EYE.

F FACE, **FACILITATE,** FACT, **FACTOR,** FAIR, FAITH, FAITHFULNESS, FALL *FALSE*, FAMILIAR, FAMILY, FAR, *FASHION*, FAST, FAVOR, **FEATURE,** FEEL, *FEMALE*, FEW, FIELD, FIGURE, FILL, FILLER, **FINAL,** FIND, FINE, FIT, FIX, **FLEXIBLE,** *FLOAT*, FOCI, **FOCUS,** FOLLOW, *FOOT*, FOOTNOTE, FORCE, FORM, *FORMAL*, **FORMULA,** *FORWARD*, *FRAME*, **FRAMEWORK,**

FREE, FREQUENCIES, *FREQUENT*, FRIEND, FRONT, FRONTED, FULL, **FUNCTION, FUNDAMENTAL, FURTHERMORE,** FUSION, FUTURE.

G *GAP*, **GENDER,** GENERA, GENERALIZATION, **GENERATE, GENERATION,** GENERATIVE, GENUINE, GIRL, GIVE, **GLOB,** GO, **GOA,** GOOD, *GOVERN*, **GRADE,** GRADIENT, *GRADUAL*, *GRAMMAR*, GRAMMATICAL, GREAT, GROUND, GROUP, GROW, **GUARANTEE,** *GUIDE*.

H HAND, *HANDLE*, HAPPEN, HARD, HEAD, HEAR, HEAVY, HEIGH, HELP, **HENCE,** HENCEFORTH, HERE, **HIERARCHY,** HIGH, **HIGHLIGHT,** HISTORY, *HIT*, HOLD, HOME, HOPE, *HOST*, HOUR, HOUSE, HOWEVER, HUMAN, **HYPOTHESIS**.

IDEA, **IDENTICAL, IDENTIFY,** IDIOSYNCRATIC, IF, **IGNORANT, ILLUSTRATE , IMAGE,** *IMAGINE*, *IMMEDIATE*, **IMPACT, IMPLEMENT, IMPLICATE, IMPLICIT, IMPLY,** IMPORTANT, **IMPOSE,** *IMPROVE*, INCLUDE, **INCORPORATE,** INCREASE, INDEED, INDEPENDENT, **INDEX, INDICATE,** INDICES, **INDIVIDUAL, INDUCE, INFER,** INFINITIVE, INFLECTIONAL, INFLUENCE, *INFORM*, *INFORMAL*, INFORMANTS, INFORMATIVE, **INHERENT, INITIAL, INNOVATE, INPUT, INSERT,** *INSIDE*, **INSIGHT, INSPECT, INSTANCE,** INSTANTIATD, INSTEAD, **INSTRUCT,** *INSTRUMENT*, **INTEGRATE,** *INTEND*, **INTENSE, INTERACT,** INTEREST, INTERESTINGLY ,INTERFACE, *INTERFERE*, **INTERMEDIATE, INTERNAL , INTERPRET,** INTERPRETABLE, **INTERVAL, INTERVENE,** INTONATION, INTRODUCE, INTUITIONS, INVENTORY, INVERSE, **INVESTIGATE, INVOLVE, INVOLVE, ISOLATE, ISSUE, ITEM.**

J JOINT, JUDGE, JUDGE, JUST, **JUSTIFY**.

K KEEP, *KEY*, KEYWORDS, KIND, KING, KNOW.

L **LABEL,** LACK, LANGUAGE, LANGUAGE-SPECIFIC, LARGE, LATE, LATTER, LAW, LEAD, LEARN, LEAVE, LEFT, LENGTH, LET, LETTER, LEVEL, LEXICAL, LEXICON, LIE, LIFE, LIGHT, LIKE, LIKELY, **LIKEWISE,** LIMIT, LINE, LINEAR, LINGUISTIC, LINGUISTS, **LINK,** *LIST*, LISTEN, LITERATURE, LITTLE, LIVE, LOCAL, **LOCATE,** LOCUS, **LOGIC,** LONG, LOSE, LOSS, *LOT*, LOVE, LOW.

M MACHINE, MAIN, **MAINTAIN, MAJOR,** MAKE, *MALE*, MAN, **MANIPULATE, MANUAL,** *MAP*, MAPPING, **MARGIN,** MARK, MARKEDNES, MASCULINE, *MATCH*, MATERIAL, MATRIX, MATTER, MAXIMAL, **MAXIMISE,** MEAN, MEASURE, **MECHANISM,** MEDIAL, **MEDIATE,** MEET, MEMBER, MEMORY, **MENTAL,** MENTION, MERE, MERGED, **METHOD,** METRICAL, MIDDLE, MIND, **MINIMAL, MINIMUM, MINOR,** MINUTE,

MIRROR, MISMATCH, *MIX*, MODAL, *MODEL*, MODERN, **MODIFY,** MOMENT, MONOLINGUAL, MONTH, MOREOVER, MORPHEME, MORPHOLOGY, MORPHOSYNTACTIC, MOTHER, *MOTION*, **MOTIVE,** MOVE, *MULTIPLY*, **MUTUAL.**

N NAME, *NARROW*, NATIVE, NATURE, NECESSARY, NECESSITY, NEED, **NEGATE,** NEIGBBOUR, **NEUTRAL,** NEVER, **NEVERTHELESS,** NEW, NODE, NOMINAL, NOMINATIVE, **NONETHELESS, NORMAL,** NORTH, NOTATION, NOTE, NOTICE, **NOTION,** *NOUN*, NOVEL, NOW, **NUCLEAR,** NUCLEUS, NULL, NUMBER.

O OBJECT, **OBJECTIVE,** OBLIGATORY, OBSRVE, **OBTAIN, OBVIOUS,** OCCASION, **OCCUPY, OCCUR, ODD,** OFFER, OFTEN, *OMIT*, **ONGOING,** ONLY, ONSET, OPEN, OPERATE, *OPPOSE*, *OPPOSITE*, OPPOSITION, OPTIMAL, **OPTION,** OPTIONALLY, ORDER, ORDINARY, ORGANIZE, **ORIENT,** *ORIGIN*, OTHERWISE, OUT, **OUTCOME,** *OUTLINE*, **OUTPUT, OVERALL, OVERLAP,** OVERT, OVERVIEW, OWN.

P *PAIR*,, PAPER, **PARADIGM,** PARADIGMATIC, **PARALLEL, PARAMETER,** *PARENT*, PART, **PARTICIPATE,** PARTICIPLE, PARTICIPLES, PARTICULAR, PASS, **PASSIVE,** PAST, *PATH*, *PATTERN*, *PAUSE*, PAY, PEAK, PEOPLE, **PERCEIVE, PERCENT,** PERCEPTUAL, *PERFECT*, *PERFORM*, PERHAPS, **PERIOD,** PERMIT, *PERSERVE*, PERSON, **PERSPECTIVE, PERSUME, PHASE, PHENOMENON,** PHONETICM, PHONOLOGICAL, PHRASE, **PHYSICAL,** *PICK*, PICTURE, PIECE, PITCH, PLACE, PLACEMENT , PLAN, PLAUSIBLE, PLAY, *PLURAL*, **PLUS,** POINT, POLARITY, POOR, **PORTION, POSE,** POSIT, POSITION, POSITIONAL, **POSITIVE,** POSSESS, POSSESSIVE, POSSIBLE, **POTENTIAL,** POWER, PPARSE, *PRACTISE*, PRAGMATIC, **PRECEDE, PRECISE,** REDICATE, **PREDICT, REDOMONANT,** *PREFER*, PREFIX, **PRELIMINARY,** PREPOSITIONAL , PRESENT, PRESS, PRESSURE, PREVENT, PREVERBAL, **PREVIOUS, PRIMARY,** PRINCE, **PRINCIPLE, PRIOR,** *PROBABLE*, PROBE, PROBLEM, **PROCEED, PROCESS,** PRODUCE, PRODUCT, *PROGRAMME*, PROGRESS, **PROHIBIT, PROJECT,** PROMINENT, PRONOUN, *PRONOUNCE*, PRONUNCIATION, PROPER, PROPERTY, **PROPORTION,** PROPOSE, PROSODIC, PROTOTYPICAL, PROVE, PROVIDE, **PSYCHOLOGY,** *PURE*, PURPOSE, **PURSUE,** PUT.

Q *QUALIFY*, **QUALITATIVE,** QUALITY, QUANTIFIED, QUANTITATIVE, QUANTITY, QUESTION, *QUICK*, QUITE, **QUOTE**.

R **RADICAL,** RAISE, **RANGE,** RANK, RANKINGS, *RARE*, RATE, RATINGS, **RATIO,** REACH, **REACT,** READ, READY, REAL, REAL, REALISE, REALIZATIONS, REALLY, REASON, RECALL,

RECEIVE, RECENT, RECOGNIZE, RECORD, **RECOVER,** RED, REDUCE, REDUNDANT, REDUPLICATION, *REFER,* REFERENCE, REFERENTS, **REFINE,** *REFLECT,* REGARD, **REGION, REGISTER,** *REGULAR,* **REJECT,** RELATION, RELATIONAL, RELATIVE, **RELEVANT, RELY,** REMAIN, REMARK, REMEMBER, **REMOVE,** *REPAIR*, *REPEAT*, *REPLACE*, REPORT, REPRESENT, REPRESENTATIONAL, **REQUIRE, RESEARCH, RESIDE,** *RESIST,* **RESOLVE, RESOURCE,** RESPECT, **RESPOND,** *RESPONSIBLE*, REST, **RESTRICT,** RESULT, **RETAIN,** RETURN, **REVEAL, REVERSE,** *REVIEW,* **REVISE,** RICH, RISE, ROBUST, **ROLE,** ROOM, *ROOT,* ROUGH, ROUND, RULE, RUN.

S SALIENT, SAME, *SAMPLE, SATISFY,* SAY, *SCALE,* **SCENARIO,** SCHEMA, **SCHEME,** SCHOOL, SCIENCE, **SCOPE,** SCORES, *SEARC,* SECT, **SECTION,** SEE, **SEEK,** SEEM, SEGMENTS, **SELECT,** SEMANTICS, SENSE, SENSITIVE, *SENTENCE,* SEPARATE, **SEQUENCE, SERIES,** SERIOUS, SERVE, SESSION, SET, SEVERAL, SHAPE, SHARE, *SHARP,* **SHIFT,** SHORT , SHOW, SIDE, SIGN, *SIGNAL,* **SIGNIFICANT,** SILENCE, **SIMILAR,** SIMPLE, SIMULTANEOUSLY, SING, SINGLE, SINGLETON, SINGULAR, SIT, **SITE,** SITU, SITUATION, *SLIGHT, SLOW,* SMALL, SO, **SO-CALLED,** SOCIAL, **SOLE,** *SOLVE,* **SOMEWHAT,** SOON, SORT, SOUND, **SOURCE,** SOUTH, SPACE, SPATIAL, SPEAK, SPECIAL, **SPECIFIC, SPECIFY,** *SPLIT* , SPONTANEOUS, SPREAD, **STABLE,** STAGE, STAND, STANDARD, START, STATE, **STATISTIC,** STATIVE, **STATUS,** STAY, *STEM,* STEP, STIMULI, STONE, STOP, STORE, STORY, **STRAIGHTFORWARD,** STRAIGHTFORWARDLY, **STRATEGY,** STRENGTH, **STRESS,** *STRICT,* STRIKE, *STRING,* STRONG, **STRUCTURE,** STUDENT, STUDY, **STYLE,** SUBJECT, **SUBSEQUENT,** SUBSET, SUBSTANCE, **SUBSTITUTE,** SUBTLE, SUCCEED, **SUFFICIENT,** SUFFIX, SUGGEST, *SUIT*, **SUM, SUMMARY,** SUPPORT, SUPPOSE, SURE, SUREPRISE, SURFACE, SURROND, **SURVEY, SURVIVE,** SYLLABLE, **SYMBOL,** SYNCHRONIC, SYNTACTIC, SYNTAX,, SYSTEM.

T TABLE, TABLEAU, TAKE, TALK, **TARGET, TASK,** TEACH, **TECHNICAL,** *TELEPHONE,* TELL, TEMPLATE, TEMPORAL, *TEND,* **TENSE,** TERM, **TERMINATE,** TERMINOLOGY, TEST, **TEXT, THEME,** THEN, **THEORY,** THER, **THEREBY,** THEREFORE, THING, THINK, THOUGH, THREE-WAY, THROUGH, THUS, *TIE, TIGHT,* TIME, TODAY, TOGETHER, TOKENS, TONAL, TONE , TOO, *TOOL,* TOP, **TOPIC,** TOTAL, TOWARD, **TRACE,** *TRACK,* **TRADITION,** TRAIN, **TRANSFER, TRANSFORM, TRANSIT,** *TRANSLATE,* TRANSPARENT, *TREAT,* TREE, **TRIGGER,** TRUE, TRY, TURN, TYPE, *TYPICAL,* TYPOLOGY.

U	**ULTIMATE,** UNACCENTED, UNCLEAR, **UNDERGO, UNDERLIE,** UNDERSPECIFIED, UNDERSTAND, UNEXPECTED, **UNIFORM, UNIFY, UNIQUE,** UNITE, UNIVERSALLY, *UNIVERSE*, UNIVERSITY, UNMARKED, UNRELATED, *UPPER*, USAGE, USE, USUAL, **UTILISE,** UTTERANCE, UTTERED**.**
V	**VALID,** VALUE, VAN, VARIETY, VARIOUS, **VARY,** *VERB*, VERBAL, **VERSION,** VERSUS, VERY, **VI,** VICE, VIEW, **VIOLATE,** VIRTUE, **VISIBLE, VISUAL,** VOCABULARY**,** VOICE, **VOLUME,** *VOWE***L.**
W	WANT, WAY, *WEAK*, WEAKER, WEEK, *WEIGH*, WELL, WELL-KNOWN, WEST, **WHEREAS, WHEREBY,** WHETHER, WHOLE, WH-QUESTIONS, WHTE, WIDE, WIN, WINDOW, WITHIN, WOMAN, WONDER, WORD, WORK, WORLD, WORTH, WRITE, WRONG
Y	YEAR, YES, YET, YIELD, YOUNG.

- Note: AWL words are bold, non- GSL/non- AWL words are underlined, 2nd GSL are italicized, and the rest are 1st GSL words.