A corpus-based lexical study on frequency and distribution of Academic Word List in English for Ethiopia Textbooks Grades 9-12

By: Bereket Gebreselassie

Abstract

This research is a corpus-based evaluation of Academic words in English for Ethiopia textbooks grade 9-12. The methodology involved compilation and analysis of textbook corpus using computer software. For its frame work of analysis Academic Word list and vocabulary profiling software were used. The study Examined the four textbooks by evaluating lexical coverage, vocabulary frequency, vocabulary range and text coverage of academic words. The result from the study indicated that the textbooks contain 498 academic words of the total 570. These 498 academic words have 3.87 text coverage. In addition, 84% of the words are shared among the textbooks.in addition 66% of academic words have adequate repetitions. overall, the finding showed Academic words are well presented in the textbooks. Finally, it was recommended that further studies are needed to give full picture of lexical nature of the textbooks. In addition, teachers and students should be informed of vocabulary research findings and tools to inform and facilitate language learning.

Introduction

In language studies, one of the major focus is determining how appropriate language input is to learners. The issue constitutes various areas including language input, modes of delivery, materials used, and context of teaching learning. One of the sources of input are school textbooks. Textbooks play a central role in school contexts because of their use as a major source of language input. instruction. The role textbooks play is twofold especially in EFL context like Ethiopia, where exposure to the target language is limited to teacher talk and textbook input**.** To play the intended role textbooks should be suitable to their audience.

Studies conducted to evaluate factors contributing to text difficulty have shown that vocabulary knowledge constitutes the main factor related to the difficulty of a text. Vocabulary knowledge is considered major determinant factors for reading and comprehension. Researchers establish this claim substantiated by Research evidence (Adams and Huggins 1985) noted word recognition skill is the major factor that separates good and less skilled readers. The vocabulary of first graders is a significant predictor of reading comprehension ten years later(Cunningham and Stanovich 1997)**.**

Researchers in corpus linguistics have been conducting that are targeted to find vocabulary threshold levels and determine the amount lexical knowledge that a language learners need for various language needs, and the relationship of vocabulary to comprehension (Laufer 1989; Nation and Waring 1997; Webb and Nation 2008; Laufer and Ravenhorst-Kalovski 2010; van Zeeland and Schmitt 2013). One central element of these studies is the use of frequency lists for research as stated before small number of vocabulary account for majority of written and spoken text. It is believed that learners acquire vocabulary in the order of their frequency and range. As stated in the literature the minimum threshold is knowledge of 95% of the vocabulary in a text for comprehension to take place and one will be able to reach this level with the knowledge of the most frequent 3000 words plus proper nouns (Hirsh & Nation, 1992; Laufer, 1989, 1992).

Academic Word List is among the important vocabulary students need due to the high text coverage. Academic words have text coverage of 4%-10% in an academic text. The Academic Word List, compiled by (Coxhead, 2000) consists of 570 word families that are not in the most frequent 2,000 words also known as the general service list of English but which appear frequently

in different academic texts. AWL is very significant for students who know the high frequency words and intend to continue their study in English medium setting (Nation & Coxhead, 2014).

Evaluating textbooks using contemporary available tools will allow to see if students have the exposure to important vocabulary items and helps to make suggestions for adjustments if the textbooks lack important lexical items. The motivation for this study is to contribute to filling the existing gap in textbook evaluation.

Classifications based on vocabulary frequency

One of the classifications of vocabulary is the use of frequency. Currently frequency data is widely used in language studies investigating various aspects of language elements and their occurrences, which is based on the power law. The power law can be found across different fields of study, in language also the frequency with which words are used appears to follow a power law widely known as Zipf's first law. According to (Newman 2005)"When the probability of measuring a particular value of some quantity varies inversely as a power of that value, the quantity is said to follow a power law, also known variously as Zipf's law or the Pareto distribution." This theory states Zipf (1949) as cited in (Booth 1967):

> The number of occurrences of each different word in a text is counted and the words are then arranged in a table in which the first word is the most frequent, the second word the second most frequent, and so on. The order of any word in the list is called its rank (r) and the number of occurrences of that word its frequency (f). Zipf's first law then states that: Rf=c where c is a constant for any particular text p. 386.

According to Zipf's law if we count up how often each word (type) of a language occurs in a large corpus and then list the words in order of their frequency of occurrence, we can explore the relationship between the frequency of a word f and RANK its position in the list, known as its rank (Manning 2008). LI (1992) explains, by choosing the word rank rather than the word length, the exponential distribution that is typical for random texts becomes a power law function. This strongly suggests that the power law as expressed by Zipf's law in natural languages is also purely due to the choice of the rank as the independent variable(Li 1992).

In computational Linguistics Vocabulary, frequency refers to **"**the number of occurrences of a linguistic item in a text or corpus. Different linguistic items have different frequencies of occurrence in speech and writing" (Richards & Schmidt, 2010). As stated above frequency has been used as one way of vocabulary classification. Paul Nation a long-term frequency-based researcher breaks vocabulary into four categories: high-frequency words, academic words, technical words and low-frequency words. That is, it is assumed that both native- and non-native-speaking learners acquire vocabulary largely in the order of its range and frequency. High-frequency and wide-range words are generally learned before lower-frequency and narrower-range words (Nation, 2006).

As one of the requirements for cognitive processes in knowledge acquisition is the availability of multiple exposure and chances for practicing the language thus learning frequently used vocabulary items enables language learning and facilitates the stage for the next stages of learning. Now days the availability of accessible tools and advancement of computer technology specifically in analysis of corpora has led to the increasing interest in corpus linguistics research. Schmitt (2010) in explaining the importance of corpus-based frequency counts states that:

> Corpus counts are objective and quantifiable, and computers are well suited to fast and accurate counting. …. corpus evidence can uncover lexical behaviour that would be difficult to intuition. However, it is useful to not become complacent and too trusting of automatized computer counts. A frequency count is only as good as the corpus it is based upon, and every corpus has limitations. (p.86)

There is easy access currently of frequency lists of language use, and consequently reliable information on which words are preferred by the speakers in different domains and communicative situations (Sánchez and Pérez 2009). As stated in the above explanations, using corpus driven data for research and teaching has various merits because the frequency results show which language element is used frequently and what there order of importance is. Even if one does not base all decisions on frequency data, it will be useful to use it in decision making related to vocabulary teaching and material preparation especially in lower stages of language teaching.

High frequency words

High frequency words as the name indicates is the list of frequently occurring vocabulary in language use**.** According to Nation (Nation, 2011) "high-frequency words make up a relatively small, very useful group of words that are important no matter what use is made of the language. Because each word in this group is frequent, the learners will get a very good return for learning them." Most of these words are content words and knowing enough of them allows a good degree of comprehension of a text.

The most widely cited high frequency words are the first 2000 words families. This classification was based on the classic list of high frequency words is Michael West's General Service List 1953(Nation & Waring, 1997).According to(Schmitt and Schmitt 2014) GSL includes a little over 2,000 headwords and has been an important resource for teachers and material writers for many decades. The 2,000 figures were based on research on oral discourse. Later in a recent study (Nation, 2006) argued that the core vocabulary should include the most frequent 3,000 word families. After this level of 3,000-word families, other opportunities to learn will avail because learner who attained this level can learn from context (Schmitts, 2014). The idea here is that high frequency words are very useful due to their frequent occurrences in language use and special focus should be given to teaching these words before passing to other vocabulary groups.

Mid- Frequency Words

The mid-frequency vocabulary covers the range between the high and low frequency vocabulary. To some extent, the academic and technical vocabulary can be classified as the mid-frequency range. (Schmitts, 2014, p. 493). The mid-frequency vocabulary is essential when operating in English across a range of different topics and situations, for instance to cope with the university studies in English. (Schmitts, 2014, p. 495).

Mid-frequency words are also used for defining other words technical words are often explained in the texts. Schmitt continues in asserting the need to distinguish mid frequency vocabularies stating that high-frequency vocabulary in English extends up to about 3,000 word families, and that low-frequency vocabulary begins at about the 9,000 frequency level. This leaves a great gap between the 3,000 and 9,000 levels which has not been systematically addressed before and suggest to call this group of vocabulary as mid frequency band." the naming is important as "it

allows the field to recognize it as a discrete phenomenon, with its own unique benefits for users, and pedagogical challenges for language practitioners (Schmitt and Schmitt 2014:495).

Low frequency vocabulary

Nation, (2011) states that the term Low frequency words refers to words that occur infrequently and are sometimes limited to specific fields of study. When the complete word frequency count is made for a text, it is found that words of high rank, that is of low frequency, occur in such a way that many words have the same frequency (Booth 1967).Nation, (2011) suggests that significant amount of class time should not be taken to teach this words. Instead, learners should learn them according to their specific need of these words. In addition, nation states once high frequency words are learned teachers should focus on equipping students with the knowledge of strategies to learn low frequency words instead of specifically focusing on them. In any text of a thousand words or more about 40% of the word types, occur only once. In a written text, about one word in every 18 or 20 is a low frequency word. While low frequency words are often met only in reading. High frequency words occur in the receptive activities of reading and listening, and also in the productive activities of writing and speaking(Nation, 1983).

Studies show that the presence of low frequency vocabulary affects reading comprehension. Marks et .al (1974) claimed if 15 per cent of words in a reading text is replaced with low frequency words comprehension decreases substantially. Nation (1983) states large number of unknown words in a reading material has negative impact on readers' comprehension and vocabulary learning.

Schmitt and Schmitt (2014) argues that it is not easy to demarcate the point where low frequency vocabulary starts 'The conceptualization as a high/low frequency dichotomy is untenable, as the vocabulary immediately beyond the 3,000 high-frequency cut-off point (i.e. at the 4,000 and 5,000 levels) is clearly too useful to be written off as low-frequency vocabulary'. Schmitt continues in asserting the need to distinguish mid frequency vocabularies stating that high-frequency vocabulary in English extends up to about 3,000-word families, and that low-frequency vocabulary begins at about the 9,000-frequency level. This leaves a great gap between the 3,000 and 9,000 levels, which has not been systematically addressed before. They suggest to call this group of vocabulary as mid frequency band 'the naming is important as it allows the field to recognize it as

a discrete phenomenon, with its own unique benefits for users, and pedagogical challenges for language practitioners' (Schmitt and Schmitt 2014).

Academic language

Before further discussions of academic vocabulary, it is essential to look at what academic language is because academic vocabulary exists in academic language. Knowledge of Academic language is taken as one of the important knowledges for everyone involved in an academic setting. Since the general term, academics might mean many things in language teaching and learning researchers have given definitions and tried to explain what constitutes as an academic language. Snow (2010) as cited in (Nagy and Townsend 2012) noted, 'There is no exact boundary when defining academic language; it falls toward one end of a continuum (defined by formality of tone, complexity of content, and degree of impersonality of stance), with informal, casual, conversational language at the other extreme'. (p. 450)

Thus, the classification of academic language emanates from the language characteristics used in academics. (Nagy and Townsend 2012) suggested a unified definition which emanates from the perspectives of academic language. Their definition is "Academic language is the specialized language, both oral and written; of academic settings that facilitates communication and thinking about disciplinary content." They explain that academic language can be both written and oral and both have similarities and differences. Despite the difference, there exists reciprocated supportive role among written and oral academic language. They describe academic setting "any context in which disciplinary concepts are being conveyed, in the traditions in which they were developed or discovered, uses academic language" and role of facilitating communication and thinking is based on the notion that academic thinking involves the cognitive processing of disciplinary concepts and phenomena, which would be near impossible without academic language. As it can be noted from the definitions due to, the focus in conveying academic communication and the role of academic setting in tuning the language to fit its need academic language differs from everyday conversation. According to (Biber, 2006), academic language differs because it contains more of the following elements:

1. Latin and Greek vocabulary
2. Morphologically complex words

3. Nouns, adjectives, and prepositions

4. Grammatical metaphor, including Nominalization

Academic vocabulary and technical vocabulary

As (Coxhead, 2000a) defined academic words as a type of specialized vocabulary that appears most frequently in written academic text like textbooks and journal articles but don't occur abundantly in other settings. Academic words may be discipline specific or cross-disciplinary in their use (Townsend et al. 2012). (Baumann and Graves 2010) state

Cross-disciplinary or General academic vocabulary is vocabulary that exists in texts across several disciplines or academic domain whereas Domain specific is the type of vocabulary unique to a certain discipline. Hirish notes that Academic vocabulary exists for two major reasons, "First, it signals the formality and seriousness of the subject matter. Second, it consists of vocabulary that allows academics to do the kinds of things that academics typically do" (Hirsh, 2004). One example of what an academic language vocabulary is the list of special purposes vocabulary is the Academic Word List (Coxhead, 2000). This list contains 570-word families believed to have paramount importance for learners. The list high frequency academic words in various academic subjects (Nation and Coxhead 2012).

Another type of specialized vocabulary is referred to as technical vocabulary. Technical words are those that are specifically used in a particular subject area, and as opposed to academic words, technical words are directly related to the topic discussed. Words such as molecule and bacteria are examples of technical words in middle school science textbooks that are specifically used in a particular subject area, and as opposed to academic words, technical words are directly related to the topic discussed(Greene 2015). They range from items which are unique to the field and do not occur elsewhere to items that have the same form as high-frequency items but specialized meanings within a field. Technical items are reasonably common within a field, but not so common elsewhere, and differ from subject area to subject area. (Schmitt 2010). Due to the nature of technical words sometimes named jargons, a reader will face difficulties if they are not familiar with the words. Flood and west (1950) explain that beyond the nature the concepts one of the reasons for difficulty in reading science texts is the existence of unfamiliar words to the reader.

"Vocabulary presents a problem even in scientific articles written for English readers" (Flood and West 1950).Thus in addition to general English high frequency words students who join a specific field will need high frequency technical vocabulary of the subjects.

The Academic Word List AWL

The Academic Word List, compiled by (Coxhead, 2000) consists of 570 word families that are not in the most frequent 2,000 words also known as the general service list of English but which appear frequently in different academic texts. The list is divided into ten sublists according to the frequency and range of the words. For example, Sublist 1 contains the most common words in the Academic Word List. Sublist 2 contains the next most common words, and so on. There are 60 families in each sublist, except for subsist 10 which as 30. The List is not restricted to a specific field of study and it does not include words that are only specific to a given field of study. AWL is very significant for students who know the high frequency words and intend to continue their study in English medium setting (. Nation & Coxhead, 2014). Coxhead highlights the importance of AWL as follows: knowledge of the highest incidence and high utility academic words in English can significantly boost a student's comprehension level of school-based reading material. Secondary students who are taught these high-utility academic words and routinely placed in contexts requiring their usage are likely to be able to master academic material with more confidence and efficiency, wasting less time and energy in guessing words or consulting dictionaries than those who are only equipped with the most basic 2000-3000 words that characterize ordinary conversation.

The corpus used in making AWL has 3,600,000 token. It was prepared by collecting large number of texts were collected from the four academic streams of humanities, science, commerce, and law. Each stream was further divided in to 7 subject areas which is a total of 28 subject areas and the texts included range from a balanced length of short and long texts for each subject and stream(P. Nation & Coxhead, 2014). Nation and Coxhead summarized the criteria for preparing AWL

The word family could not be in 2,000-word families of A General Service List of English Words by Michael West (1953). The word family had to occur in all four faculty divisions of the corpus, and had to occur a minimum of 10 times in each faculty division.

The word family had to occur in at least 15 of the 28 subject divisions of the corpus. The word family had to occur with a frequency of at least 100 occurrences in the corpus. The AWL according to (Coxhead, 2000) covers around 10% of academic corpus.

Text coverage of vocabulary lists

Research by (Liu Na and Nation 1985)has shown that this ratio of unknown to known words is not sufficient to allow reasonably successful guessing of the meaning of the unknown words. At least 95% coverage is needed for that. Research by (Laufer, 1989)suggests that 95% coverage is sufficient to allow reasonable comprehension of a text. Thus, a learner or teachers should systematically approach and utilize strategies to learn more vocabulary that help to move from 80%-95% coverage. For example, if one learns the 570 academic word lists after learning 2000 list, which equals 2570 words, can bring the coverage of an academic text up to approximately 90%. In other words, knowledge of the first 2000 plus 570 AWL words enables knowledge of 90% of the words one comes across in any academic text. As stated in the above section AWL is focused on because, it covers approximately 10% of any academic text. It appears slightly in different frequencies in arts, commerce, law, and science texts. It covers 12.0% in commerce texts , 9.3% in arts, 9.4% in law, and 9.1% in science texts(Coxhead 2000b). As it can be seem from its coverage knowing this list of words will be essential for students "AWL is the next 'band' to teach after high-frequency vocabulary, and everything after that is de facto low-frequency vocabulary, as it is rarely addressed in any principled manner" (Schmitt & Schmitt 2014).

Thus, it is essential to plan vocabulary teaching according to the language need of learners, which in turn helps to decide what kinds of vocabulary needs more emphasis: frequency information guides these decisions.

Research Questions

In order to meet the objective of evaluating Academic words in the four textbooks the following research questions were the focus of the study.

1. What is the presentation of Academic word list (AWL)in the text books?
2. What is the distribution and frequency of academic words?
3. What is the text coverage of academic words in the textbook?

Methodology

The methodology adopted for this research project is corpus-based research methodology, which involves a compilation and analysis of corpora using computer software. The four textbooks books were converted to machine-readable format. OCR devices with appropriate software can scan printed material. This method of conversion from print to electronic form is becoming more cost effective each year as OCR technology improves(Atkins, Clear, and Ostler 1992). Alternatively, text can be keyboarded manually. The output of this step will be an electronic copy of the four textbooks. Each file will be saved with its own distinct file name. Foxit phantom pdf OCR identification was used for data capturing from the original pdf format of the textbooks.

The obtained copy of the books was cleaned of unwanted elements this includes pictures, all the bibliographies, tables, numerals, and appendices. Once the text is treated, each book was compiled in to ASCII (or text) format file format, which is understood by the analysis software. This is the standard format used for texts included in corpora. The main advantage is that ASCII is a universally recognized text format, one that can be used with any word-processing program and the numerous software programs designed to work on corpora, such as taggers and parsers and concordances (Meyer 2002).

For Cleaning, unwanted contents and editing of scanning errors notepad ++ and Foxit phantom pdf and ant word profiler were used in order using phantom pdf non-textual elements like pictures tables graphs were removed, once this was achieved each textbook was converted to text file. following this the text files were entered to notepad ++ to remove un wanted textual elements and edit errors occurred during the conversion process. Using the software numbers, non-English text elements were removed then errors including spelling and misrecognized words were corrected by referring to each textbook and manually typing these words.

Analysis

Analysis was carried out using Range Software. RANGE is used to compare the vocabulary of up to 32 different texts at the same time(Anthony 2014). RANGE can be used to compare a text against vocabulary lists to see what words in the text are and are not in the lists, and to see what percentage of the items in the text are covered by the lists. It can also be used to compare the

vocabulary of two texts to see how much of the same vocabulary they use and where their vocabulary differs. The analysis was concerned on examining the Academic Words presentation of the textbooks thus Academic Word List (AWL) was used as a bases of analysis. Each output from the analysis was entered to XL spread sheet to further examine and obtain descriptive statistics of the range and frequency of academic words.

Academic Vocabulary in the textbooks

As explained in the literature AWL is an important word list that covers up to 10% in any academic text. Thus, English textbooks are expected to include these academic vocabularies. Since the textbooks under this study are expected to prepare students for higher education language need the distribution and traits of AWL is examined. The table below shows the result of AWL quantity and text coverage.

Table 1 Quantity and text coverage of AWL

Word Lists	General AWL		Passages AWL	
	Tokens %	**Word Families**	**Tokens %**	**Word Families**
Grade 9	2236/ 3.33	331	598/ 3.03	**233**
Grade 10	2088/ 3.22	334	505/2.82	192
Grade 11	4270/ 4.38	393	640/4.47	233
Grade 12	3290/ 4.23	367	664/4.39	238
AIO	11884/ 3.87	498	2407/ 3.55	399

There are 331 academic word families in grade nine textbook with a coverage of 3.33% of the lexical items. Grade Ten textbook contains 334 words covering 3.22% which is lower than the previous textbook. In terms of lexical coverage This is the third higher coverage next to the 2000 level words. Students will benefit if they learn AWL after the 2000 level is reached.

Among the textbooks grade 11 and 12 textbooks have high number of AWL both in terms of number and coverage. Grade Eleven consists 393 AW words with a 4.38% coverage and 367 AW appear in grade twelve covering 4.23% of the running words.

the table shows that reading sections contains 399 academic words. The distribution of AW indicates that Grade nine and ten consist equal amount of 233 AW. Grade TEN TB contains 192AW which is the smallest among the TB and grade twelve contains the highest number of with 238 AW.

The coverage data shows that AW cover 3.03% in grade nine, 2.82% in grade ten,4.47% in grade eleven and 4.49% in grade twelve. the data also shows that AW cumulative coverage is 3.55%. Showing that AW significant text coverage in the reading sections. The data also shows that AW coverage is higher in grade eleven and twelve. when compared in terms of coverage high coverage was observed in grade twelve and in grade ten AW coverage is low.

Distribution of Academic Words

As part of giving a complete picture range and distribution of academic words was also examined. The distribution of academic words across the textbook and the reading section is presented in Table

Table 2 Distribution of AWL

AWL by	**General**		**Passages**	
Range	**Families**	**Percent**	**Families**	**Percent**
In four books	206	41.57%	64	31%
In three books	105	21.08%	97	28%
In two books	99	19.88%	113	24%
In one book	88	17.67%	125	16%
Total	498	100%	399	100%

Over all there are 498 academic words in the four textbooks that fall in four shared range categories. The data showed Range wise 206-word families are common to all four text books,105 shared by three,99 by two of the books. This shows that 82.53% are shared academic words. The data also showed that 88 Academic Words or 17.67% are unshared Academic words. Thus, on average each book contains 22.25 exclusive academic words.

The range analysis on the distribution of AWL in the reading sections showed that there are 399 academic words in the reading sections of the textbooks. the result showed that 274 academic words which are shared academic words. shared academic words constitute 84% of the words.

Frequency of Academic words

The analysis carried out to examined to measure if the words recycling or occurrences meets the minimum repetition frequency. As stated earlier the frequency amount considered is five repetitions. Thus, the words considered high frequency words is based on their reoccurrence in a textbook the following table shows the frequency result of academic words,

Table 3 Frequency data of Academic words

Frequency	General		Passages	
	Families	**Percent**	**Families**	**Percent**
High frequency AWL	328	66%	162	*41%*
Low Frequency AWL	170	34%	237	*59%*
TOTAL	498	100%	399	100%

As shown in the table 328-word families have a frequency high frequency words and 170 word. The frequency data of AW in the reading sections of the Textbooks revealed that 162 AW are high

frequency words and 237 AW are low frequency words. The amount of high frequency vocabulary is ,66% in general and 59% in the reading sections.

Academic words in the text books

One of the issues investigated is the presentation of academic words in the textbook. This included the amount, distribution and frequency of academic words. As (Coxhead, 2000) stated academic words are type of specialized vocabulary that appears most frequently in written academic text like textbooks and journal articles but do not occur abundantly in other settings. Academic words contain 570-word families believed to have paramount importance for learners. The list high frequency academic words in various academic subjects and cover around 4%-10% of any academic text (Nation and Coxhead 2012).

The findings of this study showed that 87% of the AWL are found in the textbooks. The text coverage of the words ranged from 3.33%-4.38% in general and 2.82%-4.47% in the reading sections. This is the highest text coverage next to the 2000K high frequency words. In addition, the finding showed higher amount both in number and text coverage of AWL in grade eleven and twelve. The higher number in these grades is expected due to the expectations that students will sit for university entrance exam at the end of grade twelve and join universities if successful.

Academic literature shows if students learned academic words after mastery of the first 2000 words there will be a high return in text coverage. As (Coxhead, 2000) stated ,Secondary students who are taught these high-utility academic words and routinely placed in contexts requiring their usage are likely to be able to master academic material with more confidence and efficiency, wasting less time and energy in guessing words or consulting dictionaries than those who are only equipped with the most basic 2000-3000 words.

Regarding common academic words 82.53% of AWL are shared in general and 84% in reading sections. This indicates AWL are shared and recycled vocabulary items. The inclusion of similar AWL across textbooks will in turn elevate the likelihood of learning these words. This is also supported by the frequency findings of AWL because most words appear at a frequency of five and more times. The amount of high frequency vocabulary is ,66% in general and 59% in the reading sections. This shows that most AWL are adequately repeated and can be encountered

across the textbooks. Such organized presentation of the words will benefit students since the likelihood of learning will be increased due to repeated exposure. In addition, learning Academic words will yield higher text coverage for students who continue their education in higher institutions because Awls' high text coverage than in purely academic texts than language textbooks. Thus, academic words presentation in the textbooks is sufficient, and well organized and in consideration of lexical principles.

The use of AWL in textbooks is considered as one of the well incorporated vocabulary research findings in textbook writing. Regarding its wide use Nation states Coxhead's (2000) research on academic vocabulary is well known by teachers as well as coursebook writers because Coxhead made the lists freely available in different formats and the ease of the list to use and teachers also will be familiar to these words as part of their academic trainings (Nation 2011). Similar, vocabulary research findings and resources are made available by researchers in the field. These resources can be used by different stake holders in the academic scene including teachers and researchers. Using such resources deemed important sufficiently should be focused on and promoted by policy makers so that students can benefit more from their language classrooms and textbooks.

Conclusion

Textbooks play a central role in the teaching learning transaction among teachers and students. Due to their role teaching materials must have high standards in all aspects that make it in to the textbook. In language classrooms a well-crafted textbook will benefit students' linguistic development and enables them to meet the learning objectives of their education. For this reason, studying and evaluating textbooks is essential.

The present study evaluated the academic vocabulary aspect of Government prescribed high school English for Ethiopia textbooks by employing corpus-based research method. Regarding AWL presentation the finding showed Academic words are well presented 498 of 570 (87%) academic words are in the textbooks in addition the most academic words are shared and presented in high frequency.

The use of AWL in textbooks is considered as one of the well incorporated vocabulary research findings in textbook writing. Regarding its wide use Nation states Coxhead's (2000) research on academic vocabulary is well known by teachers as well as coursebook writers because Coxhead made the lists freely available in different formats and the ease of the list to use and teachers also will be familiar to these words as part of their academic trainings (Nation 2011). Similar, vocabulary research findings and resources are made available by researchers in the field. These resources can be used by different stake holders in the academic scene including teachers and researchers. Using such resources deemed important sufficiently should be focused on and promoted by policy makers so that students can benefit more from their language classrooms and textbooks.

The study results are limited to Academic words evaluation thus other vocabulary aspects of the books should be explored to definitively state if all vocabulary elements are presented in an appropriate manner. However, the findings from this research is a starting place to inform teachers and students how academic words are presented. It will not be feasible for teachers to compensate for every problem in the textbooks from limited class time and a limited textbook however being informed of the facts and availing resources like important word lists, text evaluating techniques, and vocabulary tests, will guide students to focus their time and energy on the important items that will assist them in the long term and enable wide use of their vocabulary knowledge to foster their academic journey.

Recommendations

The present study was an endeavour to evaluate the Academic vocabulary aspect of Government prescribed high school English for Ethiopia textbooks. Based on the findings and scope of the study the following recommendations are made. Future researches should focus on evaluating other word lists and lexical elements also study should be conducted to evaluate more textbooks of language and content area textbooks using similar methodologies. Study types may focus on replication and expansion of evaluation. In addition, studies that relate textbook evaluation with students' vocabulary knowledge will be very essential. Studies also must include testing various aspects of students' vocabulary knowledge at different grade levels so as to assist students' vocabulary learning in particular and language learning in general. Such studies and testing should not be limited to elementary and high school textbooks and students. Rather it will be useful to

conduct studied at higher education level to address the different vocabulary needs across field of studies. Findings from such studies will be useful to tailor and execute study findings to the nature and demand of specific field of study accordingly

References

Adams, Marilyn Jager, and A. W. F. Huggins. 1985. “The Growth of Children’s Sight Vocabulary: A Quick Test with Educational and Theoretical Implications.” *Reading Research Quarterly* 20(3):262.

Anthony, L. 2014. *AntWordProfiler*. Tokyo, Japan: Waseda University.

Atkins, Sue, Jeremy Clear, and Nicholas Ostler. 1992. “Corpus Design Criteria.” *Literary and Linguistic Computing* 7(1):1–16.

Baumann, James F., and Michael F. Graves. 2010. “What Is Academic Vocabulary?” *Journal of Adolescent & Adult Literacy* 54(1):4–12.

Booth, Andrew D. 1967. “A “Law≒ of Occurrences for Words of Low Frequency.” *Information and Control* 10(4):386–393.

Coxhead, Averil. 2000a. “A New Academic Word List.” *TESOL Quarterly* 34(2):213–238.

Coxhead, Averil. 2000b. “A New Academic Word List.” *TESOL Quarterly* 34(2):213–238.

Cunningham, Anne E., and Keith E. Stanovich. 1997. “Early Reading Acquisition and Its Relation to Reading Experience and Ability 10 Years Later.” *Developmental Psychology* 33(6):934.

Flood, W. E., and M. P. West. 1950. “A Limited Vocabulary for Scientific and Technical Ideas (I).” *ELT Journal* 4(4):104–108.

Greene, Jennifer Wells. 2015. “The Fundamentals of Academic Vocabulary: Essential Concepts for Middle School Students and Their Teachers.” *Perspectives on Language and Literacy* 41(3):29.

Laufer, Batia. 1989. “What Percentage of Text-Lexis Is Essential for Comprehension?”

Laufer, Batia, and Geke C. Ravenhorst-Kalovski. 2010. “Lexical Threshold Revisited: Lexical Text Coverage, Learners’ Vocabulary Size and Reading Comprehension.” *Reading in a Foreign Language* 22(1):15–30.

Li, Wentian. 1992. “Random Texts Exhibit Zipf’s-Law-like Word Frequency Distribution.” *IEEE Transactions on Information Theory* 38(6):1842–1845.

Liu Na, and I. S. P. Nation. 1985. “Factors Affecting Guessing Vocabulary in Context.” *RELC Journal* 16(1):33–42.

Manning, Christopher D. 2008. *Foundations of Statistical Natural Language Processing*.

Meyer, Charles F. 2002. *English Corpus Linguistics: An Introduction*. Cambridge, UK; New York: Cambridge University Press.

Nagy, William, and Dianna Townsend. 2012. “Words as Tools: Learning Academic Vocabulary as Language Acquisition” edited by N. Lesaux and N. Schmitt. *Reading Research Quarterly* 47(1):91–108.

Nation, I. 2006. “How Large a Vocabulary Is Needed for Reading and Listening?” *Canadian Modern Language Review* 63(1):59–82.

Nation, I. S. P. 2011. “Research into Practice: Vocabulary.” *Language Teaching* 44(04):529–39.

Nation, Paul. 1983. “Learning Vocabulary.” *New Zealand Language Teacher*.

Nation, Paul, and Averil Coxhead. 2012. “Special Purposes Vocabulary.” in *The Encyclopedia of Applied Linguistics*, edited by C. A. Chapelle. Oxford, UK: Blackwell Publishing Ltd.

Nation, Paul, and Averil Coxhead. 2014. “Vocabulary Size Research at Victoria University of Wellington, New Zealand.” *Language Teaching* 47(03):398–403.

Nation, Paul, and Robert Waring. 1997. “Vocabulary Size, Text Coverage and Word Lists.” Pp. 6–19 in *Vocabulary: Description, Acquisition and Pedagogy*, edited by N. Schmitt and M. M. Cambridge University Press.

Newman, Mej. 2005. “Power Laws, Pareto Distributions and Zipf’s Law.” *Contemporary Physics* 46(5):323–51.

Richards, Jack, and Richard Schmidt. 2010. *Longman Dictionary of LANGUAGE TEACHING AND APPLIED LINGUISTICS*. 4th ed. London: Pearson.

Sánchez, Raquel Criado, and Aquilino Sánchez Pérez. 2009. “Vocabulary in EFL Textbooks: A Contrastive Analysis against Three Corpus-Based Word Ranges.” Pp. 862–875 in *A survey of corpus-based research [Recurso electrónico]*.

Schmitt, Norbert, and Diane Schmitt. 2014. “A Reassessment of Frequency and Vocabulary Size in L2 Vocabulary Teaching.” *Language Teaching* 47(04):484–503.

Townsend, Dianna, Alexis Filippini, Penelope Collins, and Gina Biancarosa. 2012. “Evidence for the Importance of Academic Word Knowledge for the Academic Achievement of Diverse Middle School Students.” *The Elementary School Journal* 112(3):497–518.

Webb, Stuart, and Paul Nation. 2008. “Evaluating the Vocabulary Load of Written Text.” *TESOLANZ Journal* 16:1–10.

van Zeeland, H., and N. Schmitt. 2013. “Lexical Coverage in L1 and L2 Listening Comprehension: The Same or Different from Reading Comprehension?” *Applied Linguistics* 34(4):457–79.

Appendices

Academic Word Families in the Textbooks

ABANDON	APPEND	BRIEF	COMPLEMENT	CONTEXT
ABSTRACT	APPRECIATE	BULK	COMPLEX	CONTRACT
ACADEMY	APPROACH	CAPABLE	COMPOUND	CONTRADICT
ACCESS	APPROPRIATE	CAPACITY	COMPREHENSIVE	CONTRAST
ACCOMMODATE	APPROXIMATE	CATEGORY	COMPUTE	CONTRIBUTE
ACCOMPANY	AREA	CHALLENGE	CONCEIVE	CONTROVERSY
ACCURATE	ASPECT	CHANNEL	CONCENTRATE	CONVENE
ACHIEVE	ASSEMBLE	CHAPTER	CONCLUDE	CONVERSE
ACKNOWLEDGE	ASSESS	CHART	CONDUCT	CONVERT
ACQUIRE	ASSIGN	CHEMICAL	CONFER	CONVINCE
ADAPT	ASSIST	CIRCUMSTANCE	CONFINE	COOPERATE
ADEQUATE	ASSUME	CITE	CONFIRM	COORDINATE
ADJUST	ASSURE	CIVIL	CONFLICT	CORE
ADMINISTRATE	ATTACH	CLARIFY	CONFORM	CORPORATE
ADULT	ATTAIN	CLASSIC	CONSENT	CORRESPOND
ADVOCATE	ATTITUDE	CLAUSE	CONSEQUENT	COUPLE
AFFECT	ATTRIBUTE	COHERENT	CONSIDERABLE	CREATE
AID	AUTHOR	COLLAPSE	CONSIST	CREDIT
ALLOCATE	AUTHORITY	COLLEAGUE	CONSTANT	CRITERIA
ALTER	AUTOMATE	COMMENT	CONSTITUTE	CRUCIAL
ALTERNATIVE	AVAILABLE	COMMISSION	CONSTRUCT	CULTURE
AMEND	AWARE	COMMIT	CONSULT	CURRENCY
ANALYSE	BEHALF	COMMODITY	CONSUME	CYCLE
ANNUAL	BENEFIT	COMMUNICATE	CONTACT	DATA
ANTICIPATE	BIAS	COMMUNITY	CONTEMPORARY	DEBATE
APPARENT	BOND	COMPATIBLE		DECADE

DECLINE	DYNAMIC	EXHIBIT	GENERATION	INJURE
DEDUCE	ECONOMY	EXPAND	GLOBE	INNOVATE
DEFINE	EDIT	EXPERT	GOAL	INPUT
DEFINITE	ELEMENT	EXPLOIT	GRADE	INSERT
DEMONSTRATE	ELIMINATE	EXPORT	GRANT	INSIGHT
DENOTE	EMERGE	EXPOSE	GUARANTEE	INSPECT
DENY	EMPHASIS	EXTERNAL	GUIDELINE	INSTANCE
DEPRESS	ENABLE	EXTRACT	HENCE	INSTITUTE
DERIVE	ENCOUNTER	FACILITATE	HIGHLIGHT	INSTRUCT
DESIGN	ENERGY	FACTOR	IDENTICAL	INTEGRATE
DESPITE	ENFORCE	FAMILIES	IDENTIFY	INTELLIGENCE
DETECT	ENHANCE	FEATURE	IDEOLOGY	INTENSE
DEVICE	ENORMOUS	FEDERAL	IGNORANT	INTERACT
DEVOTE	ENSURE	FEE	ILLUSTRATE	INTERNAL
DIFFERENTIATE	ENTITY	FILE	IMAGE	INTERPRET
DIMINISH	ENVIRONMENT	FINAL	IMMIGRATE	INTERVAL
DISCRETE	EQUATE	FINANCE	IMPACT	INTERVENE
DISCRIMINATE	EQUIP	FINITE	IMPLEMENT	INVEST
DISPLACE	EQUIVALENT	FLEXIBLE	IMPLY	INVESTIGATE
DISPLAY	ERODE	FLUCTUATE	IMPOSE	INVOLVE
DISPOSE	ERROR	FOCUS	INCIDENCE	ISOLATE
DISTINCT	ESTABLISH	FORMAT	INCLINE	ISSUE
DISTORT	ESTATE	FOUNDATION	INCOME	ITEM
DISTRIBUTE	ESTIMATE	FOUNDED	INDEX	JOB
DIVERSE	ETHNIC	FRAMEWORK	INDICATE	JOURNAL
DOCUMENT	EVALUATE	FUNCTION	INDIVIDUAL	JUSTIFY
DOMESTIC	EVENTUAL	FUND	INDUCE	LABEL
DOMINATE	EVIDENT	FUNDAMENTAL	INFER	LABOUR
DRAFT	EVOLVE	FURTHERMORE	INFRASTRUCTURE	LAYER
DRAMA	EXCEED	GENDER	INITIAL	LECTURE
DURATION	EXCLUDE	GENERATE	INITIATE	LEGAL

LIBERAL	NEVERTHELESS	POLICY	RADICAL	REVOLUTION
LICENCE	NONETHELESS	PORTION	RANDOM	ROLE
LIKEWISE	NORMAL	POSE	RANGE	ROUTE
LINK	NOTION	POSITIVE	RATIO	SCENARIO
LOCATE	NUCLEAR	POTENTIAL	REACT	SCHEDULE
LOGIC	OBJECTIVE	PRECEDE	RECOVER	SCHEME
MAINTAIN	OBTAIN	PRECISE	REFINE	SCOPE
MAJOR	OBVIOUS	PREDICT	REGIME	SECTION
MANIPULATE	OCCUPY	PREDOMINANT	REGION	SECTOR
MARGIN	OCCUR	PRELIMINARY	REGISTER	SECURE
MATURE	ODD	PREVIOUS	REGULATE	SEEK
MAXIMISE	OPTION	PRIMARY	REJECT	SELECT
MECHANISM	ORIENT	PRIME	RELAX	SEQUENCE
MEDIA	OUTCOME	PRINCIPAL	RELEASE	SERIES
MEDICAL	OVERALL	PRINCIPLE	RELEVANT	SEX
MEDIUM	OVERLAP	PRIOR	RELUCTANCE	SHIFT
MENTAL	OVERSEAS	PRIORITY	RELY	SIGNIFICANT
METHOD	PANEL	PROCEED	REMOVE	SIMILAR
MIGRATE	PARAGRAPH	PROCESS	REQUIRE	SITE
MILITARY	PARALLEL	PROFESSIONAL	RESEARCH	SOLE
MINIMAL	PARTICIPATE	PROHIBIT	RESIDE	SOMEWHAT
MINIMISE	PARTNER	PROJECT	RESOLVE	SOURCE
MINIMUM	PASSIVE	PROMOTE	RESOURCE	SPECIFIC
MINISTRY	PERCEIVE	PROPORTION	RESPOND	SPECIFY
MINOR	PERCENT	PROSPECT	RESTORE	STABLE
MODIFY	PERIOD	PSYCHOLOGY	RESTRICT	STATISTIC
MOTIVE	PERSPECTIVE	PUBLICATION	RETAIN	STATUS
MUTUAL	PHASE	PUBLISH	REVEAL	STRATEGY
NEGATE	PHENOMENON	PURCHASE	REVENUE	STRESS
NETWORK	PHYSICAL	PURSUE	REVERSE	STRUCTURE
NEUTRAL	PLUS	QUOTE	REVISE	STYLE

SUBMIT	TRANSFER
SUBORDINATE	TRANSFORM
SUBSEQUENT	TRANSMIT
SUBSTITUTE	TRANSPORT
SUCCESSOR	TREND
SUFFICIENT	TRIGGER
SUM	ULTIMATE
SUMMARY	UNDERGO
SUPPLEMENT	UNDERTAKE
SURVEY	UNIFORM
SURVIVE	UNIFY
SUSTAIN	UNIQUE
SYMBOL	UTILISE
TAPE	VALID
TARGET	VARY
TASK	VEHICLE
TEAM	VERSION
TECHNICAL	VIA
TECHNIQUE	VIOLATE
TECHNOLOGY	VIRTUAL
TEMPORARY	VISIBLE
TENSE	VISION
TERMINATE	VISUAL
TEXT	VOLUME
THEME	VOLUNTARY
THEORY	WELFARE
THEREBY	WHEREAS
THESIS	WHEREBY
TOPIC	WIDESPREAD
TRACE	
TRADITION	