A COMPARISON OF THE ACADEMIC WORD LIST AND THE ACADEMIC VOCABULARY LIST: SHOULD THE AVL REPLACE THE AWL?

Razieh Gholaminejad[a], Mohammad Reza Anani Sarab[b]
([a]r_gholaminejad@sbu.ac.ir, [b]anani@sbu.ac.ir)

Department of English Language and Literature
Shahid Beheshti University, Tehran, Iran

Abstract: In this commentary, we begin with the discussion on a brief history of academic wordlists. Adopting a comparative perspective, then, the merits and demerits of the Academic Word List (AWL) (Coxhead, 2000) and its competing counterpart the Academic Vocabulary List (AVL) (Gardner & Davies, 2014) are presented. We also explore whether the AWL can still be considered as "the best list" (Nation, 2001, p. 12) for improving academic words, or whether its counterpart is reasonably "the most current, accurate, and comprehensive list" (Gardner & Davies, 2014, p. 325). The comparison was made in terms of twelve aspects: corpus size, types of corpus texts, sources of corpus texts, text balance, disciplines included, counting unit, wordlist items, method for excluding high-frequency words, minimum frequency, method for excluding technical words, sequence of list items and lexical coverage. The comparison reveals that the AVL is far from complete and cannot replace the AWL. The results of the comparison can have implications for practitioners and course developers.

Keywords: academic vocabulary, academic word, English for Academic Purposes/EAP, word-list, lexical coverage

DOI: http://dx.doi.org/10.15639/teflinjournal.v32i1/167-182

Recent studies demonstrate that knowledge of vocabulary can have "a direct influence on language learners' reading and writing proficiency" (Yang, 2015, p. 27). More specifically, the size and depth of vocabulary knowledge have a significant correlation with performance in writing and speaking (Kilic, 2019). The importance of vocabulary knowledge in language learning has inspired researchers to work on specific layers of lexical knowledge including academic words.

Academic words are defined as the "formal, context-independent words with a high frequency and/or wide range of occurrence across scientific disciplines, not usually found in basic general English courses: words with high frequency across scientific disciplines" (Farrell, 1990, p. 11). The idea of establishing an academic wordlist, as a list of English academic words useful for students at the tertiary level, has a long tradition in the English for Academic Purposes (EAP) (Hyland & Tse, 2007). Although under different rubrics such as *sub-technical vocabulary* (Yang, 1986), *semi-technical vocabulary* (Farrell, 1990), s*pecialized-nontechnical lexis* (Cohen et al., 1988), *non-technical terms* (Goodman & Payne, 1981), and *transdisciplinary lexicon* (Drouin et al., 2018), academic words share the attribute of high frequency across all academic disciplines.

Significant relationships have been found between the use of academic vocabulary and the quality of student writing (Csomay & Prades, 2018). Academic vocabulary knowledge also significantly affects school performance (Schuth et al., 2017). Important for its "strong enabling function" (Malmström et al., 2018, p. 29), academic vocabulary has been considered challenging to language learners (Li & Pemberton, 1994). This can be attributed to the fact that these words are not as specific as technical words (Nation, 2001). Farrell (1990) confirms this by explaining that this is the consequence of the assumption that students already know these words and therefore they do not need explicit instruction. A further reason might be the fact that they do not occur as frequently as high-frequency vocabulary items (Xue & Nation, 1984), leaving learners not as familiar with them as they are with other types of words. Thus, helping non-native learners learn the so-called challenging words has become a concern for EAP researchers. They are also problematic for native students as most of them are originally Graeco-Latin, and mostly refer to abstract ideas, thereby introducing further propositional density to texts (Corson, 1997).

In order to identify the vocabulary which students probably encounter at university, Campion and Elley (1971) developed a wordlist of 500 most common words and 3200 frequently-used words based on textbooks and lectures of 19 academic disciplines. Similarly, Praninskas (1972) developed the American University Wordlist based on a corpus of university-level textbooks of 10 disciplines, while Lynn (1973) and Ghadessy (1979) compiled lists of words assumed to be difficult for academic reading. Xue and Nation (1984) checked these four pioneering wordlists against each other, and found 70%

overlap. They combined them and formed the University Wordlist (UWL). The UWL consists of approximately 800 words supposed to have a wide range in academic texts with coverage of 8% (Nation, 1990).

The UWL was criticized by Coxhead (2000) for its lack of consistent selection principles. It was also argued that UWL included many of the weaknesses of the previous works, and was based on a very small corpus which contained an unbalanced range of topics. In order to address the aforementioned issues, Coxhead (2000) developed the Academic Wordlist (AWL). Since a controversial issue has long been about which academic words are paramount to learn and in what order, Coxhead divided the words on the AWL into ten sub-lists ranked from the most to the least frequent.

It was claimed that the AWL represented words used in academic texts, which could serve as an essential measure of learners' academic competence. Researchers, such as Nation (2015), argue that knowledge of academic words on the AWL, together with the knowledge of the General Service List (GSL) (West, 1953), can provide around 98% coverage for most types of academic texts. This value is supposed to allow unassisted comprehension of texts. That is why Nation (2001, p.12) considers the AWL as "the best list" for improving academic words.

Two decades after the development of the AWL, it is still widely used in many universities across the world. However, Gardner and Davies (2014) criticized the methodology used for development of the AWL, and established a new Academic Vocabulary List (AVL). As a rival, the AVL was situated in a stronger position to become a standard reference for academic vocabulary development. Despite this claim, it is not clear whether in practice the AWL must be substituted by the AVL. Program administrators, EAP teachers, or materials developers may now be in a dilemma in the selection of the best academic wordlist. This article is a comparative attempt to explore whether the AWL can still be considered as "the best list" of academic words, claimed by scholars such as Nation (2001, p.12), or that the AVL is reasonably more accurate, and comprehensive as its developers argue. Despite the fact that the AVL needs some time to find its place in ELT materials, a clear comparison of the two wordlists can provide useful insights for practitioners in making a decision.

COMPARISON OF THE INFLUENTIAL WORDLISTS: AWL VS. AVL

With only 31.05% overlap, more than two thirds of the items on the two lists are mutually exclusive. Nearly 26.87% of the AVL also appear in the GSL (West, 1953), and 42.07% of the AVL is not included in the GSL or in the AWL (Hartshorn & Hart, 2016). In Table 1, the two wordlists are compared in terms of twelve aspects.

Table 1. The AWL vs. the AVL

Aspect	AWL	AVL
1. Corpus Size	3.5 million tokens	120 million tokens
2. Types of Corpus Texts	Textbooks, book chapters, laboratory manuals, journal articles	Academic journals, academically oriented magazines, and newspapers.
3. Sources of Corpus Texts	64% in New Zealand, 20% in Britain, 13% in the USA, 2% in Canada, and 1% in Australia.	100% in the USA
4. Text Balance	Equal numbers of total tokens in each discipline. Also, equal numbers of short texts, medium texts, and long texts.	Not specified
5. Disciplines	28 subject areas organized into 7 general areas within 4 disciplines: Arts, Commerce, Law and Science	9 disciplines: 1) Education, 2) Humanities, 3) History, 4) Social Science, 5) Philosophy, religion, psychology, 6) Law and Political Science, 7) Science and technology; 8) Medicine and Health, 9) Business and Finance
6. Counting Unit	Word Family	Lemma
7. Wordlist Items	570 word families (3,112 items) No familiar item included	3,015 lemmas (1,991 word families) Extremely familiar items included such as ‘study’, ‘group’, or ‘system’.

Aspect	AWL	AVL
8. Method for Excluding High-Frequency Words	The word should not be included in the GSL	The frequency of the lemma must be 50% higher in the academic corpus than in the non-academic one.
9. Minimum Frequency	100 times in the whole corpus	Not specified
10. Method for Excluding Technical Words	Range	1) Range 2) Dispersion 3) Discipline measure
11. Sequence of List Items	Items are grouped into10 sub-lists arranged by frequency of word family, and each sublist is arranged alphabetically.	Items are grouped into an entire list with frequency rank of lemmas from 1 to 3015.
12. Lexical Coverage	10%	14%

Each aspect is discussed below for the two lists.

Corpus Size

The corpus used for development of the AVL is nearly 35 times larger than the AWL one. Thus, the AVL is built on a more representative data in terms of size. Gardner and Davies (2014) argued that the corpus must be large enough in order to be able to separate academic core words from both general high-frequency words and technical words. Hence, they increased their corpus size up to 120 million tokens, taken from the 425-million-word COCA.

Types of Corpus Texts

The corpus used for construction of academic vocabulary should preferably be made up of a broad range of academic genres. For construction of the AWL, various academic genres were used (textbooks, book chapters, laboratory manuals, journal articles), while the corpus used for the AVL only included academic journals, accompanied later by academically-oriented magazines and finance sections of newspapers. Textbooks are among the excellent academic genres, which are missing in the AVL development corpus.

As the target audience of an academic wordlist are EAP learners, particularly undergraduates, samples of texts that are most likely to be encountered by such an audience (i.e. textbooks) needed to also be included (Gholaminejad & Anani Sarab, 2020). Besides, the corpora of neither of the wordlists contained any samples of spoken academic language. Therefore, the lists can be considered *written* academic wordlists.

Sources of Corpus Texts

While the AVL development corpus only included texts published in the USA, the AWL development corpus comprised texts from different varieties of academic English published in a number of English-speaking countries. Although the disproportionate share of texts from New Zealand tilts the balance in the AWL, it is based on a corpus of texts published in a variety of different places. Thus, the representativeness of the AVL development corpus can be questioned, as the resulting wordlist can be considered an *American* academic wordlist.

Text Balance

Unlike the AWL development corpus for which equal numbers of short, medium, and long texts as well as equal numbers of tokens for each discipline were selected, the AVL development corpus neglected this critical feature. In addition, the AVL development corpus also contained strikingly different numbers of tokens for each individual discipline. For instance, Education included 8,030,324, Social sciences 16,720,729, and Science and technology 22,777,656 tokens. This lack of balance, according to Coxhead (2000), can increase the bias from word repetition within longer texts. Moreover, the type of genres included in the sub-corpora of the AVL had little balance. While the sub-corpora for Education and Humanities were based on journals exclusively, the ones for Business and Finance were derived from magazines and newspapers and for the rest a combination of journals and magazines were used. Since genres typically vary in terms of lexical density, this imbalance may have affected the results.

Disciplines

The AWL development corpus has been criticized for its biased composition. Hyland and Tse (2007) argue that Coxhead's selection of the texts in her corpus was "opportunistic" (p. 239), since it has not given full and

equal coverage of the range of academic disciplines. Accordingly, the AWL items do not occur with the same frequency across disciplines. As they noted, this has brought about the above-average coverage of 12% in commerce. The "inclusion of disciplines which shared greater similarities in her commerce corpus" led to the "remarkably high frequencies of words in the AWL common to finance-oriented disciplines", while there were dissimilar disciplines in the areas of arts and sciences (p. 248). This raises questions about the usefulness of the list for *all* disciplines. Hyland and Tse (2007) found that the distribution of the AWL words was not even across different disciplines, and accordingly, concluded that all of the words on the AWL are not of equal value to all students, and some words may be of no use to them at all.

As for the AVL, the selection and classification of the disciplines seem to be arbitrary. Psychology is grouped together with philosophy and religion, while history stands alone. Similarly, Humanities, Education, and Social Sciences are three distinct categories, while Science and technology are grouped together. Such an arbitrary decision may have introduced a bias in results. Thus, the criticism against the AWL composition can also be leveled at the AVL. In fact, the scope of neither of the corpora provides a full coverage of sufficient number of disciplines.

Counting Unit

The word-family approach was the counting unit for the development of the AWL. It assumes that knowledge of a root can enable understanding of its derived forms (Nation, 2001). Using this approach has been challenged, on the basis that word families often include members with extremely diverse meanings. For instance, members of a word family like 'react' or 'constitute' do not share the same core meaning. It is misleading to assume that a student who knows one family member knows all the others. Not only the inflected and derived forms of a word add considerably to its learning load, but also students usually cannot make connections between even quite closely-related forms of a headword (Gholaminejad & Anani Sarab, 2020; Schmitt & Zimmerman, 2002). Adopting a word-family approach also restricts the usefulness of the wordlist only to receptive vocabulary learning. However, the productive use of vocabulary requires more knowledge as the link from meaning to form is harder to establish than form to meaning (Durrant, 2014). Moreover, when reading, many of the coping strategies which help students make up for lacks in

knowledge are not available to them when they are writing. This would make the process of production even more difficult.

To tackle such problems, lemmas were used to construct the AVL, with the justification that knowledge of inflectional word relationships precedes the derivational one (Gardner, 2007). Whereas word families refer to headwords plus their inflectionally- and derivationally-related forms, lemmas denote headwords and their inflectionally-related forms only. Although the lemma approach can decrease some of the semantic issues of words, it fails to resolve them entirely. That is, even if the part of speech of a word such as 'process' is clarified in the lemma approach; it is not still clear for which meaning it has been selected. Besides, further research is required to investigate whether the lemma approach can, in practice, facilitate productive vocabulary development. Malmström et al. (2018) contend that when an academic wordlist is based on scholarly articles or textbooks, the wordlist usually tends to include those words which students should know *receptively*, that is to say, such wordlists are intended to enable students to *read* academic texts. Thus, using a lemma approach, by itself, cannot promise creation of a wordlist suitable for productive purposes.

Besides, compared to lemma, the word-family approach is more economical, equipping learners with knowledge of a large number of words by just learning a headword. Although it is inevitably true that family members may not share the same meaning, a short glance at the family members within the AWL manifests that the majority of the members in each family share transferrable meanings. With moderate burden of items to learn, word-family approach seems to offer more gains than losses, in terms of the required amount of learning.

Wordlist Items

The AWL was criticized by Ward (2009) for being too long for practical use. The AVL, on the other hand, consists of more than three times the number of the items on the AWL. While there are 5.46 words per family on the AWL, the AVL includes 1.76 words per family (Hartshorn & Hart, 2016). This means that an AVL learner has to undergo five times as much workload as an AWL learner does, as learning is facilitated by the transferability of meaning throughout family members. Not only does this impose a huge burden on the AVL learners, but also inclusion of extremely familiar items, found in the

26.87% overlap shared by the GSL, can make the list too tedious for the AVL learners.

Method for Excluding High-Frequency Words

Coxhead adopted a stopword approach, which refers to excluding a set of specified words from analysis. The stopword approach fits well with Nation and Waring's (1997) proposal that specialist vocabulary lists should be constructed on top of an initial knowledge of the most frequent general words. However, Gardner and Davies (2014) criticized this idea, explaining that the stopword list may include words which have specific academic uses or meanings. For instance, Durrant (2009) shows that a word such as 'address' has a general use as a noun, but an academic role as a verb. Besides, Gardner and Davies (2014) demonstrated that some general words have a much higher frequency in academic texts than in general texts. This implies that such general words ought to be included in academic wordlists. However, not using a list of stopwords has culminated in the extreme length of the AVL and inclusion of familiar words.

Minimum Frequency

The AWL has been criticized for the low threshold of word frequency necessary to be selected, by Hyland and Tse (2007). They considered items as frequent "if they occurred above the mean for all academic word items in the corpus" (p. 240). This measurement rendered 192 families in their study, which is much shorter than Coxhead's, with 570 families.

The AVL was, on the other hand, criticized by Lei and Liu (2016) for determining no minimum frequency. While the 50% higher frequency ratio ensures that the given word appears more frequently in the academic corpus than in the non-academic one, it does not guarantee that the word is actually a highly frequent one in the academic corpus (Lei & Liu, 2016). The existence of infrequent items on the AVL is also supported by a study conducted by Durrant (2016) who showed that only a small core of 427 items on the AVL was frequent across 90% of disciplines under study.

Method for Excluding Technical Words

The range criterion used for the AWL required that the word appear in at least half of the corpus. However, Gardner and Davies (2014) required that the lemma occur with at least 20% of the expected frequency in seven of the nine

disciplines. According to Lei and Liu (2016), Gardner and Davies's range ratio is much more rigorous since it not only has a higher ratio (78% vs. Coxhead's 50%) but also requires the lemma to occur at least 20% of the expected frequency, rather than simply to just occur.

Also, for the development of the AVL the measure of 'dispersion' was used to ensure selection of items with an even distribution, as well as the criteria of 'discipline measure' to exclude even more discipline-specific words (Gardner & Davies, 2014). Here the question arises as to whether these strict additional statistics resulted in more gains or losses. In addition, considering the huge number of words such as 'albeit', 'accommodate', 'accompany', 'amend', 'annual', 'bulk', 'behalf', and 'cease' that are included in the AWL but ended up being excluded from the AVL, further research is required to explore the necessity of utilizing these statistics with the selected thresholds. As a suggestion for future studies, researchers can also ask learners/teachers to rate the usefulness of these words in practice.

Sequence of the List Items

To help sequencing of vocabulary teaching, Coxhead (2000) believed that the ideal wordlist should be divided into smaller, frequency-based sub-lists. This was criticized by Hartshorn and Hart (2016), who explained that the sequencing of the AVL items provides frequency information, which is lost in the AWL. They note that keeping the frequency information matches the natural patterns of acquisition of words. They showed that the procedures in the creation of the AWL may have diluted frequency differences from one sub-list to the next, giving the AVL an advantage in this regard.

Lexical Coverage

Gardner and Davies (2014) compared 570 randomly-selected AVL word families with the AWL word families using the academic sub-corpora of the COCA and the BNC. The result was that the AVL word families achieved higher coverage (13.8%-13.7%) than the AWL ones (6.9%-7.2%) in the two corpora. Therefore, they claimed that the AVL was a better wordlist.

Lei and Liu (2016) considered this comparison questionable because the conversion of the AVL lemmas to word families could have led to including, in the word families, lemmas that were actually not in their AVL, thus inflating the number of items of the list. They added that a reason for the AWL's lower coverage is that the AWL included all the members of the word families which

occurred in the corpus used in its development without applying the selection criteria such as frequency and range for each member. If the same selection criteria applied to the selection of the headword of a word family are also applied to the selection of each of its members, wordlists developed by using the word family method may attain the same level of coverage that wordlists developed by the lemma method have been able to reach.

In addition, Qi (2016) ascribed the AVL's higher coverage to the distributional differences of the two lists at different word-frequency levels. As the AVL was not built on any stopword list, the frequency of AVL items begins at a much higher threshold than the AWL, and because of setting no criteria for minimum frequency, it ends at a lower frequency than the AWL. The more high-frequency items a wordlist contains, the higher coverage it may achieve. Thus, without identifying items at each word-frequency level for each wordlist, one cannot make a fair comparison. The AVL's higher coverage might be the consequence of containing some GSL words that the AWL had excluded.

Another reason for the unfairness of the comparison is that the intended audience of the two wordlists are academic English learners. This requires a type of corpus which is derived from the material used by target audience. However, the already available online corpora (the BNC academic sub-corpus and the COCA academic sub-corpus) were used by Gardner and Davies. Sample texts in neither BNC nor COCA are potentially prevalent in "real educational settings," which renders the comparison result invalid (Qi, 2016, p. 23).

CONCLUSIONS

This article is aimed to compare the AWL and the AVL in order to facilitate the teachers' decision-making for the selection process of the suitable academic wordlist. It was shown that the AWL has been criticized by various scholars, and the AVL is subject to different criticisms. In summary, the corpus for the AVL has the advantage of being far larger than that of the AWL. However, it is limited by the inadequate types of corpus texts and genres, and does not include any samples of textbooks or spoken academic language. Besides, it is based on undiversified sources of corpus texts, as it only includes texts published in the USA. This has, in a sense, rendered the final wordlist an American one. The imbalance among the sub-corpora may also have caused

bias from word repetition within longer texts. The corpus also suffers from arbitrary selection and classification of the disciplines. Therefore, the final wordlist may not be equally useful for students of all disciplines. It seems as if Gardner and Davies neglected Coxhead's (2011) recommendation that "Future research needs to be based on more balanced corpora that represent a wider range of subjects" (p. 357). Another issue concerns the lemma counting unit, which is uneconomical in terms of the required amount of learning time and cannot guarantee a wordlist suitable for productive purposes. Although avoiding the use of any stopword list for the development of the AVL has advantages, it has led to inclusion of familiar items. The extreme length of the AVL imposes a huge workload on learners. Using no minimum frequency for the AVL is also another issue which leads to inclusion of words which are not frequent. Among the positive points of the methodology used for the development of the AVL are the sequencing of items and the method used for excluding technical words, although the latter needs further research to attain advisability. Finally, the AVL's higher coverage, which is described by Gardner and Davies to be "nearly twice the coverage as the AWL" (Gardner & Davies, 2014, p. 323) is strictly subject to criticism.

Regarding the pedagogical implications of the study, the AVL, with a corpus which is devoid of any samples of textbooks, does not seem to be an ideal resource for teachers to use at the tertiary level. University students are mostly exposed to textbooks and are more likely to encounter this type of register. Hence, they need to study an academic wordlist derived from a corpus of textbooks. Furthermore, the AVL is not advisable to be used by EAP instructors for students of different disciplines in the same way, for the reason that the corpus of the AVL is not based on diversified sources of texts with balance among the sub-corpora or the disciplines included. Durrant (2016) supports this claim by demonstrating that the AVL is more relevant to some students than others. Finally, EAP teachers who are interested to use the AVL as a resource for students are recommended not to use the *whole* list, considering the extreme length of the AVL. While studies, such as those conducted by Masrai and Milton (2018), demonstrate that the majority of the words from the AWL fall within the 3000 most frequent words, about half of the items on the AVL have very little use (Durrant, 2016). In fact, EAP teachers can select which items of the list to focus on depending on the context, discipline, and proficiency of the students.

All in all, both the AWL and the AVL seem to be far from complete. It is not surprising then that the AVL has not been able to supersede the AWL in spite of its developers' strong claims about its superiority. Further studies are required to examine how significant the variations in these two wordlists are in practice for academic vocabulary development.

REFERENCES:

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

Cohen, A., Glasman, H., Rosenbaum-Cohen, P. R., Ferrara, J., & Fine, J. (1988). Reading English for specialised purposes: Discourse analysis and the use of standard informants. In P. Carrell, J. Devine, & D. Eskey (Eds.), *Interactive approaches to second language reading* (pp. 152–167). Cambridge University Press.

Corson, D. (1997). The learning and use of academic English words. *Language Learning, 47*(4), 671–718. https://doi.org/10.1111/0023-8333.00025

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238. https://doi.org/10.2307/3587951

Coxhead, A. (2011). The academic word list 10 years on: Research and teaching implications. *TESOL Quarterly, 45*(2), 355–362. https://doi.org/10.5054/tq.2011.254528

Csomay, E., & Prades, A. (2018). Academic vocabulary in ESL student papers: A corpus-based study. *Journal of English for Academic Purposes, 33*, 100-118. https://doi.org/10.1016/j.jeap.2018.02.003

Drouin, P., L'Homme, M. C., & Robichaud, B. (2018). Lexical profiling of environmental corpora. In N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis, & T. Tokunaga (Eds.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)* (pp. 3419–3425). European Language Resources Association (ELRA).

Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes, 28*, 157-169. https://doi.org/10.1016/j.esp.2009.02.002

Durrant, P. (2014). Discipline and level specificity in university students' written vocabulary. *Applied Linguistics, 35*(3), 328-356. https://doi.org/10.1093/applin/amt016

Durrant, P. (2016). To what extent is the Academic Vocabulary List relevant to university student writing? *English for Specific Purposes, 43,* 49-61. https://doi.org/10.1016/j.esp.2016.01.004

Farrell, P. (1990). *Vocabulary in ESP: A lexical analysis of the English of electronics and a study of semi-technical vocabulary* (CLCS Occasional Paper No. 25). Trinity College, Centre for Language and Communication Studies.

Gardner, D. (2007). Validating the construct of word in applied corpus-based vocabulary research: a critical survey. *Applied Linguistics, 28*(2), 241–65. https://doi.org/10.1093/applin/amm010

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics, 35*, 305-327. https://doi.org/10.1093/applin/amt015

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

Gholaminejad, R., & Anani Sarab, M.R. (2020). Academic vocabulary and collocations used in language teaching and applied linguistics textbooks: A corpus-based approach. *Terminology, 26*(1), 83–108. https://doi.org/10.1075/term.00043.gho

Goodman, A., & Payne, E. (1981). 'A taxonomic approach to the lexis of science'. In L. Selinker, E. Tarone, & V. Hnazeli (Eds.), *English for academic and technical purposes: Studies in honour of Louis Trimble* (pp. 23–39). Newbury House.

Hartshorn, K. J., & Hart, J. M. (2016). Comparing the academic word list with the academic vocabulary list: Analyses of frequency and performance of English language learners. *The Journal of Language Teaching and Learning, 6*(2), 70-87.

Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary"? *TESOL Quarterly, 41*(2), 235-253. https://doi.org/10.1002/j.1545-7249.2007.tb00058.x

Kilic, M. (2019). Vocabulary knowledge as a predictor of performance in writing and speaking: A case of Turkish EFL learners. *PASAA: Journal of Language Teaching and Learning in Thailand, 57*, 133-164.

Lei, L., & Liu, D. (2016). A new medical academic word list: A corpus-based study with enhanced methodology. *Journal of English for Academic Purposes, 22*, 42-53. https://doi.org/10.1016/j.jeap.2016.01.008

Li, E. S. L., & Pemberton, R. (1994). An investigation of students' knowledge of academic and subtechnical vocabulary. In *Proceedings of the joint seminar on corpus linguistics and lexicology* (pp. 183-196). Hong Kong University of Science and Technology.

Lynn, R. E. (1973). Preparing word lists: A suggested method. *RELC Journal, 4*, 25-32. https://doi.org/10.1177%2F003368827300400103

Malmström, H., Pecorari, D., & Shaw, P. (2018). Words for what? Contrasting university students' receptive and productive academic vocabulary needs. *English for Specific Purposes, 50*, 28-39. https://doi.org/10.1016/j.esp.2017.11.002

Masrai, A., & Milton, J. (2018). Measuring the contribution of academic and general vocabulary knowledge to learners' academic achievement. *Journal of English for Academic Purposes, 31*, 44-57. https://doi.org/10.1016/j.jeap.2017.12.006

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Newbury House.

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

Nation, I. S. P. (2015). Chapter 33: Which words do you need? In J. R. Taylor (Ed.), *Handbook of the word* (pp. 568-581). Oxford University Press.

Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge University Press.

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

Qi, H. (2016). *A corpus-based comparison between the Academic Word List and the Academic Vocabulary List* [Master's thesis, University of Western Ontario, London, Ontario, Canada]. Electronic Thesis and Dissertation Repository. 3721. https://ir.lib.uwo.ca/etd/3721

Schmitt, N., & Zimmerman, C. B. (2002). Derivative word forms: What do learners know? *TESOL Quarterly*, *36*(2), 145-171. https://doi.org/10.2307/3588328

Schuth, E., Köhne, J., & Weinert, S. (2017). The influence of academic vocabulary knowledge on school performance. *Learning and Instruction, 49*, 157-165. https://doi.org/10.1016/j.learninstruc.2017.01.005

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

West, M. (1953). *A general service list of English words.* Longman.

Xue, G., & Nation, I. S. P. (1984). A university word list. *Language Learning and Communication, 3*(2), 215–299.

Yang, H. (1986). A new technique for identifying scientific/technical terms and describing science texts. *Literary and Linguistic Computing, 1*(2), 93-103. https://doi.org/10.1093/llc/1.2.93

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