ISSN (P) 2058-6906 · ISSN (E) 2790-3400 DOI 10.35534/lcs.0601003

Language and Cognitive Science ─ 6(1): 44−59

Mingxian Li, Yujun Lee* and Dilin Liu

Development of a Neurology Word List: A Corpus-based Study

Abstract: The research finding that vocabulary use is often discipline-bound (Hyland and Tse 2007) has demonstrated the need for discipline-specific word lists. Driven by this need, this study aims to explore the viability and validity of developing a subdiscipline word list – a vocabulary list in neurology, a subdiscipline of medicine. A neurology corpus (NeuroC) of 6,180,718 words was compiled drawing from 970 neurological research articles systematically chosen from 10 journals. Using AntConc and AntWordProfiler as corpus tools and applying rigorous frequency and range criteria, the study selects a total of 717 words as our Neurology Word List (NeuroWL). To test the validity of the NeuroWL, the coverages of the NeuroC by the NeuroWL and by Coxhead's AWL are calculated and compared. The results reveal that the coverage by the NeuroWL (12.99%) almost doubles that by Coxhead's Academic Word List (AWL) (6.06%). Additional analyses indicate that the NeuroWL shows an overlap of only 32.4% with the AWL, 53.6% with Wang et al.'s (2008) Medical Academic Word List (MAWL), and 43.9% with Yang's (2015) Nursing Academic Word List (NAWL). These findings testify to the validity of the NeuroWL and its pedagogical significance to ESP/EFL teachers/learners in neurology and the need for developing subdiscipline word lists.

Keywords: academic English; English for specific purposes; subdiscipline word list; vocabulary learning

***Corresponding author, Yujun Lee:** Key Library of Artificial Intelligence and Cognitive Neuroscience, Xi'an International Studies University, Xi'an, China; Foreign Language and Culture Department, North Sichuan Medical College, Nanchong, China; e-mail: sammyleee@163.com

Mingxian Li: College of Foreign Languages, Shandong University of Science and Technology, Qingdao, China; e-mail: talmx@163.com

Dilin Liu: Department of English, The University of Alabama, Tuscaloosa, AL, USA; e-mail: dliu@ua.edu

1 Introduction

It is generally held that word lists should be developed specific to their respective disciplines (Mudraya et al. 2006; Hyland and Tse 2007; Hsu et al. 2014; Yang et al. 2015; Lei et al. 2016). However, no study exclusively on academic word lists for the field of neurology has been conducted so far, despite the fact that academic word lists have been compiled for other fields. As an important subdiscipline of clinical medicine, neurology has gained even greater attention with the development of brain science and neurocognitive science. Against this backdrop, the present study aims to develop a neurology academic word list with the following factors taken into consideration:

(1) Graduate students still encounter problems when reading and writing English academic articles, even after they have finished the study of general medical textbooks.

(2) Many researchers from other disciplines such as psychology, linguistics, and computer science are becoming more involved in exploring how the brain works, and mastery of neurology-specific vocabulary facilitates a more comprehensive and precise understanding of the materials concerned. Linguistic researchers, for instance, probing the mechanism of how patients with brain tumors differ from normal persons in language production will work more efficiently with sufficient and specific academic vocabulary in neurology.

(3) Previous studies have shown that there are words which are not in Coxhead's (2000) Academic Word List (AWL) but have high frequencies in specific disciplines (Yang 2015; Liu and Han 2015; Hyland and Tse 2007), which implies that the AWL is not representative enough. For example, Yang (2015) studied the most frequently used academic vocabulary in nursing and developed a Nursing Academic Word List (NAWL). Yang found that the NAWL coverage in the corpus is 13.64%, which is higher than that of the AWL coverage in Coxhead's four sub-corpora of commerce (12%), arts (9.3%), law (9.4%), and science (9.1%) and that of medical research (10.07%), as is found in Chen and Ge (2007). Furthermore, Liu and Han (2015) studied the academic word list for environmental science and found that the AWL is not entirely useful for learners in this discipline because the coverage of some word families is narrow and some words with high frequencies in environmental texts are not covered.

Four word lists have been recognized as important milestones in the history of vocabulary studies. The General Service List (GSL) is the first of these. In 1953, West developed the GSL from a 2.5 million-word corpus according to the criteria of frequency, universality, utility, and usefulness. It has been extensively examined to determine how representative it is of written texts (Nation 2004; Newman 2016). The paradigm of word family still has impacts on academic word studies in spite of the limitations later identified by researchers. The second list is the University Word List (UWL). In 1984, Xue and Nation created this large-scale academic word list named UWL, which has an influence in a variety of teaching and research contexts. Another word list that followed is the AWL, a larger and more modern corpus of academic materials compiled by Coxhead (2000). Coxhead explored lexical items and their different frequencies in different texts and worked out the coverages of AWL in academic and nonacademic corpora.

The AWL corpus consists of articles from journals, textbooks, and texts from Bauer's (1993) Wellington Corpus of New Zealand Written English, covering 28 areas in the domains of arts, commerce, law, and science. About 3.5 million running words primarily of New Zealand English from these domains were collected, with approximately 875,000 words in the disciplines of arts, commerce, law, and science. The AWL was developed following three criteria: (1) the first 2,000 word families should be excluded; (2) word families had to be found more than 10 times in each of the four sections of the corpus and in at least 15 of 28 subject areas represented; and (3) the members of a word family had to occur 100 times or more in the entire corpus, with an average of 25 times in the four outlined sections of the corpus (arts, commerce, law, and science).

The three criteria have had far-reaching influence on research thereafter. However, the representativeness of AWL has been questioned, for some of the words in the first 2,000 word families actually have academic meanings, which means they should have been included in the academic word list. For instance, Gardner and Davies's study (2014) displayed that many high-frequency general words, such as *interest, capital, and rate*, have special or different meanings in academic contexts, and some high-frequency general words actually boast even higher frequency in an academic English context. Therefore, two types of academic word lists should be developed: the more general interdisciplinary ones and discipline-specific ones. This is because many studies have confirmed that the coverage of AWL, usually about 6%, is relatively low in discipline-specific corpora (e.g. Lei and Liu 2016). Up to now, some influential academic discipline-specific word lists have been created, such as Hsu's (2014) Engineering Academic Word List, Martínez et al.'s (2009)

Agricultural Academic Word List, Yang's (2015) Nursing Academic Word List, and Wang et al.'s (2008) Medical Academic Word List (MAWL). However, a neurology academic word list has not yet been developed.

Besides developing the NeuroWL, this study also attempts to address the following questions on establishing the NeuroWL: (1) What is the AWL's coverage of the neurological academic research article corpus (NeuroC)? (2) Are word families in the NeuroWL unique to the AWL? (3) Compared with the MAWL and NAWL, are word families in the NeuroWL unique?

The development of NeuroWL should also have implications for textbook design and dictionary compilation in the field of neurology and contribute to more efficient ESP/EFL teaching and learning, as well as academic research in neurology.

2 Method

2.1 The Corpus Used in the Study

Based on previous literature (Liu and Han 2015; Sinclair 1991), we established a neurological academic research article corpus (NeuroC) as the database for this study. First, we searched neurology journals in Web of Science, which produced more than 70 journals. Second, 70 journals were rated by two experts from the Neurology Department at a Chinese medical university. Then, the top 10 rated journals were selected as the data source of NeuroC (see Table 1) based on the principle of evaluation consistency. In the end, the 110 most-cited articles from each of the 10 journals were chosen, which was done automatically using the search functions of Web of Science.

The NeuroC consists of 6,180,718 running words, 108,669 types, and 970 written texts from 2006 to 2015. We had intended for the corpus to comprise 1,100 texts (110 articles * 10 journals = 1100 texts), but only 970 texts were ultimately selected because some journals are new publications. For example, the first issue of JAMA NEUROLOGY was published in 2013, hence only 30 articles could be selected for the research. What is worth noting is that some review articles are also included in the texts selected, a practice that we adopted from Lei and Liu (2016), who hold the view that reviews contain important academic information with even deeper insight and diverse perspectives.

Table 1: Top 10 journals in the neurology field

No.	Name of journal	Acronym
1	ACTA NEUROPATHOLOGICA	AN
2	ALZHEIMER'S AND DEMENTIA	AD
3	ANNALS OF NEUROLOGY	AON
4	BRAIN	BRAIN
5	JAMA NEUROLOGY	JN
6	LANCET NEUROLOGY	LN
7	MOLECULAR NEURODEGENERATION	MN
8	NATURE REVIEWS NEUROLOGY	NRN
9	NEUROLOGY	NE
10	NEUROSCIENTIST	NS

2.2 Data Processing and NeuroWL Item Extraction and Selection

Our data processing first incorporated standardization of the articles and reviews before storage in the corpus and the retrieval of words for the database. Then, we manually created 10 files for the 10 journals and another 10 files for the unpolluted texts by deleting charts, diagrams, numbers, appendices, bibliographies, equations, and other distracting textual components in the articles and classifying them on the basis of the journals.

Next, the 970 texts of the NeuroC were loaded into Laurence Anthony's (2015) free software program AntConc (3.4.4w) for the extraction and selection of our NeuroWL. We followed Coxhead's (2000) and Yang's (2015) practice by adopting frequency and range as selection criteria on the basis of word family. Range was used to ensure that our list would not include items that met the frequency criterion but appeared in only a few journals and articles. Specifically, in terms of frequency and range, the selected words in the NeuroWL beyond the first 2,000 most frequently occurring word families had to boast a minimum frequency of 200 and occur in at least 9 of the 10 journals. The AntConc program helped identify 1,567 word families that each had a frequency of 200 or more.

These 1,567 word families and the 10 journal files were then loaded into Laurence Anthony's (2015) AntWordProfiler (1.4.0) to check the occurrence of these words in the ten journals. Those of the 1,567 items that occurred in at least 9 journals (our range measure mentioned above) were chosen for the final NeuroWL.

After the establishment of the NeuroWL, we used the AntWordProfiler to compare the coverages of the NeuroC by the GSL, the AWL and the NeuroWL respectively, as well as the AWL and the NeuroWL coverages in the ten journals, to learn their coverage distribution patterns. Finally, coverage comparisons were made between NeuroWL and MAWL, and between NeuroWL and NAWL.

3 Results

In this section, we first briefly describe the NeuroWL that we developed based on the procedures and criteria described above. Then we report and compare the results of the coverages of the NeuroC by the GSL, the AWL and our NeuroWL.

3.1 The Neurology Word List

Based on the criteria mentioned above in the methodology section, 717 word families met all the criteria and were ultimately selected, and formed the NeuroWL as shown in Appendix1. Table 2 shows the top 10 and bottom 10 word families in the NeuroWL (words in bold type also occurred in the AWL). We compared the items of the AWL with those in our NeuroWL. The results show that only 233 of the AWL's 570 word families are also in the NeuroWL, that is to say, there is only a 32.4 overlap.

Table 2: The ten top and bottom word families in the NeuroC

Headword	Range	Freq.	Headword	Range	Freq.
patient	10	27354	medial	10	218
range	10	23480	span	10	217
clinical	10	12756	***enable***	10	217
novel	10	10999	***expert***	10	216
dementia	10	10459	***adequate***	9	215
cell	10	10315	***insight***	10	213
Alzhheimer's	10	9877	disrupt	10	210
data	10	9596	polymorphisms	10	207
cognitive	10	9279	***conflict***	10	204
tau	10	8990	simultaneously	10	202

3.2 The GSL and the AWL Coverages in the NeuroC

According to Hsueh-chao and Nation (2000) and Nation (2001), adequate reading comprehension depends on a knowledge of 90% to 95% of the words in a text. It is thus very important to know the coverages of the NeuroC by the GSL, the AWL, and our NeuroWL. The results of the analysis of the GSL and the AWL coverages of the NeuroC are reported in Table 3.

Table 3: Coverage of the NeuroC by the GSL and the AWL

Word list	Coverage (%)	Number of word families
GSL's first 1000 words	54.05	987
GSL's second 1000 words	5.08	872
AWL	6.06	566

Regarding the GSL, its first and second most frequently occurring word families account for 59.13% in the NeuroC, a coverage that falls far below the 90%–95% required for successful comprehension of research articles in neurology. In other words, with a mere mastery of general words (i.e. the first 2,000 most frequent words), a reader will not be able to comprehend research articles in the field of neurology. Concerning the AWL, of its 570 word families, 566 appeared in the NeuroC, covering 6.06% of the running words, which is in line with the coverage (6.27%) of the previous medical text study by Cobb and Horst (2002). However, the coverage is different from those shown in studies on word list developments in other fields (Chen et al. 2007; Coxhead and Byrd 2007; Martínez et al. 2009; Vongpumivitch et al. 2009). For example, the AWL covered 9.06% of the tokens in agriculture research articles (Martínez et al. 2009) and 10.7% of the tokens in medical research articles (Chen et al.).

3.3 AWL and NeuroWL Coverage in the NeuroC and in the Different Journals

The AWL and NeuroWL coverage in the NeuroC and in each of the ten journals is displayed in Table 4. As shown in Table 4, while the AWL's coverage in the NeuroWL is 6.06%, the NeuroWL's coverage in it is 12.99%, which more than doubles the former. This means that the AWL is not adequate and hence not useful enough to equip students for study in the neurology science discipline. Therefore, this result demonstrates the necessity for developing the NeuroWL, although the AWL has its virtues for academic English study in general, which cannot be ignored.

Table 4: AWL and NeuroWL coverage in the NeuroC and the ten journals

Corpus	NeuroC	Different Journals									
		AN	AD	AON	BR	JN	LN	MN	NRN	NE	NES
AWL coverage (%)	**6.06**	5.38	6.42	6.35	6.06	6.77	6.05	6.22	5.4	6.13	6.79
NeuroWL coverage (%)	**12.99**	13.71	12.5	13.67	13.5	14.04	13.13	12.91	13.87	12.42	11.34

Regarding coverage in the 10 journals, the percentages of coverage of the AWL and the NeuroWL vary noticeably across the ten journals. The AWL boasts a coverage of 6.79% (its highest) in the NES journal but only has a coverage of 5.38% (its lowest) in the AN journal. Similarly, the NeuroWL shows a coverage of 14.05% (its highest) in the JN journal but displays a coverage of 11.34% (its lowest) in the NES journal. These findings confirm the need for choosing articles from many different journals in a field to compile a discipline-specific corpus, such as our NeuroC. In other words, an adequate number of journals should be included for creating a discipline-specific academic corpus.

The following passage with 31 lines and 246 number tokens is a paragraph randomly selected from one neurological research article (Attems et al. 2005) to establish that NeuroWL is representative enough as a neurology academic word list. The NeuroWL's coverage in the text is 13.8%, slightly higher than the 12.99% coverage in the NeuroC as reported above. NeuroWL words are bold-faced.

> In **summary**, our results **indicate** that progressing AD **pathology** significantly increases the severity of **overall** CAA when statistics are calculated for cohorts including both CAA-**positive** and -negative cases. This increase is mainly caused by increasing **occipital** CAA. By contrast, in cohorts consisting of CAA-**positive** cases alone, increasing AD **pathology** fails to (statistically) increase the severity of **overall** CAA, but only increases the severity of **occipital** CAA. Progressing AD **pathology** thus shifts the topographical distribution of CAA towards the **occipital** cortex. Severe CAA is sometimes present in the total absence of AD **pathology**, while some cases with severe AD lack any CAA, suggesting that **neuritic** AD pathology and CAA might represent different entities. If present together in the first place, however, they seem to strongly

> influence each other. This is supported by recent findings in AD brains of more severe **tau** labeling around Ab-laden **cerebral** vessels than around those without Ab **deposition**, suggesting that both CAA and perivascular deposition of hyper phosphorylated tau are a consequence of elevated levels of **soluble** Ab around **cortical** blood vessels. It was suggested that CAA develops in a hierarchical way, **similar** to the topographical distribution pattern of parenchymal Ab **deposits** with initial involvement of leptomeningeal and parenchymal vessels in the **neocortex**, followed by involvement of allocortical regions, **cerebellum**, and deep brain structures. This can not explain the frequent and **prominent** involvement of the **occipital** region by CAA, nor the strong **correlation** of its increasing severity with progressing AD **pathology**.

Our comparison of our NeuroWL's coverage in the NeuroC with the coverage of other discipline-specific academic word lists in their respective discipline corpora indicates that the NeuroWL's 12.99% coverage in the NeuroC is comparable to or consistent with the MAWL's 12.24% medical corpus coverage reported by Wang et al. (2008) and the NAWL's 13.64% nursing corpus coverage reported by Yang (2015). We also directly compared the items in our NeuroWL with those in Wang et al.'s (2008) MAWL and Yang's (2015) NAWL to determine to what extent the items on the lists overlap. The results show that the overlap rates are not very high, with the overlaps being 53.6% between NeuroWL and the MAWL and 43.9% between the NeuroWL and the NAWL. These fairly low overlap rates between the discipline-specific vocabulary lists further demonstrates the importance and necessity of a discipline-specific academic word list in neurology.

4 Discussion

The study has developed a neurological academic word list. The results in Table 4 indicate that the NeuroWL's coverage in the NeuroC is significantly different from and also higher than the coverage offered by Coxhead's AWL. The coverage of the AWL in the NeuroC is 6.06%, and its coverage varies from 6.79% to 5.38% in the 10 different journals. The coverage of the NeuroWL in the NeuroC is 12.99%, which more than doubles the coverage of the AWL. Therefore, it is reasonable to argue that the NeuroWL is more representative of the vocabulary used in neurology than the AWL. Furthermore, the study shows that the NeuroWL does not have a very high overlap with the MAWL and the NAWL respectively, for its overlap with the MAWL is 53.6% and its overlap with the NAWL is only 43.9%. The NeuroWL was expected to have a high overlap with the MAWL and the NAWL, because neurology science, medicine, and nursing science are all subdisciplines of the life sciences, but the results of our study do not seem to support this presupposition.

These results support the argument for developing discipline- and subdiscipline-specific academic word lists in addition to general academic word lists. Quite a number of scholars have argued that we should develop more discipline-/subdiscipline-specific academic word lists, as each academic discipline or even subdiscipline has its own way of explaining experience, its own forms of argumentation, and its own preferred forms, meanings, and collocation patterns (Samraj 2002; Hyland and Tse 2007; Martínez et al. 2009; Yang 2015). For instance, Hyland and Tse (2007: 247–248) believe that different research practices and different ways of seeing the world are associated with different forms of argument, preferred forms of expression, and specialized uses of lexis. As evidence, Yang (2015), in her development of the NAWL, found that the AWL's coverage in nursing was low and attributed this finding to subdiscipline-specific word preferences.

In fact, the need for both general academic and discipline-/subspecific academic word lists can also be explained from the perspective of cognitive lexical semantics. In cognitive linguistics, lexical semantics (word meaning) has *profile and domain*. Profile refers to the concept symbolized by the word in question. *Domain* refers to any system of concepts related in such a way that to understand any of their concepts. It is necessary to understand the whole structure in which it fits (Croft and Cruse 2004: 14–19). In other words, knowledge or conceptual structure (*domain*) presupposes the profiled concept, and profile cannot be separated from domain. Put differently, the meaning of a word will not be available without a certain domain. For instance, GROUND and LAND denote (profile) the same thing. It is difficult to distinguish them in profile. However, we can distinguish them according to their different domains. LAND describes the dry surface of the earth in contrast with SEA, while GROUND describes the dry surface of the earth in contrast with AIR (Fillmore 1982a: 121). Here, SEA and AIR are their domain or base respectively. This implies that word meaning is modulated by domains. In other words, experts of a given specific discipline form a unique community, and the expertise they share is a unique conceptual structure, which is crucial for accessing the academic words in their field. In this study, a specific discipline or subdiscipline is conceptualized as domain. Therefore, the meaning of a given word should be established by a specific discipline or subdiscipline. Hence, more specific disciplinary/subdisciplinary word lists are needed.

Similarly, the fact that some of the words in the discipline-specific word list are common words with discipline-specific meanings (i.e. polysemous words) also provides support for this cognitive semantic perspective. The different meanings of these words come from different domains rather than profiles. For example, the word "mouth" denotes one profile, the opening to a container. However, its domain can be BODY, BOTTLE, CAVE, and RIVER (Croft and Cruse 2004: 19). The word "mouth" is polysemous because it has a sense for each of the profile-base pairings. This fact further supports the need for developing discipline-/subdiscipline-specific academic word lists.

5 Implications and Limitations: A Conclusion

The NeuroWL developed in this study may be helpful to material writers, teachers, and students in the field of neurology. Material writers may use the list as a guide in developing teaching materials (including textbooks) by making sure that the lexical items of the word families in the list are covered and by prioritizing especially those with a very high frequency. Similarly, teachers may employ the list to help concentrate their teaching on the items on the list based on their frequency and to establish a frequency- and needs-based teaching sequence. For students, the NeuroWL, with fewer than 800 items, may serve as a vocabulary study list to facilitate learning.

For future research, this NeuroWL may need to be further tested against other neurology corpora, especially those covering a larger number of journals and a longer time period. This is because, in this study, we selected articles from only ten journals over just a ten-year period. In other words, the data of our NeuroC may not be representative enough. Furthermore, although we have made the suggestion that target words for academic word lists should be chosen on the basis of domain or frame, how to determine domains or frames in specific disciplines with machine learning has not yet been discussed and remains a subject for future studies.

References

Attems, Johannes, Kurt A. Jellinger & Felix Lintner. 2005. Alzheimer's disease pathology influences severity and topographical distribution of cerebral amyloid angiopathy. *Acta neuropathologica* 110(3). 222–231.

Bauer, Laurie. 1993. Introducing the Wellington corpus of written New Zealand English. *Te Reo* 37. 21–28.

Chen, Qi & Guang-chun Ge. 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.

Cobb, Tom & Marlise Horst. 2002. Is there room for an academic world list in French? (Draft chapter) [Electronic version]. In Paul Bogaards & Batia Laufer (eds.), *Vocabulary in a second language: Selection, acquisition & testing* (Language Learning & Language Teaching 10), 15–38. Amsterdam: John Benjamins Publishing Co.

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

Coxhead, Averil & Pat Byrd. 2007. Preparing writing teachers to teach the vocabulary and grammar of academic prose. *Journal of Second Language Writing* 16. 129–147.

Croft, William & D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.

Fillmore, Charles. 1982. Frame semantics. In The Linguistic Society of Korea (ed.), *Linguistics in the morning calm*, 111–137. Seoul, Korea: Hanshin Publishing Company.

Gardner, Dee & Mark Davies. 2014. A new academic vocabulary list. *Applied Linguistics* 35. 305–327.

Hsueh-chao, Marcella Hu & Paul Nation. 2000. Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language* 13(1). 403–430.

Hsu, Wenhua. 2014. Measuring the vocabulary load of engineering textbooks for EFL undergraduates. *English for Specific Purposes* 33. 53–64.

Hyland, Ken & Polly Tse. 2007. Is there an "academic vocabulary"? *TESOL Quarterly* 41. 235–253.

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

Liu, Jia & Lina Han. 2015. A corpus-based environmental academic word list building and its validity test. *English for Specific Purposes* 39. 1–11.

Martínez, Iliana A., Silvia C. Beck & Carolina B. Panza. 2009. Academic vocabulary in agriculture research articles: A corpus-based study. *English for Specific Purposes* 28. 183–198.

Mudraya, Olga. 2006. Engineering English: A lexical frequency instructional model. *English for Specific Purposes* 25. 235–256.

Nation, Paul. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, Paul. 2004. A study of the most frequent word families in the British National Corpus. In Paul Bogaards & Batia Laufer (eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (Language Learning & Language Teaching 10), 3–13. Amsterdam: John Benjamins Publishing Co.

Newman, Jacob Andrew. 2016. *A corpus-based comparison of the Academic Word List and the Academic Vocabulary List*. MA thesis, Brigham Young University.

Samraj, Betty. 2002. Disciplinary variation in abstracts: The case of behavior and conservation biology. In John Flowerdew (ed.), *Academic discourse*, 4–56. London: Pearson.

Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Vongpumivitch, Viphavee, Ju-yu Huang & Yu-Chia Chang. 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. 33–41.

Wang, Jing, Shao-lan Liang & Guang-chun Ge. 2008. Establishment of a medical academic word list. *English for Specific Purposes* 27. 442–458.

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

Xue, Guoliang & Paul Nation. 1984. A university word list. *Language Learning and Communication* 3. 215–229.

Yang, Ming-nuan. 2015. A nursing academic word list. *English for Specific Purposes* 37. 27–38.

Bionotes

Mingxian Li

Mingxian Li (1978) is Associate Professor and master supervisor at the College of Foreign Languages in Shandong University of Science and Technology, and a visiting scholar at the University of Alabama. Her research interests include description and teaching of semantic and pragmatic meanings of language from a cognitive point of view. Her recent book (in Chinese) is *On the memory of English words*. She has published articles in many Chinese journals.

Yujun Li

Yujun Li(1978) is Associate Professor in the English Department of North Sichuan Medical College and Neural Disease Institute of North Medical College. His research interests include neurolinguistics and language rehabilitation. He has published articles in leading international journals, such as *Neuroreport* and *Brain and Behaviour*.

Dilin Liu

Dilin Liu (1953) is Professor Emeritus of Linguistics/TESOL in the English Department at the University of Alabama. His research interests include corpus/cognitive linguistic approaches to the study and teaching of English grammar and vocabulary. He has published extensively, including seven books

(six authored and one edited) and numerous articles in leading international journals, such as *Applied Linguistics, Cognitive Linguistics*, and *Modern Language Journal*. His most recent book (co-authored) is *Conducting sentiment analysis* (CUP, 2021).

Appendix

Neurological Word List (listed in alphabetical order)

Letter	Word family
A	aberrant; abnormality; abstract; abundant; academy; accelerate; access; accurate;
achieve; acid; activation; acute; acta; additionally; adequate; adhesion; adjacent; administer;	
adult; adverse; affect; affinity; ageing; agitation; aggregate; agonist; algorithm; alpha; alter;	
alternative; Alzheimer; amino; amygdale; amyloid; amyotrophic; anatomical; annual;	
ANOVA; antagonist; anterior; antibody; antigen; apathy; aphasia; APOE; apoptotic; apparent;	
approach; appropriate; area; array; assay; assess; astrocytes; astrocytic; asymptomatic;	
atrophy; attenuate; atypical; auditory; autoimmune; autopsy; autosomal; available; axonal;	
B	barrier; basal; baseline; benefit; beta; bias; bilateral; biochemical; biological; biology;
biomarker; biomedical; biopsy; blot; Braak; brainstem; brief; buffer; bulbar; burden;	
C	calcium; cancer; candidate; capable; capacity; cardiac; cardiovascular; carrier; cascade;
category; causal; cell; cellular; cerebellar; cerebellum; cerebral; cerebrospinal; cerebrovascular;	
challenge; channel; cholesterol; chromosome; chronic; cingulate; circuit; classic; clearance;	
clinic; clinical; clinician; cluster; cognition; cohort; community; complement; complex;	
component; compound; comprehensive; concentration; concept; concomitant; confirm;	
conflict; connectivity; consecutive; consent; considerable; consortium; constant; context;	
contrast; contribute; cord; core; correlation; cortical; criteria; crucial; culture; cycle; cytokine;	
cytoplasmic;	
D	data; database; decade; decline; defect; deficiency; deficit; define; definite;
degeneration; degradation; deletion; dementia; demographic; demonstrate; dendritic; density;	
dentate; depletion; deposit; design; despite; detect; deterioration; developmental; diabetes;	
diagnosis; differentiate; diffuse; disability; disorder; display; disrupt; disruption; distal; distinct;	
diverse; DNA; domain; dopamine; dopaminergic; dorsal; dose; downstream; drug; duration;	
dynamic; dysfunction;	
E	efficacy; elan; elevate; embed; embryonic; emission; empathy; emotional; enable;
encephalitis; encode; endogenous; endothelial; enhance; enrich; enroll; entorhinal;	
environment; enzyme; epidemiology; epilepsy; episode; equivalent; error; establish; estimate;	
etiology; evaluate; evident; evoke; excitatory; exclude; executive; exhibit; expert; external;	

Letter	Word family
extracellular;	
F	facilitate; factor; familial; feature; fetal; fibre; figure; final; finding; fisher; fluid ;
fluorescence; fluorescent; focal; focus; foundation; fraction; fragment; frequency; frontal;	
frontotemporal; function; fundamental; furthermore;	
G	GABA; gad; ganglia; gender; gene; generation; genetic; genome; genotype; glial;
glucose; glutamate; glutamatergic; goal; gradient; granule; rant; gray; gyrus;	
H	haemorrhage; hallucination; healthcare; hemisphere; hence; heterogeneity;
heterogeneous; hexanucleotide; hippocampal; hippocampus; histological; homeostasis;	
hypertension; hypothesis;	
I	identical; identify; IgG; image; imaging; immune; immunohistochemistry;
immunoreactivity; Immunostaining; impact; impaired; impairment; incidence; inclusion;	
incubate; incubation; index; indicate; individual; induce; infarct; infect; inferior; inflammation;	
inflammatory; infusion; inherit; inhibit; inhibitor; initial; injection; input; insight; insoluble;	
instance; institute; intact; integrity; intense; interact; interestingly; interleukin; intermediate;	
interval; intracellular; intracranial; intravenous; intrinsic; investigate; involve; ischemic; isoform;	
issue;	
J	junction;
K	kDa; kinase; kit; knockout; knockdown;
L	label; laboratory; lancet; lateral; layer; lesion; levodopa; Lewy; ligand; limbic; linear;
link; lipid; lobar; lobe; localized; locus; longitudinal; lymphocyte; lysosomal;	
M	magnetic; magnitude; maintain; major; mammalian; manifestation; manual;
manuscript; mapping; markedly; matrix; mature; mayo; MCI; mechanism; media; medial;	
median; mediate; medical; medication; medium; membrane; mental; metal; meta; metabolic;	
method; methyl; microglia; microscopic; microscope; microtubule; mini; minimal; minimum;	
mitochondrial; MMSE; mode; modulate; molecule; monoclonal; morphology; mortality;	
mortem; MRI; MRNA; muscle; mutation; myelin;	
N	necrosis; neocortex; neocortical; neonatal; nerve; network; neural; neuritic; neuro;
neurodegeneration; neurofibrillary; neuroimaging; neurol; neurology; neuronal;	
neuropathological ; neuroprotective; neuropsychiatric; neuropsychological; neuroscience;	
neurosurgery; neurotoxic; neurotransmitter; neurotrophic; nevertheless; nigra; nitric; NMDA ;	
node; nonetheless; normal; notion; novel; nuclei; nucleus;	
O	obesity; objective; observational; obtain; obvious; occipital; occur; olfactory;
oligodendrocyte; oligomer; ongoing; onset; optic; optimal; oral; outcome; output; overall;	
overexpression; overlap; overnight; oxidative; oxygen	
P	palsy; panel; paradigm; parallel; parietal; parenchyma; parietal; parkinsonism;
participate; passive; pathogenesis; pathogenic; pathology; pathophysiology; pathway; peak;	
peptide; percent; period; peripheral; perspective; pharmaceutical; pharmacological; phase;	
phenomenon; phenotype; phosphate; phosphorylate; physical; physician; physiology; placebo;	
plaque; plasma; plasticity; plus; poly; polyclonal; polymorphisms; pooled; portion; positive;	
positron; posterior; postmortem; postsynaptic; potential; precise; preclinical; precuneus;	

Letter	Word family
precursor; predict; predictor; predominant; preferentially; prefrontal; preliminary; presynaptic;	
prevalence; previous; primary; principal; prior; process; profile; prognosis; progression; project;	
proliferation; prolong; prominent; promote; proportion; protease; protein; protocol; psychiatry;	
publication; putative; pyramidal;	
Q	quantification; quantitative;
R	randomize; range; ratio; reactivity; recall; receptor; recombinant; recruit; recurrent;
reference; region; regression; regulate; relapse; release; remit; repetitive; require; research;	
residual; resistant; resonance; respiratory; respond; retrieval; retrospective; rev; reveal;	
reverse; robust; roche; rodent; role; routine;	
S	scan; schizophrenia; sclerosis; score; secretase; section; seizure; sensory; sequence;
serial; series. serum; settings; sex; shift; signaling; significant; similar; simultaneously; slice;	
sodium; soluble; somewhat; source; span; spatial; species; specific; spectrum; spinal;	
spontaneous; sporadic; stable; standardize; status; stimulation; strategy; stress; striatum;	
structure; subcortical; subgroup; subjective; subsequent; subset; substantia; substrate; subtle;	
subtype; subunit; sufficient; suggestive; sum; summary; superior; surgery; survey; susceptibility;	
symptom; synaptic; syndrome; synthesis; synuclein; systemic;	
T	tangle; target; task; tau; technique; technology; temporal; termed; text; thalamus;
theory; therapeutic; therapy; thereby; threshold; timing; tissue; tomography; toxic; tracer;	
tract; trans; transcription; transfer; transgenic; transient; translational; transport; traumatic;	
trend; trigger; tumor; tyrosine;	
U	unclear; undergo; underlie; unique; unrelated; uptake
V	van; vary; vascular; ventral; verbal; version; versus; via; visible; visual; vitro; vivo;
voltage; volume; voxel; vulnerability;	
W	weight; whereas; widespread; worsen;
Z	zone;