

Contents lists available at [ScienceDirect](http://www.sciencedirect.com)

Journal of English for Academic Purposes

journal homepage: www.elsevier.com/locate/jeap

A new medical academic word list: A corpus-based study with enhanced methodology

Lei Lei ^a, Dilin Liu ^{b,*}

^a Department of English, School of Foreign Language Studies, Huazhong University of Science and Technology, #1037 Luoyu Road, Wuhan, Hubei, 430074, PR China

^b Department of English, The University of Alabama, Box 870244, Tuscaloosa, AL, 35487, USA

ARTICLE INFO

Article history:

Received 31 March 2015

Received in revised form 24 January 2016

Accepted 25 January 2016

Available online 3 February 2016

Keywords:

Corpus

Academic English

Medical English

Vocabulary/word lists

ABSTRACT

Vocabulary use often varies significantly across academic disciplines. Hence, it is important to develop discipline-specific academic vocabulary lists. So far, all of the existing discipline-specific word lists have been developed using Coxhead's (2000) method that excluded general high-frequency words. However, Gardner and Davies's (2014) recent work on a new academic vocabulary list has challenged such a practice because many general high-frequency words have a much higher frequency in academic English than in general English and often have special meanings in academic English. Drawing on and combining methods and procedures from Coxhead (2000) and Gardner and Davies (2014), this corpus study developed a new medical academic vocabulary list, the MAVL. Based on the results of a series of comparative analyses, the MAVL boasts a much better coverage of medical English while being 53% shorter than the existing medical academic word list developed by Wang, Liang, and Ge (2008). Thus, this new list should better serve the needs of medical English learners. This study also provides evidence for the need to include, in discipline-specific vocabulary lists, general high-frequency words that have a significantly higher frequency and special meanings in the discipline than in general English. Pedagogical and research implications are also discussed.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Academic vocabulary refers to the lexical items that are relatively frequent across a wide range of academic texts but are infrequent in other genres (Coxhead & Nation, 2001). Although it is key to language teaching and learning, academic vocabulary has proved challenging to language learners (Li & Pemberton, 1994; Shaw, 1991; Thurston & Candlin, 1998), either because academic words do not occur as frequently as general-service words do (Worthington & Nation, 1996; Xue & Nation, 1984), or because they are “not likely to be glossed by the content teacher” (Flowerdew, 1993, p. 236), a point also highlighted by Coxhead (2000). As a result, several academic word lists have been developed to facilitate the learning and teaching of general academic vocabulary. For example, Xue and Nation (1984) compiled the University Word List (UWL) of 836 word families from Campion and Elley (1971), Praninskas (1972), Lynn (1973), and Ghadessy (1979). Following the UWL, Coxhead (2000) developed the now well-known Academic Word List (AWL) of 570 word families, whose frequency accounted for

* Corresponding author.

E-mail address: dliu@ua.edu (D. Liu).

approximately 10% of tokens in academic texts. The AWL was also assessed in a wide range of academic disciplines, and results showed that it registered a fairly stable coverage of roughly 10% in various disciplines (e.g. Chen & Ge, 2007; Martinez, Beck, & Panza, 2009; Vongpumivitch, Huang, & Chang, 2009).

However, some scholars (e.g., Hyland & Tse, 2007) have challenged the usefulness of a general academic vocabulary list, such as the AWL, based on corpus research findings that the use of vocabulary varies significantly across disciplines in terms of range, frequency, collocation, and meaning. Accordingly, they argue for the need to develop academic vocabulary lists for various specific disciplines. In recognition of this need, several discipline-specific academic word lists have been developed, such as Hsu's (2014) list of 729 most frequent word families in engineering compiled from engineering textbooks; Mudraya's (2006) Engineering Academic Word List boasting 1200 word families; Wang, Liang, and Ge's (2008) Medical Academic Word List (MAWL) registering 623 word families; and Yang's (2015) Nursing Academic Word List (NAWL) containing 676 word families in the nursing discipline.

One important point to note regarding these discipline-specific word lists is that all of them followed Coxhead's (2000) practice in the development of the general AWL by excluding the general high-frequency words such as those in West's (1953) 2000-item General Service List (GSL) or those in the BNC's first and second 1000 most frequent words. Such a practice seems to have been motivated by the assumption that language learners generally grasp high-frequency words before they learn low- or lower-frequency words and that academic words are of a lower frequency in comparison with general high-frequency words. This assumption may have been influenced by Nation's (2001) classification of words into four levels: (1) high-frequency words, (2) academic words, (3) technical words, and (4) low-frequency words. It is important to note, however, that, recognizing the overlap among high-frequency, academic, and technical words, Nation (2013) has restructured his classification system by adding mid-frequency vocabulary and putting academic and technical vocabulary under the "specialized" category. Based on Nation's original classification, learners should master the high-frequency words (for example, the 2000 most frequent words in the GSL) before they learn academic words. Such a classification and the underlying learning-order assumption have, however, been questioned because several studies (e.g. Cobb, 2010; Neufeld, Hancioglu, & Eldridge, 2011; Gardner & Davies, 2014) have found that some AWL items were among the most frequent words in the British National Corpus (BNC) and Davies's (2008) Corpus of Contemporary American English (COCA), challenging a clear-cut division between the high-frequency words and the academic words based only on frequency.

Furthermore, as Gardner and Davies's study (2014) has shown, many general high-frequency words, such as *interest*, *capital*, and *rate*, have special or different meanings in academic contexts from those in general English. In addition, Gardner and Davies (2014) have also found that some general high-frequency words actually boast a much higher frequency in academic English than in general English. Therefore, they argue that such general high-frequency words should be included in academic word lists.

To address the aforementioned problems in the existing academic word lists including the AWL, Gardner and Davies (2014) adopted a new method to define and identify academic words and used it in developing a new Academic Vocabulary List (the AVL) of 3015 words in the form of lemmas. The reason why Gardner and Davies (2014) used the lemma form to report or list the words in the AVL, rather than the word family form used by Coxhead (2000) in the AWL, is that the lemma form is much more informative and user-friendly than the word family form, for the latter does not indicate the parts of speech of the listed words. Without the parts of speech information, the usefulness of the listed words is limited and questionable. For example, the family head word *abstract* in the AWL could be used in three parts of speech: as an adjective meaning *disassociated from any specific instance*, a verb meaning *to extract/remove or make a summary*, and a noun meaning *the summary of a book or an article*. It is not clear which word class/es of the word *abstract* is/are referenced in the AWL.

According to Gardner and Davies (2014), academic words are those that each occur 50% more frequently in the academic texts than in non-academic texts with even distribution across a wide range of academic disciplines and which do not occur unexpectedly too frequently in any single discipline. Therefore, academic words must be identified by using four systematic rigorous selection criteria. The first is *ratio*: the frequency of an academic word must be at least 50% higher in the academic corpus than in the non-academic portion of the corpus; the second is *range*: it must occur at least 20% of its expected frequency in at least seven of the nine disciplines in the corpus; the third is *dispersion*: the word must have a dispersion with at least 0.8 of the Juilland D value (Juilland, Brodin, & Davidovitch, 1970); the fourth is *discipline measure*: the word cannot occur more than three times the expected frequency in any discipline in order to eliminate discipline-specific and technical words (Gardner & Davies, 2014). However, there is a weakness in Gardner and Davies's method. While the frequency ratio ensures that the given word occurs far more frequently in academic English than in general English, it does not guarantee that the word is actually a highly frequent one in academic English. To make sure the words that meet the frequency ratio requirement are truly highly frequent words in academic English, it seems necessary to employ a measure of minimum frequency, a practice Coxhead (2000) used in her method.

Gardner and Davies (2014) evaluated the representativeness of their AVL in terms of its coverage in academic English texts against its coverage in non-academic texts and also compared its coverages with those of the AWL. The results show that the AVL covers 13.8% and 13.7% of the academic sub-corpora in the BNC and the COCA respectively, much higher than its coverage in the newspaper and fiction sub-corpora (8.0%/7.0% and 3.4%/3.4% respectively), providing solid evidence that words in the AVL are indeed those used mainly in academic English. Regarding the coverage comparison between the AVL and AWL, because, as noted earlier, the two lists are in different forms, Gardner and Davies (2014) first randomly selected 570 word families from the 3015 lemmas in the AVL to make the comparison. The results indicate that the randomly chosen 570 word families from the AVL boast a 14% coverage in the academic sub-corpora of the BNC and the COCA, which doubles the 7%

Download English Version:

<https://daneshyari.com/en/article/360140>

Download Persian Version:

<https://daneshyari.com/article/360140>

[Daneshyari.com](https://daneshyari.com)