Accounting Academic Word List (AAWL): A Corpus-Based Study

Reza Khany
Department of English Language and Literature, Faculty of Literature and Human Sciences, Ilam University, Ilam, Iran
r.khany@ilam.ac.ir

Behrooz Kalantari*
Department of English Language and Literature, Faculty of Literature and Human Sciences, Ilam University, Ilam, Iran
b.kalantari@ilam.ac.ir

Abstract

The aim of this study was threefold: it aimed to develop a field-specific academic word list for accounting, to find the degree of coincidence between the word list and Coxhead's academic word list (AWL), and also to compare the occurrences of the most frequently used words in the list with six available word lists in different disciplines. A large corpus of accounting research articles was compiled and analyzed. We recognized 658 academic word families with the highest frequency in the corpus which we called Accounting Academic Word List (AAWL**)**. These 658-word families accounted for 10.16 % of the whole corpus. Further analysis indicated that out of these high-frequency word families we identified, only 354 coincided with those listed in AWL. Moreover, 50 most frequently used words in the list accounted for 3.98 % of the whole corpus. These words appeared in six available word lists in different disciplines with different degrees of occurrences which is a starting point for the development of a composite word list. Generally, this study confirmed the significance of subject-specificity of corpus-based word lists. The findings of this study suggest that AAWL can be used as a reference for the accounting community.

Keywords: Academic Word List; Accounting Academic Word List; corpus analysis; Accounting Students; English for Academic Purposes

Received: 2021-01-18 Accepted: 2021-03-10
Available Online: 2021-03-15 DOI: 10.22034/efl.2021.268643.1070

*Corresponding author

1. Introduction

Because of the vital role of vocabulary in reading and writing academic texts, the knowledge of vocabulary seems to bring about educational achievement. Also, the acquisition of vocabulary is an indispensable component of students' academic competence. Therefore, academic word and collocation lists are widely considered to be important in education especially within the domain of language learning, particularly English for Academic Purposes (ESP) (e.g., Campion & Elley, 1971; Coxhead, 2000; Dang, 2018; Deveci, 2019; Ghadessy, 1979; Jablonkai, 2020; Jahangard 2007; Khani & Tazik, 2013; Lynn, 1973; Martinez, Beck & Panza, 2009; Praninskas, 1972; Valipouri & Nassaji, 2013; Wang, Liang, & Ge, 2008; Ward, 2009; Xue & Nation, 1984; Yang, 2015). It has been argued that academic word lists can be classified into two types: "general academic word lists and field-specific academic word lists" (Liu & Han, 2015, p. 1). General academic word lists are composed of words related to various disciplines (Campion & Elley, 1971; Coxhead, 2000; Gardner & Davies, 2014; Ghadessy, 1979; Lynn, 1973; Praninskas, 1972; Xue & Nation, 1984). These words can be acquired and used by the majority of ESP students as a prerequisite for their university studies.

Having combined four previously developed word lists, Xue and Nation (1984) established a comprehensive University Word List (UWL). The list contained 800 high-frequency non- General Service List (GSL) words across various disciplines. This word list was used for more than fifteen years and gained considerable attraction at that time. The word list was criticized by Coxhead (2000). In her seminal article, she mentioned the need for a more comprehensive academic word list and proposed Academic Word List (AWL). Coxhead explained that UWL lacked consistent selection principles and the corpora were small and did not contain a balanced range of topics. She emphasized a need for an academic word list based on the data gathered from a large, well-designed corpus of academic English. Her corpus contained 3.5 million words in 28 sub-disciplines of four main disciplines of Art, Commerce, Law and, Science. AWL consists of 570 word families which has become a standard and appropriate vocabulary list in English language education for many years, but several studies have found some weaknesses of AWL since its advent (e.g., Hyland & Tse, 2007; Gardner & Davies, 2014).

Gardner and Davies (2014) developed an academic vocabulary list (AVL) while raising some concerns regarding the use of AWL. They draw our attention to two aspects of AWL that seem to be problematic: "the use of word families to determine word frequencies and the relationship of the AWL with West's (1953) General Service List (GSL)" (p. 3). By pointing out some key considerations for AVL such as using lemmas, not word families and including different academic disciplines, they established an AVL which derived from a 120-million-word

academic subcorpus of the 425-million-word Corpus of Contemporary American English (COCA; Davies, 2012). The corpus was almost 35 times larger than Coxhead's AWL corpus. In order to create the AVL, they used four criteria to distinguish the academic core including Ratio, Range, Dispersion, and Discipline Measure in which Ratio helps to exclude general high-frequency words from an academic 'core', while Criteria of Range, Dispersion, and Discipline Measure help to exclude technical words and words that occur mainly in one or two disciplines. They concluded that the AVL discriminated between academic and other materials and that it covers 14% of academic texts in both COCA and the British National Corpus.

The above-mentioned studies aimed to propose core academic words in their lists, but Hyland and Tse (2007) described that identifying core academic words from different academic disciplines can be questionable because these "lexical items often occur and behave in different ways across disciplines in terms of range, frequency, collocation, and meaning"(p. 235). Therefore, it seems essential to produce field or discipline-specific words to understand academic discourses. Field-specific academic word lists include words commonly found in different subject areas of a particular discipline (e.g., Beck & Panza, 2009; Khani & Tazik, 2013; Martinez, Valipouri & Nassaji, 2013; Wang, Liang, & Ge, 2008; Yang, 2015).

Medical academic word list (MAWL) was created by Wang, Liang, and Ge (2008). Eliminating GSL word families, MAWL contained 623 word families, which accounted for 12.24% of their corpus. Ward (2009) used a 271,000 word-corpus in order to develop a word list for foundation engineers. He intended to have a word list that could be utilized by low level learners of English and used by all disciplines of engineering. The word list was named Basic Engineering List (BEL) consisting of 229 words. He criticized that high school education does not equip engineering students with the ability to read English language textbooks in colleges or universities. Martinez, Beck and, Panza (2009) conducted a study in order to identify the academic words in a corpus of agriculture research articles (RAs). By adopting both quantitative analysis and qualitative observations, they found that the GSL (67.53%) and the AWL (9.06%) provided a cumulative coverage of 76.59% for the whole corpus. In their study the idea of specificity in EAP in general and the specificity of the meanings and behaviors of the words in particular has been emphasized. An academic word list for applied linguistics including 773 academic word types was developed by Khani and Tazik (2013). Of 773 words, 573 found in Coxhead's (2000) AWL. They described that GSL and AWL covered 88% of tokens in their corpus. Valipouri and Nassaji (2013) conducted a study to analyze a corpus of 1,185 chemistry RAs including 4 million words from different chemistry subject areas. They identified 1400 academic word families. They explained that, of 1400 word families, 327 overlapped with word families in AWL which provided coverage of 9.60% of the tokens in their

corpus. Out of 1400 words families, 390 used frequently in chemistry RAs which has not been in GSL and AWL.

Yang (2015) analyzed 252 English nursing research articles to find out the most frequently used nursing academic vocabularies. He established a Nursing Academic Word List including 676 word-families which accounts for 13.64% of the coverage in the nursing research articles. He stressed the necessity for generating field-specific academic word lists for EFL nursing students to strengthen their academic reading and writing proficiency.

Most aforementioned studies mainly used a text analysis approach for their purposes. The significance of texts analysis in EAP is associated with the idea that the texts used in particular specialist environments have particular characteristics that distinguish them from other texts and from the generalized summaries of linguistic features that arise from an approach to text analysis that uses a corpus of differing texts (Dudley-Evans, 1994).

Research articles as academic texts can be considered to be an important source of academic and technical information in order to investigate vocabularies related to specific fields of study. Specialists as actual writers of research articles try to use field-specific vocabularies in order to convey their messages to the readers and to publish their articles in well-known journals. This has persuaded researchers to make attempts not only to develop and establish word lists for different purposes including general, academic, or technical but also to present models and frameworks for identifying and categorizing words. For these reasons, accounting students and professors need to have a word list containing specific vocabularies related to the research articles of their field. To date, no study has been done to analyze accounting research articles to develop a word list.

One main finding among above-mentioned studies is that vocabulary can help increase students' advancement in academic writing in their area of study and lack of vocabulary can really affect the quality of their writing. The variability of vocabulary use can be attributed to the variability of academic disciplines. In other words, various disciplines have some words with a high frequency and a wide range of occurrences which are not usually found in basic general English texts (Farrell, 1990 as cited in Liu and Han, 2015). That is, the meanings of most of these words are closely related to a particular subject area (Valipouri & Nassaji, 2013). This is the main requirement for developing field-specific academic word lists and also a few studies have been done to identify technical vocabulary in different disciplines (Chung & Nation, 2003, 2004; Kwary, 2011; Ha & Hyland, 2014). Thus, EAP practitioners should work closely with special vocabulary to gain an understanding of discourses and courses related to their disciplines which confirms the idea of specificity of EAP vocabularies and texts (Dudley-Evans, 1993; Hyland, 2006). It is also challenging to adhere to general academic word lists such as AWL as versatile lists suitable and useful for all levels and

disciplines. Therefore, every field such as should have its own specific academic word list. Accounting as a practical and helpful field does not have a specific academic word list and attempts were made in the present study to establish such as word list.

2. The present study

Accounting students especially graduates need to read and write research articles written in English in order to be aware of the latest development and research trends in the field. One prerequisite to this awareness is to know words that are common in accounting research articles. Additionally, there are some words such as *disclosure*, *assets*, *forecast*, *cash*, *equity* which are frequently used in accounting research articles that are not in any general or field-specific academic word lists. As previously mentioned, many word lists have been developed for several specific disciplines, for example, applied linguistics; medical academic word list (MAWL); chemistry academic word list (CAWL); nursing academic word list (NAWL); engineering English word list; academic vocabulary in agriculture.

Regarding accounting, no one to the best of our knowledge has developed a word list specific to the field of accounting. To make sure, the researchers conducted interviews with three professors teaching English for accounting at three universities in Iran. They confirmed that there is no comprehensive accounting academic word list available for their students to be able to use it to read accounting research articles and to use it as a guide to write papers in English. Thus, an accounting academic word list is of practical significance. Accordingly, this research article is an attempt to develop a more restricted, discipline-based lexical repertoire of accounting. In developing the word list, we have tried to answer the following questions.

1. What are the most frequently used academic words in the accounting research articles?
2. How many academic words in accounting research articles coincide with those of AWL?
3. Which of the most frequent words identified in this study appeared in six various word lists across different fields?

The rationale behind the third question is that comparing different word lists from different fields of study to find out common words can lead to designing a composite word list useful for many disciplines. That is, the purpose of this investigation is to compare and contrast many word lists in order to determine the degree of overlap among various word lists (Yorkston, Dowden, Honsinger, Marriner, & Smith, 1988).

3. Methodology

3.1. The corpus

For the study, the corpus consisted of a particular genre (research article)-2,098 accounting research articles were compiled. The criteria for building the specialized corpus were determined considering the specialty of the texts, their sizes, representativeness, and dates of publication (Liu & Han, 2015; Sinclair, 1991, 2005 as cited in Liu & Han, 2015). The corpus was specialized texts on accounting and provided information about vocabulary typically used in the field. Moreover, a very large corpus was used including 29,441,460 words. Finally, to achieve an acceptable level of representativeness of the corpus, the research articles written by qualified researchers and experts and published in top accounting journals from 2008 to 2017 were used.

Based on the purpose of our study, a large corpus was gathered to make sure of a reasonable number of occurrences of academic words. As Coxhead (2000) pointed out, "more language means that more information can be gathered about lexical items and more words in context can be examined in depth" (p. 216). All the research articles were downloaded from the five most influential journals in academic accounting. Bonner, Hesford, Van der Stede, and Young (2006) found that these five journals *Accounting, Organizations and Society, Contemporary Accounting Research, Journal of Accounting and Economics, Journal of Accounting Research, and The Accounting Review journal* rank consistently as the top journals in the field. The articles in *The Accounting Review journal* were not accessible from our university (Ilam University) at the time of the study, thus this journal was replaced by *Review of Accounting Studies Journal* based on expert view. The information regarding each journal is illustrated in Table 1. All the articles were collected from the journals published from 2008 to 2017 from the journals databases. The rationale behind selecting these journals was related to the idea of representation to include texts written by a variety of writers to neutralize bias that may result from the idiosyncratic style of one writer (Atkins, Clear, & Ostler, 1992; Sinclair, 1991, cited in Coxhead, 2000). Additionally, increasing the number of lexical items in the corpus maximize the degree of representativeness of the corpus (Sutarsyah et al., 1994, cited in Coxhead, 2000). As Bonner et al. (2006) mentioned, various specialty areas of accounting were among these five journals and "the proportion of articles devoted to the specialty areas of accounting has remained relatively constant" (p. 683).They mentioned the proportion of the articles related to four subject areas of accounting including financial accounting, management accounting, systems and tax, and auditing. For example, articles related to financial accounting appear in high numbers for all journals except *Accounting, Organizations, and Society* and auditing articles emerge in all journals except in the *Journal of Accounting and Economics*. It seems that articles in all subject areas can be seen at least in three of these five

journals. In other words, the research articles in different subject areas of accounting are the focus of analysis.

Table 1

Number of Articles and Words in Each Journal

journals	Number of articles	Number of words
Accounting, Organizations and Society	409	5898593
Contemporary Accounting Research	511	6665233
Journal of Accounting and Economics	489	6850905
Journal of Accounting Research	325	5239783
Review of Accounting Studies	364	4786946
total	2098	29441460

3.2. Word Selection Criteria

A lot of studies on academic vocabulary utilized word families as the unit of analysis (e.g. Coxhead, 2000; Liu & Han, 2015; Valipour & Nassaji, 2013). A word family consists of a base word and all its derived and inflected forms that can be understood by a learner without having to learn each form separately" (Bauer & Nation, 1993, p. 253). As Nagy, Anderson, Schommer, Scott, and Stallman (1989) mentioned, word families are an important unit in the mental lexicon. If learners know a base word, learnig its common inflected or derived members of the family does not involve much more effort. Additionally, Range (the program used to analyze the corpus) includes three ready-made word lists that are the first and second 1000 most frequent words in the GSL and AWL are in word families. Therefore, in this study the words are classified as word families based on Level 6 of Bauer and Nation's (1993) Scale.

Following Coxhead (2000), three criteria (specialized occurrence, frequency and range) were adopted. For specialized occurrence, the word families included in AAWL had to be outside the first 2000 most frequently occurring words of English of GSL West (1953). Furthermore, we assigned a cut-off frequency and range procedure in which each word should occur equal to or higher than 839 times in the whole corpus and for range, each word should appear at least 50 times in each journal.

3.*3. Procedure*

For analysis, the corpus composed of 29,441,460 running words from top influential journals of academic accounting was given to Range (Heatley, Nation,

& Coxhead, 2002) to have a word list. The software can be downloaded and used for free (https://www.wgtn.ac.nz/lals/resources/vocabulary-analysis-programs). It is noted (Nation, 2005, p. 2) that Range can be adopted for several purposes "to find the coverage of a text by certain word lists, to create word lists based on frequency and range, and to discover shared and unique vocabulary in several pieces of writing". After downloading the research articles from the journals databases, the downloaded files which were in pdf format, were converted into txt files. This is necessary as a corpus program (Range) can only analyze data in the form of txt. Additionally, aspects unrelated to the lexical analysis, charts, diagrams appendices, bibliographies, equations, abbreviations, function words, articles, propositions, and symbols were removed for the analysis. Txt files were entered into the software to find out the number of occurrences of each word, its range (i.e. how many texts the word occurs in) plus the words shared with the AWL and GSL word lists. Rang was run and the output consisted two main parts namely a table showing the number of words shared with the AWL and GSL and a word list with frequency and range of each word in the corpus. Then, those words below the cut-off frequency and range procedure were deleted. Also, function words such as pronouns, prepositions and numbers were excluded from the list. Finally, by removing all GSL words from the list, AAWL was developed. Finally, to compare the word list with different word lists, six well-known word lists were selected based on the expert views in different fields. They were Coxhead's (2000) AWL, Wang, Liang and Ge's (2008) medical academic word list, Ward's (2009) engineering English word list, Martinez, Beck, and Panza's (2009) Academic vocabulary in agriculture, Valipouri and Nassaji's (2013) chemistry academic word list, and Khani and Tazik's (2013) academic word list for applied linguistics.

4. Validating the List

To make sure that the AAWL is appropriate and to determine that the list is more useful than AWL for the field of accounting, a validating test was performed. Coxhead (2000) mentioned that "the real test is how the list covers a different collection of similar texts"(p. 224). Thus, a small-sized validating corpus containing accounting research articles was compiled to investigate the AWL's coverage and that of the AAWL in the new texts.

First, three accounting professors were requested to choose randomly 30 accounting research articles from those five academic accounting journals, but the articles were different from those in our corpus regarding the years of publication. To be exact, the research articles in our corpus belonged to the years 2008 to 2017, while the validating corpus was selected from the articles published in 2018. The corpus contained 356,523 running words. As Liu and Han (2015) mentioned, the same criteria for selecting, collecting, and processing the validating corpus should be applied to build the specialized corpus to ensure the consistency of the corpus structure. Besides, "Texts in the validating corpus

should be different from those in the original corpus to guarantee the reliability of the testing results" (Liu and Han, 2015, p. 9). Therefore, the AWL and AAWL were validity tested by comparing their coverage in the validating corpus which is shown in Table 2.

Table 2

The AWL's Coverage and the AAWL's Coverage in Validating Corpus

	The AWL's coverage	the AAWL's coverage	Difference
Validating corpus	11.37%	4.08%	15.42

Although, AAWL coverage was a little lower than that of AWL in some individual articles, on average as Table 2 indicated, the AAWL covered the validating corpus better that did the AWL. The AAWL's coverage of the validating corpus is 15.42% while the AWL's coverage of the corpus is 11.37%.As Coxhead (2000) pointed out, "A frequency-based word list that is derived from a particular corpus should be expected to cover that corpus well" (p. 224). Therefore, it can be concluded from the difference (4.08%) in coverage that, the AAWL seems to be more useful than the AWL in the field of accounting.

5. Results and Discussion

This study was set out to develop an academic word list for the field of accounting, to find the degree of coincidence between the word list and the AWL, and to compare the occurrences of 50 most frequent words in the list with six available word lists in different disciplines. As mentioned previously, all words with frequencies of 839 or higher in the corpus and range of 50 in each journal were regarded as frequent. All words below these criteria were excluded; therefore, the remaining words were 3,172. After removing aspects unrelated to the lexical analysis, we had 2,129 words.

Table 3

The Coverage by the Different Kinds of Vocabulary in the Corpus

WORD LIST	TOKENS	% of the whole corpus	content words	AAWL
GSL	8,711,440	29.58%	1,471	0
AWL	3,288,707	11.17	354	**354**
not in the lists	17,441,313	59.24	304	**304**
Total	29,441,460	100	2,129	**658**

As can be seen in Table 3, 2000 word families of GSL made up 8,711,440 tokens (29.58%) of the whole corpus. Additionally, AWL (Coxhead, 2000) accounted for 3,288,707(11.17%) of tokens in the corpus. Also, there were 17, 441,313 words (59.24) which did not belong to any lists. Within GSL 2000 words families, 1,471 words met our frequency criterion (higher than 839 times). Because 2000 GSL words are considered as background knowledge for English learners and thus should be eliminated before the consideration of the establishment of accounting AWL, we removed all 1,471 words in our word list. That is, there aren't any of GSL words in our final word list. Ultimately, there were 658-word families that constituted the accounting academic word list (AAWL) (appendix 1). These 658 words occurred 2,989,923 times which account for 10.16 % of the whole corpus.

The results showed that the AAWL made up a large proportion of the running text. The high coverage indicated subject-specificity of the AAWL which reinforce the practical significance of the list in order to read and write in the field of accounting particularly research articles in the top accounting academic journals. Therefore, as explained by Xue and Nation (1984), every field needs to have a specific word list which can meet the needs of its native and non- native speakers who have to read and write academic research articles.

Regarding the coverage words in the list with those of the AWL, out of 658 word families in AAWL, 304 did not belong to AWL. They are italicized in the word list. Out of 570 AWL word families, 354 met our frequency criterion which account for 54.46 % of the AAWL. That is, only 354 of 658 high-frequency word families we identified coincided with those listed in AWL. These words are shown in bold in the list. The degree of occurrence of AWL words in AAWL indicates the significant role of the AWL in accounting research articles. In addition, the existence of 304 non AWL words with high frequency in the corpus supports the idea of specificity in ESP in which every discipline has its own special vocabulary. Thus, every discipline needs to have a specific word list which derives from target texts and accounting is no exception. In other words, the difference between the AAWL and the AWL supports the idea that different disciplines like accounting require a more restricted discipline-based lexical list. This represents challenges to the usefulness of general academic word lists across different disciplines because academic words are not used similarly across disciplines.

Table 4

The Degree of Occurrence of 50 Most Frequent Words in Six Different Word Lists

Lists top 50 words	Coxhead ,2000	Wang, Liang, and Ge ,2008	Ward ,2009	Martinez, Beck, and Panza,2009	Valipouri and Nassaji ,2013	Khani and Tazik,2013	The number of occurrences

Research	*	*		*	*	*	5
Audit		*					1
Financial	*						1
Period	*	*	*	*	*	*	6
Panel	*	*					2
Disclosure							0
Evidence	*			*		*	3
Assets							0
Forecast							0
Analysis	*		*	*			3
Cash							0
Prior	*	*		*	*	*	5
Equity							0
Economics	*			*			2
Negative	*	*	*	*		*	5
Positive	*	*	*	*	*	*	6
Investor	*						1
Significant	*				*	*	3
Consistent	*			*		*	3
Income	*						1
Corporate	*	*					2
Accruals							0
Section	*	*	*		*	*	5
Regression		*					1
Compensation	*						1
Incentives	*						1
Participant	*	*			*	*	4
Percent	*	*		*	*	*	5
Coefficient					*		1
Similar	*	*	*	*	*	*	6
Estimate	*	*	*	*	*	*	6
Role	*	*		*	*	*	5
Empirical	*				*	*	3
Statistics	*	*		*	*	*	5
Median							0
Loss			*		*		2

Annual	*	*		*			3
Ratio	*	*	*	*	*	*	6
Hypothesis	*	*		*	*	*	5
Impact	*	*		*	*	*	5
Internal	*	*			*	*	4
Indicate	*	*	*	*	*	*	6
Announcement							0
Errors	*	*			*	*	4
Contemporary	*						1
Institutional	*					*	2
Specific	*				*	*	3
Theory	*	*	*		*	*	5
Abnormal							0
Potential	*	*		*	*	*	5

In order to find answers to the third question, fifty most frequent word families in our list which are shown in Table 4 were selected. It is very necessary to note that these 50 family words occurred 1,173,488 times in our corpus which accounted for 3.98 % of the whole corpus. It was intended to find out whether these words have occurred in the above popular word lists. If a word has appeared in different word lists, it can be called a multi-purpose academic word but if a word just occurs in one word list it can be called a field-specific word special to that field. As shown in Table 4, we compared 50 most frequent words with words in six different word lists in different fields to find out the degree of coincidence among these seven word lists.

Out of 50 words in our list, 37 appeared in Coxhead's (2000) AWL, 25 in Wang, Liang and Ge's (2008) medical academic word list, 11 in Ward's (2009) engineering English word list, 20 in Martinez, Beck, and Panza's (2009) academic vocabulary in agriculture, 25 Valipouri and Nassaji's (2013) chemistry academic word list, and 26 in Khani and Tazik's (2013) academic word list for applied linguistics. Therefore, our word list shares the most words with AWL; while it shares the least words with the engineering English word list. Of 50 words, there are 6 words (*Period, positive, similar, estimate, ratio and indicate*) that occurred in all 7 word lists and 11 words appeared in 5 word lists. There are nine words (bold words in Table 4) just occur in our word list. It can be said that there are some words in all or at least in most available word lists which can be called all-purpose academic words, because they are frequent in various disciplines. That is, comparing all words in various word lists can indicate the importance of field-specific word lists which once more necessitates the establishment of a word list specific for the field of accounting. Although in this study we could compare fifty words of the AAWL, establishing a word list from comparing different word lists

can be useful and practical for all disciplines. This new academic word list will be more comprehensive than the AWL including words from all academic disciplines. Coxhead (2000) pointed out that words in AWL are frequent in many disciplines, but some words are not in AWL appearing in other word lists, for example, the word *audit* is a frequent word in both in our list and in Wang, Liang and Ge's (2008) medical academic word list.

This reinforces the argument that although vocabulary lists differed from one another in which the majority of words are unique to a single list, there are fairly overlap between various pairs of vocabulary lists. Six out of 50 words occurred in all word lists. Therefore, a composite word list can be developed for future research by carefully scrutinizing all words in the word lists.

6. Pedagogical Implications

The findings of this study can be used for accounting students, instructors and material developers. The AAWL is subject-specific repertoire comprising words that are used frequently in accounting texts, thus, it can be used as a reference for the accounting community. Additionally, the word list can be useful for those who like to read and write accounting RAs. Researchers in the field of accounting who are interested in publishing their papers in topmost influential accounting journals can include thses academic words in their studies to increase the possibility of accepting and publishing their articles. Special attention should be paid to the words in designing materials and developing a curriculum for accounting ESP courses based on the frequency order of the words in the list. Also, the efficiency of language-focused learning is of paramount importance (Nation & Hunston, 2018). Accordingly, accounting teachers and professors can have a repertoire of vocabularies at hand to use them in their classes and to recommend their students learn a reliable word list. To be exact, the AAWL can be taught directly by accounting teachers and professors. Finally, once accounting students begin their university studies, they encounter accounting RAs in order to do research or to be aware of the latest development in the field, therefore there is an urgent need to know the most frequently used words in their field of study.

7. Conclusion

The present study aimed to establish an academic word list specific to the field of accounting. To this end, an accounting academic corpus was used. Although the importance and need for AWL in the field of accounting was considered, a field-specific word list would better cover the accounting academic corpus. Analyzing a large corpus collected from research articles in five most influential accounting journals led to Accounting Academic Word List (AAWL). Therefore, subject-specificity of a word list is of paramount importance. Additionally, the need for the development of a composite word list was emphasized by analyzing the amount of overlap between various word lists. Thus, composite word lists may serve as a source of "core" vocabulary for use in all the above disciplines. Finally,

focusing on academic vocabulary supports the students in their academic studies at university, especially their academic writing and errors in vocabulary use affects the quality of their writing (Coxhead, 2012; Laufer, 1994; Leki & Carson, 1994). Further research is needed to be done in different aspects such as accounting collocations by using this corpus to complement finding of this study.

References

Bauer, L., & Nation, P. (1993). Word families. *International journal of Lexicography*, *6*(4), 253-279.

Bonner, S. E., Hesford, J. W., Van der Stede, W. A., & Young, S. M. (2006). The most influential journals in academic accounting. *Accounting, Organizations and Society*, *31*(7), 663-685.

Chung, T. M., & Nation, 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.

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

Coxhead, A. (2012). Academic vocabulary, writing and English for academic purposes: Perspectives from second language learners. *RELC Journal*, *43*(1), 137-145.

Dang, T. N. Y. (2018). The nature of vocabulary in academic speech of hard and soft- sciences. *English for Specific Purposes*, *51*, 69-83.

Davies, M. (2012). The corpus of contemporary American English: 425 million words, 1990- 2012. Available from https://www.english-corpora.org/coca/

Deveci, T. (2019). Frequently Occurring Words in Education Research Articles Written in English: A Preliminary List. *The Asian ESP Journal*, *15*(1), 8-38.

Dudley-Evans, T. (1993). Subject specificity in ESP: How much does the teacher need to know of the subject?. *ASp. la revue du GERAS*, (1), 1-9.

Dudley-Evans, T. (1994). Genre analysis: An approach to text analysis in ESP. In M. Coulthard (Ed.), *Advances in written text analysis* (pp. 219-228). London: Routledge.

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics*, *35*, 305-327.

Ha, A. Y. H., & Hyland, K. (2017). What is technicality? A Technicality Analysis Model for EAP vocabulary. *Journal of English for Academic Purposes*, *28*, 35-49.

Heatley, A., Nation, I. S. P., & Coxhead, A. (2002). RANGE and FREQUENCY programs. retrieved from https://www.wgtn.ac.nz/lals/resources/vocabulary-analysis-programs

Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.

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

Jablonkai, R. R. (2020). Leveraging professional wordlists for productive vocabulary knowledge. *ESP Today*, *8*(1), 165-181.

Jahangard, A. (2007). Which word types (technical or general) are more difficult to retain by the Iranian high school learners. *The Asian ESP Journal*, *3*(2), 6-23.

Khani, R., & Tazik, K. (2013). Towards the development of an academic word list for applied linguistics research articles. *RELC journal*, *44*(2), 209-232.

Kwary, D. A. (2011). A hybrid method for determining technical vocabulary. *System*, *39* (2), 175-185.

Leki, I., & Carson, J. G. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL quarterly*, *28*(1), 81-101.

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

Martinez, I. A., Beck, S., & Panza, C. B. (2009). Academic vocabulary in agricultural research articles: a corpus-based study. *English for Specific Purposes*, *28*, 183–198.

Nagy, W., Anderson, R., Schommer, M., Scott, J. A., & Stallman, A. (1989). Morphological families in the internal lexicon. *Reading Research Quarterly, 24*, 262–281.

Nation, I. S. P. (2005). Range Program with GSL/AWL List. Retrieved from

www.wgtn.ac.nz/lals/resources/vocabulary-analysis-programs.

Nation, P., & Hunston, S. (2018). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Valipouri, L., & Nassaji, H. (2013). A corpus-based study of academic vocabulary in chemistry research articles. *Journal of English for Academic Purposes*, *12*(4), 248-263.

Wang, J., Liang, S., & Ge, G. (2008). Establishment of a medical academic word list. *English for Specific Purposes, 7*,442–458.

Ward, J. (2009). A basic engineering English word list for less proficient foundation engineering undergraduates. *English for Specific Purposes, 28*,170–182.

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

Xue, G., & Nation, 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.

Yorkston, K., Dowden, P., Honsinger, M., Marriner, N., & Smith, K. (1988). A comparison of standard and user vocabulary lists. *Augmentative and Alternative Communication*, *4*(4), 189-210.

Appendix: Accounting Academic Word List (AAWL) ranked by frequency of occurrence

1. Research
2. Audit
3. Financial
4. Period
5. Panel
6. Disclosure
7. Evidence
8. Assets
9. Forecast
10. Analysis
11. Cash
12. Prior
13. Equity
14. Economics
15. Negative
16. Positive
17. Invest
18. Significant
19. Consistent
20. Income
21. Corporate
22. Accruals
23. Section
24. Regression
25. Compensation
26. Incentives
27. Participant
28. Percent
29. Coefficient
30. Similar
31. Estimate
32. Role
33. Empirical
34. Statistics
35. Median
36. Loss
37. Annual
38. Ratio
39. Hypothesis
40. Impact
41. Internal
42. Indicate
43. *Announcement*
44. Errors
45. Contemporary
46. Institutional
47. Specific
48. **Theory**
49. *Abnormal*
50. Potential
51. Items
52. *Client*
53. Adjusted
54. *Conservatism*
55. Credit
56. *Volatility*
57. Issue
58. Factors
59. Fees
60. *Ability*
61. Alternative
62. Define
63. *Statements*
64. Approach
65. *Column*
66. Individual
67. *Compustat*
68. Focus
69. *Extent*
70. Bias
71. *Proxy*
72. Target
73. Overall
74. *Fiscal*
75. Index
76. *Discretionary*
77. *Versus*
78. *Fraud*
79. *Sox*
80. Affect
81. Strategy
82. *Portfolio*

83. *Score*
84. *Securities*
85. **Voluntary**
86. **Design**
87. **Interaction**
88. **Options**
89. **Selection**
90. **Finally**
91. **Accuracy**
92. **Contrast**
93. **Structure**
94. *Leverage*
95. **Contract**
96. *Knowledge*
97. *Services*
98. *Executive*
99. *Shareholders*
100. **Environment**
101. **Monitoring**
102. *Correlation*
103. *Litigation*
104. **Distribution**
105. **External**
106. **Benefits**
107. **Bond**
108. *Liquidity*
109. **Deviation**
110. *Asymmetry*
111. **Identify**
112. **Previous**
113. **Initial**
114. *Means*
115. *Dummy*
116. *Implications*
117. *Turnover*
118. *Covenants*
119. **Required**
120. **Predicted**
121. *Mandatory*
122. **Professional**
123. *Variation*
124. **Function**
125. **Aggregate**
126. **Range**
127. *Equilibrium*
128. **Enforcement**
129. *Optimal*
130. *Outside*
131. **Subsequent**
132. **Revenue**
133. **Legal**
134. *Magnitude*
135. *Robust*
136. **Primary**
137. **Context**
138. **Volume**
139. **Partner**
140. **Hence**
141. **Survey**
142. **Consequences**
143. **Project**
144. *Reaction*
145. **Series**
146. **Relevant**
147. **Components**
148. *Capture*
149. *Tenure*
150. **Acquisition**
151. *Dividend*
152. **Conference**
153. **Obtain**
154. **Method**
155. **Document**
156. **Investigate**
157. **Task**
158. **Equation**
159. *Consensus*
160. **Underlying**
161. **Persistence**
162. *Intercept*
163. *Transparency*
164. *Expertise*
165. **Outcomes**
166. *Untabulated*
167. **Proportion**
168. **Framework**
169. **Source**
170. *Default*
171. **Benchmark**
172. **Fund**
173. *Database*
174. **Summary**
175. **Found**
176. **Perspective**
177. *Accountability*
178. *Bonus*

179.Constant

180.*Lagged*

181.Status

182.Manipulation

183.Global

184.Crisis

185.*Peer*

186.*Liability*

187.Assume

188.*Profitability*

189.*Customer*

190.Inferences

191.*Transaction*

192.Major

193.*Propensity*

194.Release

195.*Minus*

196.*Francis*

197.Principal

198.Evaluation

199.*Residual*

200.Assess

201.Transfer

202.*Pension*

203.Core

204.Perceive

205.*Variance*

206.Complexity

207.*Compliance*

208.*Incremental*

209.Construct

210.Resources

211.Access

212.*Hedge*

213.Conduct

214.*Dispersion*

215.*Timeliness*

216.Plus

217.*Premium*

218.*Classification*

219.*Goodwill*

220.Media

221.*Typically*

222.Precision

223.Capacity

224.Regime

225.*Digit*

226.Code

227.Unique

228.*Realized*

229.Marginal

230.*Inventory*

231.Exclude

232.Filing

233.Team

234.*Logarithm*

235.*Pressure*

236.Create

237.*Critical*

238.Commission

239.Authors

240.*Senior*

241.*Clustered*

242.Corresponding

243.Mechanisms

244.Exhibit

245.Link

246.Rely

247.Technology

248.Decline

249.*Procedures*

250.Ensure

251.*Decile*

252.*Reliability*

253.Objective

254.*Provisions*

255.Fundamental

256.Implementation

257.Involved

258.*Merger*

259.Abstract

260.Domestic

261.*Insurance*

262.Revision

263.*Reserved*

264.Comments

265.Appropriate

266.Interpretation

267.*Engage*

268.Notion

269.Exposure

270.*Bankruptcy*

271.*Cumulative*

272.Via

273.Categories

274.Impairment

275.Instance

276.*Discount*

277.Job

278.*Outstanding*

279.Computed

280.*Quintile*

281.*Adverse*

282.*Budget*

283.*Capitalization*

284.Attributes

285.Despite

286.Labor

287.*Feedback*

288.*Expenditures*

289.Conclusion

290.*Substantial*

291.Occur

292.Oversight

293.*Mitigate*

294.*Metrics*

295.Constraints

296.Trend

297.Goal

298.Psychology

299.Comprehensive

300.*Tone*

301.Concentration

302.Allocation

303.Intensity

304.Approximately

305.Communication

306.Features

307.Area

308.Norms

309.*Superior*

310.Assurance

311.*Technical*

312.*Depreciation*

313.*Linear*

314.Maturity

315.*Aggressive*

316.Contribute

317.Denote

318.Random

319.Thereby

320.Sustainability

321.Insights

322.Credibility

323.*Pseudo*

324.*Commitment*

325.*Federal*

326.Mutual

327.*Maximum*

328.*Negotiation*

329.*Perceptions*

330.Generate

331.Identity

332.Shift

333.Minimum

334.Normal

335.Innovation

336.Traditional

337.Principles

338.Network

339.*Segment*

340.*Idiosyncratic*

341.Aspects

342.Reveal

343.*Reform*

344.*Sentiment*

345.Text

346.Portion

347.*Agencies*

348.Achieve

349.Established

350.Authority

351.*Optimistic*

352.Violation

353.*Quantitative*

354.Explicit

355.*Interviews*

356.*Univariate*

357.Validity

358.Elements

359.Restricted

360.*Distress*

361.Cycle

362.*Intangible*

363.Grant

364.Criteria

365.Emphasis

366.*Calendar*

367.Derivatives

368.*Utility*

369.Inspection

370.Academic

371.Published

372.Demonstrate

373.*Drift*

374.*Billion*

375.Concept

376.*Endogeneity*

377.Dimensions

378.Gender

379.Motivated

380.*Incorporate*

381.Community

382.Exogenous

383.Emerging

384.Purchase

385.*Partial*

386.Output

387.*Deferred*

388.Preceding

389.*Deflated*

390.*Background*

391.Maintain

392.*Distinct*

393.*Percentile*

394.*Downward*

395.Prohibit

396.Debate

397.*Brokerage*

398.*Modified*

399.*Throughout*

400.Assigned

401.Equipment

402.*Formation*

403.Implicit

404.Expert

405.*Extensive*

406.*Composition*

407.Productivity

408.*Quartile*

409.*Anonymous*

410.Scope

411.*Logistic*

412.Inputs

413.Rational

414.*Anomaly*

415.Conflicts

416.*Fraction*

417.*Entry*

418.Consumer

419.Equivalent

420.*Gross*

421.Version

422.*Sue*

423.Facilitate

424.*Cognitive*

425.Consultants

426.*Accelerated*

427.Parameter

428.Techniques

429.Seek

430.*Subjective*

431.Identical

432.Detect

433.Location

434.*Reconciliation*

435.Sufficient

436.*Conjecture*

437.Reverse

438.Culture

439.Neutral

440.*Momentum*

441.*Payoff*

442.Dynamic

443.*Stakeholders*

444.Enhance

445.*Matrix*

446.*Career*

447.Medium

448.*Sophisticated*

449.*Surplus*

450.*Recall*

451.Ethical

452.*Legitimacy*

453.*Dual*

454.Incidence

455.*Reliance*

456.*Simultaneously*

457.Conformity

458.*Enterprise*

459.*Variability*

460.*Interim*

461.Administrative

462.*Inverse*

463.*Binary*

464.Entity

465.*Proprietary*

466.Acknowledge

467.*Rotation*

468.Eliminate

469.Circumstances

470.*Winsorized*

471.Proceeds

472.*Lawsuits*

473.*Archival*

474.*Goods*

475.Nevertheless

476.*Retail*

477.Specified

478.Welfare

479.Phase

480.*Inflation*

481.Mental

482.Scrutiny

483.*Extant*

484.Logic

485.Lease

486.Lobbying

487.Conventional

488.Region

489.*Footnote*

490.*Tier*

491.Scheme

492.Site

493.Payout

494.Flexibility

495.Geographic

496.Somewhat

497.Respective

498.Retain

499.*Allowance*

500.Intervention

501.Namely

502.Chairman

503.Null

504.Inherent

505.Exceed

506.*Fraudulent*

507.Coordination

508.Willing

509.Agement

510.Takeover

511.Induce

512.Availability

513.Partition

514.Scenario

515.*Hazard*

516.Highlight

517.Subsidiary

518.Characteristic

519.Baseline

520.Vesting

521.Opacity

522.Penalties

523.Aware

524.Grade

525.Diversity

526.Subordinates

527.Ultimately

528.Dominant

529.Illustrate

530.Monetary

531.*Discourse*

532.Imposed

533.*Probe*

534.Supplement

535.*Switch*

536.Prospects

537.Scandals

538.Challenge

539.Contingent

540.*Officers*

541.Confirm

542.Estate

543.*Materiality*

544.*Macroeconomic*

545.Commit

546.Interval

547.Anticipate

548.*Competence*

549.Stable

550.*Thresholds*

551.Format

552.Tendency

553.*Severance*

554.Practitioners

555.Exploit
556.Dense
557.*Moody*
558.Alter
559.Academy
560.Style
561.Schedule
562.Integrated
563.Deficiencies
564.*Subset*
565.Forthcoming
566.Outliers
567.Affiliated
568.Duration
569.Internet
570.Software
571.Essentially
572.Outsourcing
573.Workshop
574.Backdating
575.Vice
576.*Setters*
577.Foundation
578.Preliminary

579.Intuition
580.Congress
581.Transition
582.*Dye*
583.Remove
584.Concurrent
585.Apparent
586.Premiums
587.Brevity
588.*Slack*
589.Ambiguity
590.Recurring
591.Guidelines
592.Stewardship
593.Statutory
594.*Proximity*
595.*Corruption*
596.*Narrative*
597.Intrinsic
598.*Collateral*
599.Energy
600.Oriented
601.*Embedded*
602.*Aversion*

603.*Mortgage*
604.Temporary
605.*Talent*
606.*Sociology*
607.*Numerous*
608.*Regress*
609.Contrary
610.Phenomenon
611.*Confounding*
612.Promotion
613.*Personnel*
614.*Skewness*
615.*Covariance*
616.Extract
617.Chartered
618.*Exit*
619.*Cell*
620.*Cutoff*
621.Topic
622.Ongoing
623.*Electronic*
624.*Amortization*
625.Stress
626.*Justice*

627.*Incumbent*
628.*Retrieved*
629.*Insolvency*
630.*Frank*
631.*Bureau*
632.*Headquarters*
633.*Hereafter*
634.Legislation
635.*Minority*
636.*Corollary*
637.*Convey*
638.Substitute
639.*Plausible*
640.Divisions
641.Skepticism
642.Vector
643.Hybrid
644.*Divergen*ce
645.Symbolic
646.Assistance
647.*Dissemination*
648.*Certified*
649.Normative
650.*Alignment*

651.*Hail*

652.Visible

653.*Consecutive*

654.*Era*

655.Obvious

656.*Pessimistic*

657.*Comply*

658.*Handbook*