T.C.
AKDENIZ UNIVERSITY
INSTITUTE OF EDUCATIONAL SCIENCES
DEPARTMENT OF FOREIGN LANGUAGES EDUCATION

MA THESIS

LEXICAL CONTENT ANALYSIS OF A CORPUS-INFORMED ELT TEXTBOOK AND A TRADITIONALLY-PREPARED TEXTBOOK

MÜRÜVVET SÖĞÜT

ENGLISH LANGUAGE TEACHING
MASTER'S PROGRAM

Antalya, 2022

T.C. AKDENIZ UNIVERSITY
INSTITUTE OF EDUCATIONAL SCIENCES
DEPARTMENT OF FOREIGN LANGUAGE EDUCATION
ENGLISH LANGUAGE TEACHING MASTER'S PROGRAM

LEXICAL CONTENT ANALYSIS OF A CORPUS-INFORMED ELT TEXTBOOK AND A TRADITIONALLY-PREPARED TEXTBOOK

MASTER'S THESIS

Mürüvvet SÖĞÜT

Supervisor: Asst. Prof. Dr. Ersen VURAL

Antalya, 2022

ACKNOWLEDGEMENT

First of all, I would like to express my profound gratitude to my supervisor, Asst. Prof. Dr. Ersen VURAL who was abundantly helpful and offered invaluable assistance, and guidance. Completion of this thesis would not have been possible without his remarkable support.

In addition, I would like to express my sincere thanks to all the members of the monitoring committee, Assoc. Prof. Dr. Hüseyin KAFES and Asst. Prof. Dr. Feyzanur EKİZER, whose expertise and recommendations added to the richness of this research.

I am also grateful to my family, especially to my parents, for giving me constant encouragement and help anytime I needed during this study.

ABSTRACT

LEXICAL CONTENT ANALYSIS OF A CORPUS-INFORMED ELT TEXTBOOK AND A TRADITIONALLY-PREPARED TEXTBOOK

Söğüt, Mürüvvet

MA, Foreign Language Teaching Department

Supervisor: Asst. Prof. Dr. Ersen VURAL

June 2022, 143 pages

This study investigates the vocabulary profile of a corpus-informed ELT textbook and a traditionally-prepared ELT textbook in terms of their coverage of 2000 high-frequency words of New General Service List (NGSL). It also aims to demonstrate the similarities and differences in the lexical frequency profiles of each book, using a corpus-based approach. The sample course books analyzed are Touchstone 3 and Headway Pre-intermediate. Vocabulary Profiler Compleat and Frequency tools are utilized to profile the lexical content of the textbooks. The overall profile reveals that vocabulary coverage of the most frequent 2000 words of English (NGSL-1 and NGSL-2) is 74.2% in the corpus-informed textbook and 82.2% in the conventional course book, indicating that they are below the desired level for good comprehension and the lexical coverage of high frequency words in corpus-informed textbook is lower. Moreover, this study examines the type-token ratios (TTR) of the textbooks to find out the recycling rate of the vocabulary. The result shows that both course books, especially the corpus-informed textbook, need more systematic coverage and recycling of high frequency words.

Key Words: Vocabulary, word frequency, corpus-informed textbook, high-frequency words, vocabulary profiler, course book evaluation.

ÖZET

DERLEM DESTEKLİ BİR İNGİLİZCE DERS KİTABI İLE GELENEKSEL OLARAK HAZIRLANMIŞ BİR DERS KİTABININ KELİME DAĞARCIĞI PROFİLLERİNİN İNCELENMESİ

Söğüt, Mürüvvet

Yüksek Lisans, İngiliz Dili Eğitimi Ana Bilim Dalı

Danışman: Dr. Öğr. Üyesi Ersen VURAL

Haziran 2022, 143 sayfa

Bu çalışma, derlem destekli bir İngilizce ders kitabı ile geleneksel olarak hazırlanmış bir İngilizce ders kitabının kelime içeriklerinin genel İngilizcede en sık karşılaşılan 2000 kelimeyi (NGSL) kapsama oranlarını araştırmaktadır. Ayrıca bu araştırma derlem tabanlı bir yaklaşım kullanarak her kitabın sözcüksel sıklık profillerindeki benzerlikleri ve farklılıkları göstermeyi amaçlar. Analiz edilen örnek ders kitapları Touchstone 3 ve Headway Pre-intermediate'dir. Ders kitaplarının sözcüksel içeriğinin profilini çıkarmak için 'Vocabulary Profiler Compleat' ve 'Frequency' araçları kullanılmıştır. En sık kullanılan 2000 İngilizce kelimenin (NGSL-1 ve NGSL-2) kapsama oranlarının derlem tabanlı ders kitabında %74,2 ve geleneksel ders kitabında %82,2 olduğu ve bu kelimelerin kapsama oranlarının her iki kitapta da istenen seviyenin altında bulunduğu ortaya çıkmıştır. Bu sonuç aynı zamanda İngilizcenin en sık kullanılan kelimelerinin içeriğinin derlem tabanlı ders kitabında daha düşük olduğunu da göstermektedir. Ayrıca, bu çalışma, kitaplardaki kelimelerin tekrar oranını bulmak için ders kitaplarındaki kelime türlerinin (type) faklı kelimelerin (token) toplamına oranını (TTR) da incelemektedir. Sonuç, her iki ders kitabının, özellikle de derlem tabanlı ders kitabının, yüksek frekanslı kelimelerin daha sistematik bir şekilde kapsanmasına ve tekrar edilmesine ihtiyaç duyduğunu göstermektedir.

Anahtar Sözcükler: Kelime dağarcığı, kelime frekansı, derlem tabanlı ders kitabı, yüksek frekanslı kelimeler, kelime profili oluşturucu, ders kitabı değerlendirme.

TABLE OF CONTENTS

CHAPTER III

RESEARCH METHODOLOGY

CHAPTER IV

FINDINGS

CHAPTER V

RESULTS, DISCUSSION AND RECOMMENDATIONS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BNC: British National Corpus

CEFR: Common European Framework of Reference for Languages

CLT: Communicative Language Teaching

EFL: English as a Foreign Language

ELT: English Language Teaching

ESL: English as a Second Language

GSL: General Service List

L2: Second Language

NGSL-1: New General Service List I

NGSL-2: New General Service List II

NGSL-3: New General Service List III

AWL: Academic Word List

NAWL: New Academic Word List

CHAPTER I

1. INTRODUCTION

Learning a second or foreign language is becoming an essential aim for a large number of people with different purposes these days. As well as helping in everyday conversations when travelling, it may be a significant advantage for better career opportunities or getting education abroad. Moreover, acquisition of a new language is known to have a lot of mental benefits such as improving memory, creativity, thinking and also problem-solving skills.

Depending on the goal of learning an additional language, the skill needed to be focused may vary. No matter what area of the language such as reading, listening, grammar or writing is studied having a good command of it would require sufficient comprehension of the context. Regarding the comprehensibility of a text, having a grasp of lexical and grammar structures is necessary to decode the message in the minds of language learners. However, as Lewis (1993) argues "vocabulary or lexis carries more of the meaning of a text than does the grammar" (p.17). In other words, for a meaningful communication, correct use of words is more essential than correct use of grammar.

In their study based on a corpus of learners' errors, Doughty and Long (2003) observed that most of the mistakes made while communicating in another language derive from gaps in learners' knowledge of L2 lexis. In their study, vocabulary errors were found to occur in learners' productions three times more than grammar mistakes.

The significance of lexis can be better realized when we consider that a sentence with a grammatical mistake could still be understood but a lexical error would seriously impede the communication process (Gass, 1988). That can explain why vocabulary learning and adequate word knowledge for developing both spoken and written language skills are essential. Laufer and Sim (1985) emphasize the same point by noting that without sufficient vocabulary knowledge our understanding of the input would be hampered severely.

In traditional teaching contexts the emphasis had long been on improving learners' grammar knowledge, hence the significance of vocabulary used to be disregarded for a long period of time. In recent years, a growing number of studies have been researching the importance and the nature of vocabulary acquisition e.g., Carter and McCarthy (1988), Nation

(1990), Arnaud and Bejoint (1992), Coady and Huckin (1997), Schmitt (1997, 2000), Koosha and Sharifi (2004). In one such study, Meara (1980) commented that it is a discrimination to overlook vocabulary acquisition in the field of language teaching.

As a possible explanation for linguists' tendency to disregard the importance of lexical acquisition, Laufer (1986) proposed that vocabulary can be considered as an open set rather than closed system, which means that grammar and phonology lend themselves to generalization and abstraction. For this reason, there has been a tendency to work more on learners' mastering L2 grammar and phonology as they are more manageable aspects of the language for learners. According to Laufer (1986) another reason is that "development in semantics, sociolinguistics and pragmatics changed the view of language from abstract to that of a more idealized, social and functional phenomenon. Therefore, the interest has shifted from sounds and structures to meaning, discourse and speech acts" (p. 71).

Furthermore, with the change of focus from accuracy to fluency and communicative competence in language development, conveying the meaning and lexis became a crucial aspect of language teaching. Improving fluency was also emphasized by Krashen and Terrell (1983) especially at the time when Krashen (1981) put forward his ideas on comprehensible input, a hypothesis which suggests that the input we get in learning process should be understandable and a little beyond our current level in order to acquire the language effortlessly.

Although views on vocabulary teaching evolved with the developments in language teaching field and technology, one thing that has remained unchanged is the use of textbooks in the learning process, especially in language classrooms. It is widely agreed that ELT textbooks play a central role in most of language learning environments.

Even though much of the input learners receive is based on the ELT textbook in most language learning contexts, course books are criticized for presenting inauthentic language and vocabulary in their content. In this regard, it is not hard to see a gap in our knowledge about the extent to which language teaching course books reflect authentic use of language. There is very limited study on the nature of lexical content of course books. This is especially true for the corpus-based analysis of the authentic use of English language in their content.

Researchers have been using corpus-based approaches for the evaluation and the development of materials in language teaching field since the 1990s. A corpus is formed by gathering spoken or written data of native speakers' use of language from different sources

such as TV shows, radio broadcasts, magazines, newspapers and books. It provides invaluable information for applied linguists and many other researchers in related fields about how native speakers use of language in different contexts.

A corpus allows us to analyze language data and search for specific lexical or other linguistic information such as their context of use, frequency of occurrence, collocational relationships between other words.

Apart from showing language use in different situations, corpus use is crucial for teachers, researchers and material developers in their attempt to find out the most useful words and structures based on their frequency or to check how spoken language differs from written one.

A word frequency refers to how many times it occurs in a context or discourse. Frequency of use is an important aspect of lexical information, and it must be taken into consideration in all language learning processes. That is because there are countless words in a language, and we cannot commit all of them to our memory. It would be wise to concentrate on the most frequently occurring words and phrases in language in the process of selecting the content of language teaching materials.

1.1. Problem Statement

ELT textbooks are commonly used as a source in the field of language learning and teaching (Harmer 1991) to get learners engage with language in meaningfully designed tasks. ELT course books provide learners with most of the language input in the learning process; however, their vocabulary content is not usually analyzed exhaustively, especially from a corpus-based perspective. Researchers, teachers and textbook designers alike can benefit from the findings of the corpus studies and corpus-based analysis in their attempt to determine to what extent the content of the textbooks reflects the real use of the language and how the structures or lexis in course books deviate from that of native speakers.

In the last few decades the use of corpus-based approaches for the development and the evaluation of materials in language teaching field has gained importance. While research on corpus and corpus-informed materials has increased in recent years, many of the textbook authors do not draw on the results of these studies as the basis for the creation of the course books.

Moreover, studies reveal that there is a lack of systematic approach in the coverage, selection and distribution of vocabulary input in textbooks. It is claimed that for an efficient learning process, learners must master 1,500-2,000 high-frequency words in a shorter time span in the initial stages of learning. As these words are considered to be the most useful vocabulary, it is argued that they deserve more attention and a systematic approach in the teaching and learning process (Nation, 2006).

The use of corpora to determine the lexical content of textbooks, particularly high-frequency word lists that are derived from corpora, has challenged and revolutionized the conventional approaches to textbook design and analysis. However, only very few textbook authors have adopted corpus-based approaches in the process of developing language teaching course books. This apparent lack of interest in corpus informed approaches to developing materials and textbooks indicates that course book writers mostly rely on their intuition to select and compile the lexical content of the books.

In this regard, it is essential to analyze the difference between the lexical content of corpus-informed and traditionally prepared ELT textbooks for the second and foreign language instruction. Presentation of authentic and most frequent words of English in textbooks is an important aspect of the design process because it would facilitate communication for learners as they can readily use current words and phrases in everyday situations.

Given the importance of including high frequency words in language teaching course books, there is a need to carefully examine the lexical coverage of textbooks and materials. Certain course books available to learners in recent years have been touted as 'corpus-informed' by their authors. There has not been much research done to ascertain this claim about the lexical coverage of those textbooks. It is still not known whether course books defined as 'corpus-informed' vary by some means from conventional ones in terms of their lexical coverage.

Based on the need for investigating that aspect of lexical selection process in course book design, the present study is carried out to compare the coverage of high-frequency words included in both types of books and checking whether they are distributed and recycled in a systematic fashion.

1.2. Purpose of the Study

This study investigates presentation of vocabulary items in a 'corpus-informed' language teaching course book, *'Touchstone Level 3, Student's Book'*, and a traditional textbook, *Headway Pre-intermediate Level, Student's Book*, in terms of their lexical coverage and recycling patterns of the lexical items.

In order to evaluate their effectiveness in terms of vocabulary content, the two books will be compared based on the most frequent 2,000 words of New General Service List (NGSL).

The research questions addressed in this study are:

1. What is the coverage level of the most frequent 2,000 words of NGSL in a corpus-informed textbook *(Touchstone Level 3, Student's Book)* and a conventional course book *(Headway, Pre-intermediate level Student's Book)?*

2. What is the recurrence rate of words in a corpus-informed and a traditionally-prepared course book?

3. What is the recycling rate of the randomly selected words in two textbooks?

3.1. To what extent does occurrence frequency of randomly selected words in two textbooks reflect corpus-based frequency ordering?

1.3. Significance of the Study

With the advancements in the use of computers in ELT for collecting language data and forming corpora, it has become possible to see and reflect the common use of everyday English in many different contexts. It is crucial to integrate the results of corpus-based studies to design more effective corpus-informed textbooks.

Although there are clear guidelines in Common European Framework of Reference for languages about word selection, word frequency and grading, many of the textbooks do not adhere to such criteria. There is lack of a corpus-based and systematic approach in most of the textbooks.

In the process of textbook design, the following considerations are of utmost significance, to which the authors of textbooks are usually expected to pay attention:

i) how many words should be included in the textbook,

ii) which vocabulary items to cover and how to select them,

iii) required recycling rate and how to distribute the words in different units.

In this regard, a comparison of the lexical content of corpus-informed ELT textbooks with traditionally prepared textbooks may improve presentation of vocabulary items in terms of better coverage, distribution and more systematic recycling of vocabulary items.

In Turkish context, one can observe that only a limited number of studies adopted a corpus-based approach for vocabulary analysis of textbooks. A review of related literature on this subject would show that there is relatively little research addressing the comparative analysis of lexical content of a corpus-informed and traditionally prepared course book.

Results of this study is expected to raise awareness of teachers, researchers and textbook writers of the fact that there are high-frequency word lists extracted from corpora that should be given priority in determining the vocabulary profile of ELT books. Moreover, these high-frequency word lists may be used to determine which lexical items are infrequent and less useful in books and accordingly to exclude them in the design process.

Studies on teachers' perspective of corpus use would affirm that language teachers are not cognizant of the existence of such a rich authentic source, thus they are not expected to use corpora for developing materials. There is an obvious need for informing teachers about the significance of lexical selection regarding the materials adopted for classroom use. For this reason, helping teachers in the first place to analyze the textbook they use and to supplement them could facilitate the task of wisely selecting most useful lexical items in instructional materials used in the classroom.

1.4. Limitations of the Study

In this research the data is collected from the lexical content of two different English textbooks: the first is a corpus-informed course book and the other one is a traditionally-prepared course book.

A corpus-informed textbook is selected assuming that it is more sensitive to common and authentic use of English and a traditionally-prepared textbook is chosen since it is among the well-known English course books and it is easy to reach. In order to make the research content more manageable reading sections of the books are examined to profile the lexical coverage of the most frequent 2000 words of NGSL in two types of ELT books.

Therefore, both the number of books and sections investigated within the books may be considered to be extended. Lastly, the same type of comparative vocabulary profile analysis of textbooks can carried out by using different high-frequency word lists as the basis.

1.5. Glossary

Lexical content: Vocabulary components included in textbooks.

Vocabulary coverage: The number of high-frequency words included in a corpus-informed or a traditionally prepared textbook based on the corpus-assisted lists.

Vocabulary selection: The types of words that are selected and which criteria are used to choose them.

Vocabulary distribution: The number of types a vocabulary item is repeated in the textbook and the pattern of recycling within the different units.

Token: Total number of words in a text in which every repetition is included.

Type: The number of all the different words in a text in which the same words will be counted only once.

Type-token ratio: The rate of total number of items and different words in a text, which indicates whether there is diversity or repetition in vocabulary coverage of the textbook.

NGSL-1: The most frequent 1000 high frequency words of New General Service List.

NGSL-2: Second most frequent 1000 high frequency words of New General Service List.

NAWL: The most common 960 academic words of English.

Off-list: Textbook words which do not appear in any of NGSL and NAWL lists.

CHAPTER II

2. LITERATURE REVIEW

2.1. Acquisition of Vocabulary

One of the cornerstones in language learning is developing adequate knowledge of vocabulary. In order to master a language, knowing a considerable number of words is necessary. For this reason, rather than carrying grammar books, learners carry dictionaries with them (Krashen, 1989). Moreover, for the comprehension of sentences in texts and discourse, understanding words is the most important point. According to Lewis "Language consists of grammaticalized lexis, not lexicalized grammar"(1993). That is, in creating meaning words play a central role and the role of grammar is secondary.

Even though mastering a wide range of vocabulary is needed to achieve a meaningful communication, systematization of the vocabulary has been a neglected aspect of the foreign language teaching. Studies in second language acquisition has typically researched the size of vocabulary learners have (breadth of the vocabulary) rather than how well (the depth) the words are known (Vermeer, 2001). The nature of vocabulary acquisition, how it relates to other language skills and its role in accuracy and fluency began to be further researched when Communicative Language Teaching attracted more attention. Importance of vocabulary in language proficiency has been studied by a growing number of researchers especially since the emergence of CLT (Laufer &Hulstijn 2001; Nation, 2001). The reason behind it is that language fluency has been associated with having a good command of the vocabulary in the language (Laufer & Goldstein, 2004).

According to Nation (2001), a word's meaning, form and its use all make up the knowledge of a word. However, there are also other factors such as the style of the language, associations and collocations (Nation, 2013). Nordlund (2016) remarks that "learning new words involves the acquisition of vocabulary breadth as well as depth" (p. 50). Vocabulary breadth is acquired by getting the knowledge of different ways of expressing yourself and synonyms, and depth of vocabulary is important to gain the ability of idiomatic use of language such as selecting the correct collocations while communicating (Nordlund, 2016).

When some of the mentioned points are neglected in vocabulary learning process, learners cannot have the complete word knowledge, which may lead to an imbalance between

receptive and productive abilities. A learner may not be able to produce the word orally, but recognize it in a context, which is described as “degrees of word knowledge” by Melka (1997, p.88).

To be able to use the lexical items automatically, for gaining receptive and productive word knowledge (Tyler, 2012) encountering words a lot of times in varied contexts is essential.

Our brain stores the words in semantic networks, connected to each other (Aitchison, 2012). When the frequency of the context-rich encounters increases, only then learners construct deeper connections. This concept of forming deeper connections is also related to the levels of processing (Craik and Lockhart, 1972), a principle which asserts that when words are more deeply processed, it will result in a better storage in the long-term memory, which, in turn, will lead to easier retrieval of the words. Thus, this process results in internalizing and remembering the vocabulary in a better way (Nation, 2008).

2.2. Frequency Effect in Language and Vocabulary Learning

With corpus linguistics gaining attention in the field of language teaching and advances in computer technology, research on the role of word frequency on retention of words in memory grew in number (Nation, 1993a, Nation 1993b, Schmidt 2000). The term ‘frequency’ here is used to refer to the number of times a word is repeated in a context. The most frequent words in language are the ones that are used most frequently in daily language such as the word “*and, the, a, you*”, while low frequency words are the ones that are not commonly used in everyday language such as the word “*thine*”. It is therefore to be expected that learners’ level of familiarity with L2 words that occur frequently in language is higher than those that do not. This has a wide implication for lexical learning as it has strong impact on memorization, lexical access, word storage and retrieval.

The likely reason for the positive influence of high word frequency on memorization is the high amount of exposure learners are subject to (Laufer 1997). Similar to that, how learners access lexical items in their mental lexicon (lexical access) and how the lexis can be most effectively organized (lexical storage) are positively affected by frequency of vocabulary. Based on that, it can be concluded that to optimize lexical processing and retrieval, the number of times the word is repeated must be high.

As regards the biological bases of acquisition of knowledge, frequency and repetition have been shown to be determinants of language learning as have been evidenced by growing body of research in neurolinguistics (Tomasello& Brooks 1999; Bybee 2006; Anderson 2010). As Criado and Sancez (2012) explain "Neurologically, repetition of the same action affects the structure of specific synapses or connections among neurons, thus rendering future and similar tasks easier, and allowing for reaching automaticity after several trials."

How the tasks are performed automatically with repetition effect can also be understood by looking at Newell & Rosenbloom's (1981) Power Law of Practice theory, in which the role of practice frequency on learning is defined. In his explanation of Power Law of Practice, Ellis (2002a) states that the number of times a task is practiced is related to the speed of performance. As the word construction or memory is strengthened with each repetition, remembering high frequency words becomes easier.

Moreover, word recognition, processing and production stages get automatic and the task involved in the process are readily performed in parallel with frequency of word exposure in meaningful context. Thus, language fluency is thought to be facilitated with repeated practice even when the learner is not processing the information consciously as in the case of incidental lexical acquisition.

Divjak (2019) further support the argument of some researchers that by encountering a word more often, the threshold of activation in our memory gets lower. Threshold level is the minimum level of proficiency that the learners should reach to benefit from the use of a word meaning. With a lower threshold level for a word, learners would need less additional activation to recognize the word. In other words, recalling the word would take shorter time and effort.

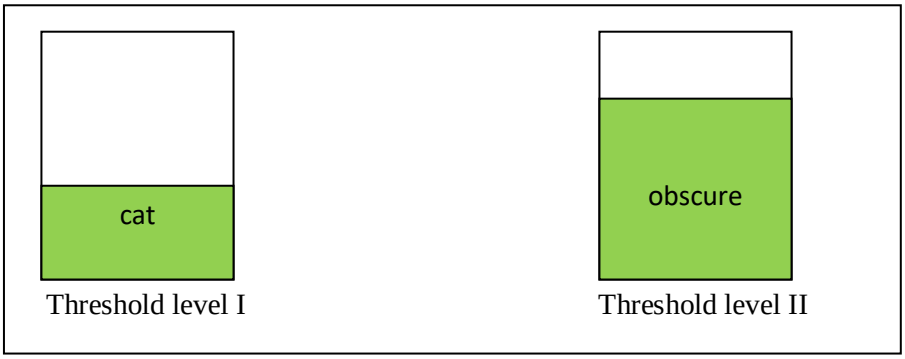

Figure 2.2 Comparison of Threshold Levels of Two Words

As can be seen in Figure 1, ‘cat’ is a word with higher frequency, so the activation level needed to reach this word is low. However, ‘obscure’ occurs less frequently in English contexts when compared to the other word and it requires a higher level of activation to recall this word. Therefore, it can be said that lower frequency words have higher threshold levels.

Research studies in psycholinguistics and cognitive sciences indicate that language acquisition is actually the result of a pattern-detection process, in which learners discover, gather, abstract and strengthen the rules unconsciously (Atkinson & Shiffrin 1968; Sinclair 1991; Ellis 2002; Tomasello 2003). Thus, it can be said that exposure to recurrent language samples results in form-meaning and form-function associations in the brain. Through frequent linguistic input, linguistic codes emerge and are strengthened through interactions during communication.

Repetition of experiences plays the biggest role in most learning situations and human activities for making the knowledge permanent and performance accurate (Criado and Sancez 2012). Also, in usage-based approach scholars like Tomasello (2003) and Behrens (2009) emphasize the importance of input frequency claiming that learners construct language knowledge by getting exposed to the language very often. When a word or structure is less frequently encountered, it is perceived as more complex by learners.

It is agreed that encountering words several times is necessary for gaining the receptive and productive ability to use a word; however, there is no agreement on the number of times needed this to happen. Waring and Takaki (2001), for example, maintains that 8 times or more repetition is required, whereas Cameron (2001) suggest 5-6 encounters would be enough to learn the new words. On the other hand, Schmitt (2008) disputes that even 8-10 exposures would only improve learners’ recognition of the words. The number suggested by Nation and Wang (1999) is ten times and they also assert that meeting lexical item more times would result in better retention. Despite the differences in numbers suggested for retention, it is clear that learners need to meet the words more often to be able to use them appropriately with ease rather than just understanding their meanings.

In addition to frequency as a condition for lexical retention there are other considerations involved. According to Nation (2008) and Matsuoka (2012) for long term lexical retention to occur learners should also benefit from ‘the spacing effect’. That is, spreading vocabulary learning over long periods can reinforce the teaching process and yield better results.

In the past, lexical frequency information was already employed in Audio-lingual method to decide on what vocabulary to choose for teaching at different levels (Sanchez, 2009).

2.2.1. Frequency in Corpus studies

Informed by the recent findings of corpus studies, many applied linguists such as Nation (2006) have posited that the most frequent words in language should be prioritized in language teaching. It stems from the fact that high-frequency words cover a big percentage of any type of text or discourse and they are learnt more quickly as compared to the low-frequency words. As a result of his analysis of nine spoken and written corpora, Nation (2006) found out that coverage of BNC' first 1000 word families is almost 80%.

In addition, language proficiency, communicative efficiency, is greatly determined by a learners' knowledge of most frequently used lexical items. This can be taken as the basis for the assumption that design of language teaching textbooks could also benefit from such information in order to reflect the real language usage. In more precise terms, it can be said that high-frequency word lists are expected to facilitate language learning process. Nevertheless, few textbook designers are guided by the findings revealed by corpus studies which attempt to analyze actual language use on the basis of lexical frequency. Lexical frequency bands created with the help of corpus studies should serve as criteria for vocabulary selection in the process of textbook design.

For elementary level students, the most frequent 800-1000 words were assumed to be necessary to learn, whereas the next 800-1000 or 2000 words were considered to be introduced to intermediate and advanced students. As the vocabulary analysis tools have advanced up to the present, helpful and reliable results are available for both teachers and textbook authors (Criado and Sanchez, 2012). Today it is quite easy to reach reliable information on the frequency lists of English vocabulary or any word frequency list of a specific context.

Important role of frequency in vocabulary acquisition is unquestionable (Webb 2007; Schmitt 2008; Nation 2008; Ellis 2013). Ellis (2012) emphasizes the frequency effect as "a key determinant of acquisition" (p.144).

In 1944, Thorndike and Lorge published a list of English word frequencies called 'The Teacher's Word Book of 30,000 Words'. Howes and Solomon (1951) conducted an

experimental study on the effect of word frequency which concluded that recognition of frequent words is much faster than the others. Another research showed that word recall and word frequency correlated positively (Hall, 1954). Cortese and Scarborough (1977) observed that frequently encountered words are accessed more readily than words with lower frequency. MacKay (1982) discovered that retrieval of high-frequency items is less subject to error.

2.2.2. Frequency Effect in Vocabulary learning

With growing body of research on the effects of exposure frequency on vocabulary learning, many studies drew attention to the importance of repeated exposure on processing and acquiring words, even in learning grammar.

In order to control the frequency and exposure effect on language, frequency lists were introduced. Research in this area emerged from corpus studies emphasizing that most frequent words in the language must be prioritized to meet the needs of the learners more efficiently (Horst, 2013).

Nation and Waring (1997) claim that with growing interest in teaching communicative aspect of the language, less emphasis is put on choosing and checking of language forms in designing courses. As attitudes change regarding the importance the form-focused instruction, benefits of explicit teaching of language forms such as syntax and lexicon were realized. This shift in attitude is also expected to change the way vocabulary teaching has conventionally been viewed. Today there is a consensus that textbook design process should make better use of frequency studies and vocabulary lists deriving from corpus studies.

2.3. Selection of Vocabulary in ELT pedagogy

Some of the well-known vocabulary lists that could be helpful for both learning and teaching English are: The Teacher's Word Book of 30,000 Words (Thorndike and Lorge, 1944) based on large written corpus of 18,000,000 words containing nearly 13,000 word families; The General Service List (West, 1953), based on a written corpus of 5,000,000 words, consists of 2,000 headwords; The Brown (Francis and Kucera, 1982) is based on 1,000,000 word written corpus; the University Word List (UWL) (Xue and Nation, 1984; Nation 1990) including 836 general academic words which could be studied as an additional list to the GSL for especially academic purposes.

Among the aforementioned word lists, West's General Service List (1953) is widely considered to be the most practical one for learners, teachers and textbook and course designers. Many years after the publication of the West's (1953) General Service List (GSL), an updated version of the list called New General Service List (NGSL) was prepared based on a larger (273 million-word subsection of the 2 billion word Cambridge English Corpus) corpus with a better coverage by Dr. Blent Culligan, Joseph Phillips and Dr. Charles Browne in 2013. With the creation of NGSL the coverage rate of GSL (84%) was increased to 92%, which provides understanding of the most of the vocabulary in many general English texts.

Word lists created with the use of corpus frequency information started to be used for educational purposes in the past. Since they can be used to expand learners' vocabulary knowledge in a targeted way, they gained popularity.

Nation and Meara (2002) maintain that studies of word frequencies provide a cost-benefit way in vocabulary teaching and learning. They also suggest that in addition to the most frequent words of English, different kinds of special purpose word lists can be prepared considering the learners' needs.

A common result of frequency counts is that 2,000 words of highest frequency cover almost 85% of the lexical items of any kind of texts in a book. Concentrating on widely used words would bring a useful return since their coverage is even bigger in spoken English (Nation and Newton, 1997).

2.4. Effect of Textbooks on Learners' Vocabulary Knowledge

Textbooks are the most common tools in almost all language learning contexts and they are used as the basis for both teaching and testing the content. Therefore, it is crucial that what is presented in materials and textbooks should be supported with the research results and pedagogical considerations. There is, however, very limited study on the lexical content of course books and especially on the corpus-based analysis of their authentic use of English. Although much of the input learners receive is based on textbooks they seem to fall short of reflecting actual hierarchy of lexical importance in language use.

EFL learners may not need to know and use all vocabulary in the course book. Therefore, studies analyzing the course books in terms of usefulness of their vocabulary perspective are growing in number and they deserve special attention.

When deciding which words to introduce and how to select these words, *Vocabulary Profiler* (www.lextutor.ca) can be utilized. It profiles the vocabulary coverage in a textbook by comparing them with the list of high- frequency words. A well-established word list, the New General Service List (http://www.newgeneralservicelist.org) can be used for comparison to provide information about the proportion (percentage) of high-frequency words used in a textbook. In addition, the output provides information about negative vocabulary word frequency (high-frequency words that aren't included).

Systematic analysis of lexical input in course books can be effectively used for the development of students' vocabulary knowledge and designing more effective teaching materials. In their review of the research on vocabulary component of teaching materials, Nordlund and Norberg (2020) observed that studies revealed three common lacking points in materials: words are not selected with a systematic approach, repetition of the new words is not sufficient within the book, and lexical items do not adequately reflect the authentic use of English.

There are both non-corpus studies and corpus-based studies carried out for assessing the textbooks. The first group of studies focusing on the lexical content in ELT materials and textbooks is non-corpus studies. From a non-corpus based perspective, Scholfield (1991) conducted a study to find the rate of newly-introduced words and recycling of old words in EFL course books and concluded that there is a conflict about the rate of new items to be introduced within the different units and among the books. Martinez (1999) worked on the vocabulary activities in textbooks to see to what extent authors took the language learning theories into account. The conclusion was that textbook writers ignore much of the research on language learning.

As regards corpus-based research on the analysis of vocabulary input in ELT materials and textbooks, the studies can be divided into three groups: (1) Research on EFL syllabus design with the integration of corpora (Sinclair&Renouf 1988; Willis& Willis 1988; Flowerdew 1996), (2) Studies on application of concordancing programs in ELT (Johns 1989; Tribble 1990; Stevens 1991; Gavioli 1997), (3) investigation of existing ELT materials and textbooks.

There are more studies on the use of concordancing programs and corpora to design teaching materials when compared with the corpora use for the analysis of materials and textbooks. The research on the assessment of existing ELT materials and textbooks is limited in which attention has been paid mostly to the number of words included in the textbooks. The studies differ in their focus concerning the textbook type, their scope and the language researched.

While the corpora use has increased making the authentic language access much easier than in the past, Harwood argues that "the language taught in commercial materials differs markedly from the language that is actually used in spoken and written discourse" (2010, p. 9).

Lexical content of EFL textbooks was studied in a variety of research conducted (Takala 1984; Kaszubski 1998; Miranda 1990; Benitez Perez & Zebrowski 1993). However, findings demonstrate that the rationale behind the vocabulary selection in the current EFL/ESL context is not fixed yet. Takala (1984) and Miranda (1990) analyzed the vocabulary component of secondary school textbooks. Another study focusing on the distribution of vocabulary items in Spanish textbooks is Benitez & Zebrowski's (1993) research. Investigation of vocabulary profile of English textbooks from different disciplines has been addressed by Kennedy (1994). These studies on the comparison of corpora and vocabulary profile of textbooks revealed a large number of infrequent words in everyday English and a disagreement on number, selection and repetition of lexical items in textbooks.

Academic texts aimed at English native speakers were analyzed in terms of their vocabulary input in another study (Sutarsyah, Nation& Kennedy, 1994). In 1991, Ljung compared the authentic use of English in textbooks in her study. Criado (2009) examined the recycling rate of different words in English teaching materials and reported that 58% of the words were repeated once or twice. Likewise, Erman (2009), Fujimori (2005), Norberg and Nordlund (2018) determined that the number of times words occurred in textbooks did not provide sufficient exposure for language acquisition.

Similar conclusions were made in a study by Konstantakis and Alexiou (2012) analyzing the lexical content in five EFL primary school textbooks used in Greece. They determined that the books were loaded with mid- and low-frequency words, which led to inclusion of only insufficient number (between 74% and 85%) of the high-frequency words on the BNC list.

According to Thornbury (2002) main problem with textbooks is deciding on which words to include from countless words in language. In many studies taking word frequency into consideration is suggested for course book authors and teachers. On the other hand, Reda (2003) identified that vocabulary content of many ELT books are *topic-based* and range of the items contained are quite limited in different books since the topics are very similar. Similar results were observed in the research related to vocabulary conducted in Turkey which mostly worked on teaching techniques. Also, studies focusing on textbooks neglected the lexical aspect and assessed types of tasks, gender and appropriateness of curriculum instead (Ünlü, 2012).

In another study, Agan (2008) investigated how comprehensible the EFL book used in public high schools was by comparing the lexical content of the book with the vocabulary size of the students. The findings showed that the book was not appropriate enough to the vocabulary levels of the students.

The research conducted in Turkey by Ünlü (2012) on loading, distribution, and repetition patterns of the 2000 high-frequency words of general English in an EFL course book revealed that the book covered only limited number of the most frequent words and neglected dimensions of vocabulary component in the textbook requires considerable attention.

The number of the words should also be considered as well as which words to include in course books. For this purpose, corpora-based high-frequency word lists might be useful in determining the words and Nation's (2001, 2006) way of counting words as token, type, word family, lemma could be used for quantifying the vocabulary component. In Nation's classification of the words, every word in a text is referred as a "token" even if the same word is repeated twice. For instance, in the sentence 'It was her decision to talk to her' number of tokens is eight. When a word is repeated more than once in a context, but counted only once it is referred as "type". In the previous example, 'It was her decision to talk to her' the number of types is seven. Two words are referred as "lemmas" when a headword occurs with its inflected forms and they are counted as single, for example student and students? "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). Therefore, the words come, comes, came and coming can be accepted from the same word family.

Among all of the studies mentioned above, none of them worked on a corpus-informed English textbook. Moreover, there is no research addressing the comparative analysis of lexical content of a corpus-informed and traditionally-prepared course book.

CHAPTER III

3. RESEARCH METHODOLOGY

3.1. Research Design

Use of corpora in language teaching has been gaining more attention recently since the emergence of corpus linguistics. Taking a corpus-based approach in language learning has important advantages. There are studies demonstrating how corpora can be used for textbook and material development and course book evaluation. Among these, most practical use of corpora is the application of it for the purpose of curriculum and textbook development and evaluation since the use of corpus gives us the opportunity to compare the content of the textbook with authentic data and update the English course books and by replacing their content with more realistic data.

In order to use corpus linguistics methodology, technology needed is a computer to store collected text files and software that is applied to the corpus of texts to produce word frequency lists. Also, software can sometimes be used to compare the words in a corpus with already existing word frequency lists. Thanks to the use of computers, extensive analysis may be carried out on a corpus through simple programs that save time.

In this research, corpora containing two different English textbooks will be investigated using a quantitative corpus linguistics methodology and purposive sampling method focusing on the lexical content of each book by means of a vocabulary profiler tool on a website, Compleat Lexical Tutor (www.lextutor.ca). With the help of this corpus linguistics program, detailed information will be obtained from both English textbooks about the frequency of vocabulary items, the range of words across the books, and their recycling rates. The results of two different corpora of textbooks are compared to analyze their authentic use of English language.

3.2. Instructional Materials Analyzed

3.2.1. Corpora of Two EFL Course Books

Two EFL corpora were compiled to analyze the data extracted from two different ELT textbooks.

The first book analyzed in the present research was *Headway Pre-intermediate Student's Book*. This traditionally-prepared textbook was written by Liz & John Soars and Paul Hancock. The book was published by Oxford University Press and the first edition of the book was printed in 1998. The fifth edition printed in 2019 was selected for this study. It was chosen because the book is published and used globally and it does not claim to be a corpus-informed textbook. While the new edition has been updated in terms of its methodology, text contents, integration of authentic language and digital resources, it is still a topic-based book.

The content of the book is claimed to be based on a well-balanced grammar and skills syllabus. The book is structured in 12 units. In each unit, there are six sections: grammar, vocabulary, reading-speaking, listening-speaking, everyday English, writing. (See table 1)

Unit 1	**Content**
Section 1	Grammar
Section 2	Vocabulary
Section 3	Reading-Speaking
Section 4	Listening-Speaking
Section 5	Everyday English
Section 6	Writing

Table 3.2.1a Headway Pre-intermediate, Unit 1 Content

The second book examined in this study was a textbook that claims to be corpus-informed, *Touchstone 3*, completing level A2 and entering B1 level of the Common European Framework of Reference for Languages (CEFR). Touchstone 3 was written by Michael Mccarthy, Jeanne Mccarten, Helen Sandiford and was published by Cambridge University Press.

The content of Touchstone 3 is based on the North American English portion of the Cambridge International Corpus (currently known as The Cambridge English Corpus). This corpus contains multi-billion word data from different sources including written and spoken, British and American English to build a syllabus reflecting how people actually use English.

In the introduction of the book, it is claimed that authors found out the most useful and frequent grammar and vocabulary as well as typical communicative uses in everyday situations, spending years searching the Cambridge International Corpus. Therefore, the book is considered to contain reliable and authentic information for teachers and learners.

Touchstone 3 is structured in 12 units. In each unit, there are four sections: Lesson A, Lesson B, Lesson C and Lesson D. All of the lessons in different units mostly have similar content. (See table 3.2)

Unit 1	**Content**
Lesson A	Getting started, grammar, speaking naturally
Lesson B	Building vocabulary and grammar, grammar, listening and speaking, vocabulary
Lesson C	Conversation strategy, strategy plus, talk about it
Lesson D	Reading, listening, writing and speaking, free talk

Table 3.2.1b Touchstone 3, Unit 1 - Lesson Content

For the purposes of practicality and their representative property this study focused only on ‘reading spots ‘of the books and the results were generalized to the whole book. That would be a reasonable decision given that there exists a strong relationship between vocabulary learning and reading practice, which forms the bulk of the classroom activities. In some classroom practices it could be observed that parts of the units like listening, writing and speaking are not prioritized as much as reading sections.

3.3. Data Collection Tools

3.3.1. Lextutor Program for Vocabulary Profile Analysis

Profiling lexical content of the two books was carried out using a corpus linguistics program called Compleat Lexical Tutor. Designed by Tom Cobb, the software has a collection of tools for researchers, teachers and learners to study vocabulary and to analyze texts in terms their lexical content. Compleat Lexical Tutor is free of charge and available on a website

(http://www.lextutor.ca/). The present study attempts to research the use of most frequent words of English in two ELT textbooks, hence the use of such software allows us to upload our textbook content and identify the word frequency and recurrence rate of lexical items. Instructions on how to use the program are provided on the screen of various features selected.

To answer Research Question I, to find and compare vocabulary coverage in two different textbooks' content, Vocabulary Profiler Compleat tool within Compleat Lexical Tutor program was used. The tool reveals the complexity of texts by comparing them with the lists of high-frequency and academic English words incorporated in the program.

As a result of the comparison, an output file is generated. The results give us the percentage of the frequent and academic words in our corpus. The output file also shows which of the high-frequency and academic words are included in the textbook content. Words such as abbreviations, specialist vocabulary and proper nouns are named off-list in the output. Also, textbook words which do not appear in any of the lists are classified as off-list. Users are provided with concordance lines (contexts including certain words) based on the uploaded corpus on Compleat Lexical Tutor, too.

There are several word lists available within the program to compare your text with. These corpora are the Brown Corpus, the British National Corpus (BNC), the New General Service List (NGSL) and New Academic Word List (NAWL).

3.3.2. Vocabulary Profiler Tool and the Base Word Lists

A quantitative approach with content analysis design was used in this research. The data was analyzed with the help of lexical analyses using two different vocabulary profilers on Lextutor, an online lexical analysis software accessible at http://www.lextutor.ca/.

To answer Research question I, *Vocabulary Profiler Compleat* (VP-Compleat) was used. It freely gives users access to classify the vocabulary items in the textbook into frequency levels.

To perform lexical analyses, VP Compleat tool makes use of several high-frequency lists that serve as *base lists* such as *New General Service List (NGSL)* (Browne et al, 2013) and *New Academic Word List (NAWL)* (Browne et al, 2013). They contain well-established, comprehensive and up-to-date high-frequency lists of English and Academic English words.

New General Service List is an updated version of the West's (1953) General Service List (GSL). NGSL was prepared based on a larger (273 million-word subsection of the 2 billion word Cambridge English Corpus) corpus with a better coverage (Browne et al, 2013). With the creation of NGSL the coverage rate of GSL (84%) rose to 92%. The high NGSL coverage level in any text means that leaners can comprehend most of the vocabulary in many general English texts.

Coxhead developed the Academic Word List (AWL) consisting of 570 word families that are frequently included in academic texts. The NAWL is a list different from that of Coxhead's. Developed by Culligan, Phillips and Browne (2013), NAWL contains nearly 960 words taken from a 288 million word corpus including the contents of academic course books and academic lectures. It was created as an extension of the NGSL and the combination of these two up-to-date lists gives a comprehensive coverage of the most common English words. These reference word lists are available within the program along with their inflected forms and the headwords.

The NGSL-1 consists of the most common 1000 words that appear in a given English text. The NGSL-2 words include second most common 1000 English words and they are found less frequently in a text. Thus, NGSL-1 words may be considered as easier to understand and more frequent when compared with NGSL-2 words. The most common words in academic texts are given in the New Academic Word List (NAWL).

VP-Compleat tool containing those base word lists are used to identify which words in the textbooks belong to the most frequent 1000 English words (NGSL-1) and 2nd most frequent 1000 English words (NGSL-2) in the study. The total number of the words in the New Academic Word List (NAWL) in the books was researched as well as off-list words in the textbook which do not appear in any of the NGSL and NAWL.

The website offers many different tools helping to learn and analyze vocabulary. Among these tools on Lextutor, *Vocabulary profiler* and *Frequency* tools were used for lexical profiling and lexical frequency analyses for the present research.

By selecting the 'VocabProfile' section, the lexical profile analysis tools can be accessed. There are three profilers on the VocabProfile page: VP Classic, VP-kids, and VP-Compleat.

VP-Compleat is the most-recently developed vocabulary profiler that gives more precise results about lexical frequency of the words. Therefore, it was chosen for the lexical content analysis in this study.

On Lextutor website, overall lexical profile of the text is also presented by giving the proportions of vocabulary as lemmas (when a headword occurs with its inflected forms and they are counted as single), types (when a word is repeated more than once in a context, but counted only once) and tokens (when every word in a text is counted separately even if they are repeated).

3.3.3. Frequency Tool on Lextutor

To answer Research Question III, 20 words with different frequency levels were selected from the most frequent 2000 words of English (NGSL1 and NGSL2). Lexical frequencies of these 20 words were compared with the frequency ordering of the words in the lists NGSL1 and NGSL2 by using the *Frequency* tool found on the website Lextutor. This tool reveals the frequency of each word within a text and lists them in the order of their recurrence rate from the most used to the least used..

A frequency list of all the words in the text is created in the output file. It also gives the rank order of the vocabulary items, their frequency figure (how many times they appear within the text) and cumulative percentage of their frequency.

3.4. Data Collection Procedure

3.4.1. Preparing the Texts for the Vocabulary Profile Analysis

Headway Pre-intermediate Student's Book and Touchstone 3 Student's Book are the sample textbooks selected for vocabulary profile analysis. Reading spots in two textbooks were the focus of this study. Since this study sampled two textbooks, comparing reading spots of a conventional textbook with reading texts in a corpus-informed textbook could give us sufficient results to have an idea on their vocabulary content. Vocabulary load analysis was done in the reading spots of the whole book based on the 2000 high frequency words of New General Service List (NGSL1 and NGSL2) and New Academic Word List (NAWL).

Reading texts were extracted from all of the units in the books considering that they would represent the whole textbook. However, Headway Pre-intermediate Student's Book consists of longer reading texts. In order to compare equal number of words from the two books, nearly 300 words from each unit were taken equaling to almost 3700 words in total in each book.

The textbooks were in pdf file format. Therefore, the first step of preparing the texts was transforming pdf files to word documents using Abby Fine Reader version 14.0. Abby Fine Reader is a PDF transformer program. However, during this process there were problems with some parts due to the illustrations on the pages or page formats. Therefore, a few parts of the reading spots were written by hand on Microsoft word program. The texts had to revised and prepared for the analysis after completing transforming and writing process. It was done in order to obtain more reliable results with the vocabulary profiler tool.

Firstly, the texts were checked for the spelling mistakes and corrected. The tool VP-Compleat requires an empty space after each punctuation mark to run smoothly. Including missing spaces was the second step in the text revision process.

It's good to have family, friends, and a partner. But you can't always be in the company of others, so it helps if you can really enjoy your own company and feel confident on your own. So, where are you on the scale of dependence and independence? This quiz will help you find out! In a group, can you express an opinion that all the others disagree with strongly? a) No, because I don't like feeling that I'm not part of the group. b) Yes. I enjoy being the one with a completely different view on things. c) Yes. I find it difficult to do, but I can't say anything.

Table 3.4.1a The Original Extract from the Headway Pre-intermediate - Unit 2

As can be seen in Table 3.4.1a and 3.4.1b, thirdly, words in short forms were changed into long forms.

It is good to have family, friends, and a partner. But you cannot always be in the company of others, so it helps if you can really enjoy your own company and feel confident on your own.
So, where are you on the scale of dependence and independence? This quiz will help you find out! In a group, can you express an opinion that all the others disagree with strongly?

a) No, because I do not like feeling that I am not part of the group.
b) Yes. I enjoy being the one with a completely different view on things.
c) Yes. I find it difficult to do, but I cannot say anything.

Table 3.4.1b The Modified Text from the Headway Pre-intermediate - Unit 2

As can be seen in Table 3.4.1c and 3.4.1d, proper names were deleted from both of the textbooks in the next step.

Alicia Keys – Pianist and singer-songwriter
Alicia Keys was born and raised in New York. She started playing the piano at the age of five. Her mother was very supportive and always encouraged Alicia to continue playing. She got excellent grades in high school and graduated at the age of 16.A versatile performer, Alicia released her first CD when she was 19. She wrote one song on the CD when she was 19.Her secret talent? She is good at swimming.

Table 3.4.1c The Original Extract from Touchstone 3 - Unit 1

……. ……. – Pianist and singer-songwriter

……. ……. was born and raised in ……. ……. She started playing the piano at the age of five. Her mother was very supportive and always encouraged ……. to continue playing. She got excellent grades in high school and graduated at the age of 16. A versatile performer, ……. released her first CD when she was 19. She wrote one song on the CD when she was 19. Her secret talent? She is good at swimming.

Table 3.4.1d The Modified Text from Touchstone 3 - Unit 1

Eliminating the instruction parts was the last step of the preparation process since this study focused directly on the reading spots.

The modified texts of two course books were saved on separate word documents. And then they needed to be copied into the tools *Vocabulary Profiler Compleat* and *Frequency* in sequence for analysis on Lextutor website.

3.4.2. Vocabulary Profile Analysis

The vocabulary load analysis of the textbooks was done in terms of the frequency of the vocabulary items. The texts from both course books were ready in separate documents as word-files.Only 'reading spots' of the textbooks were taken and revised to get more accurate results.

Firstly, *VP-Compleat tool* was used to analyze the coverage of high-frequency words in both course books. Instructions are given in the white space that is used for feeding the text into the profiler. The text in word-file format can be copied and pasted in the space after deleting the instructions.

Contents of textbooks were copied and pasted on the VP-Compleat tool's website after being revised. There are different high-frequency word lists within the program that are used as base lists. For this study, NGSL and NAWL were selected because they reflect comprehensive and updated common English words. The output of the analysis is presented after choosing the base lists and clicking on 'Submit Window' button. The output provides information about the number and percentage of lexical items contained in each base list.

All the reading spots of *Touchstone 3* were profiled as a whole and the same process was repeated for the reading spots in *Headway Pre-intermediate Students' Book*.

The result of the analysis shows the percentage and the number of the most frequently used words: NGSL-1, NGSL-2, NGSL-3 and NAWL in the course book and off-list words. The output file reveals which of the high-frequency and academic words are included or which words do not belong to any of these lists. This gives us an idea about the complexity and efficiency of the course books.

Lemmas, types and tokens are given in the first row of the output file. Headwords in a text are lemmas, for instance: the headword of *information* and *informing* is *inform.* Different words are types, for example: *speak* and *speaking* are counted as different words. Token is the total number of words in a text, for instance: if the text includes the words common (2), take (3) and people (1), the number of tokens is 6. Numbers under these headings assert how many words in the text belong to each of these categories.

Different frequency levels are given in the first column of the output file. The words in the texts are grouped according to their frequency levels as *NGSL-1* (the most frequent 1000

words), *NGSL-2* (second most frequent 1000 words), *NGSL-3* (third most frequent 1000 words) and *NAWL* (the most common Academic words of English) and off-list words that do not belong to any of these categories.

In this study, the main aim was to compare the proportion of the most frequent English words from different aspects. Therefore, the results were used to answer Research Question I, which aimed to compare coverage levels of high-frequency words in two ELT textbooks.

In addition to the lexical frequency figures, VP-Compleat tool shows some related ratios and indices in the output file. For this study, type-token ratio (TTR) is taken into consideration. TTR is obtained when the total number of word types is divided by the total number of tokens. The result is a number that falls within 0 and 1. It allows the researcher to see to what extent the words in the textbook are repeated, which is the answer to the Research Question III.

3.4.3. Selection of 20 Words from the 2000 High Frequency Word Lists

The present study focuses on the vocabulary profiles of two ELT textbooks based on the most frequent 2000 words of English. A set of 20 randomly sampled words with their inflected forms was taken from NGSL-1 and NGSL-2 to answer the Research Question III, which aimed to find out to what extent the frequency of 20 randomly selected words in two textbooks reflected their NGSL based frequency ordering. The words were chosen from both NGSL-1 and NGSL-2 to increase the validity of the results.

In the selection process the list standings of the words were the main criteria. The words were selected within a certain distance on the list between each other i.e. It was made sure that there was min. 46 and max. 141 words between two words selected. To illustrate this we can take two selected words *make* and *home* as examples. The list standing of the word *make* is 48 and it falls within the range 0-100 most frequent words. The other word *home* stands as the 161th word on the frequency list. The distance between two is 113 words.

NGSL words & their frequencies	Inflected forms
1. make (48)	**made, makes, making**
2. home (161)	**homed, homes, homing**
3. buy (242)	**bought, buying, buys**
4. reach (380)	**reached, reaches, reaching**
5. shop (437)	**shopped, shopping, shops**
6. listen (573)	**listened, listening, listens**
7. forget (651)	**forgets, forgetting, forgot, forgotten**
8. marry (775)	**married, marries, marrying**
9. suffer (837)	**suffered, suffering, suffers**
10. maintain (942)	**maintained, maintaining, maintains**
11. hang (1083)	**hanged, hanging, hangings, hangs, hung**
12. quick (1193)	**quicker, quickest**
13. trend (1279)	**trended, trending, trends**
14. usual (1383)	-
15. afford (1486)	**afforded, affording, affords**
16. unique (1592)	-
17. faith (1684)	**faiths**
18. multiple (1730)	**multiples**
19. bother (1842)	**bothered, bothering, bothers**
20. rural (1961)	-

Table 3.4.3 20 Randomly Selected Words from the 2000 High Frequency Words

Lastly, Excel software was used to form the graphs based on the output files obtained from VP-Compleat and Frequency tools to draw conclusions.

CHAPTER IV

4. FINDINGS

To answer Research Question I, the first part in each section presents the overall profile of vocabulary coverage of the textbook by showing the proportions of vocabulary from different frequency levels. Cumulative proportions of their vocabulary profile are given, too. Each course book was examined separately by using *VP-Compleat tool* on Lextutor to get these results.

In the first part type-token ratios of two books were also examined. These ratios are presented within the output of the same tool, VP-Compleat. Type-token ratios were used to compare the recycling rate of the books, which revealed the answer to Research Question II.

To answer Research Question III, the second part of each section demonstrates the comparison of recurrence rates of 20 randomly-selected words in two ELT textbooks. These words were selected from the 2000 high frequency words lists (NGSL1 and NGSL 2). *Frequency tool* on Lextutor was applied to each book separately to identify how many times those 20 words were repeated in each book as a whole.

4.1. Vocabulary Profile of the Corpus-informed Textbook, Touchstone 3

Investigating lexical load of two textbooks was the aim of the Research Question I. The first textbook examined, Touchstone 3, claims to be corpus-informed. In this part, vocabulary profile analysis of Touchstone 3 with VP-Compleat tool is presented.

Frequency Level	TYPES (%)	TOKENS (%)	CUMULATIVE TYPES (%)
NGSL-1 WORDS	702 **(57.6%)**	2964 (80.4%	57.6 %
NGSL-2 WORDS	202 **(16.6%)**	287 (7.8%)	**74.2%**
NGSL-3 WORDS	95 (7.8%)	140 (3.8%)	82 %
NAWL	30 (2.4%)	38 (1.0%)	84.4 %
OFF-LIST	187 (15.3%)	256 (6.9%)	100
TOTAL	1217	3686	100

Table 4.1a VP-Compleat Output – Vocabulary Coverage of Touchstone 3

Table 8 shows the total number of types, tokens and their cumulative percentages according to frequency levels: NGSL-1, NGSL-2 and NGSL-3 in *Touchstone 3.* However, only the cumulative of the most common 2000 words (NGSL-1 and NGSL-2) is taken into consideration to answer Research Question I.

The total of common academic words and off-list words are also given in the same table. Academic words are the ones that are widely used in most of the academic texts. Words that do not appear in any of these frequency groups are classified as off-list in the table.

The statistics in Table 8 displays that in Touchstone 702 types (57.6%) out of 1217 are among the first 1000 high frequency words (NGSL-1). These words make up most of the vocabulary in the first textbook.

As can be seen in the table, the total of second 1000 most frequent words (NGSL-2) drops down to 202 (16.6%). There is a big difference between the coverage of NGSL-1 and NGSL-2 list words. When the coverage of NGSL2 words (16.6%) is added, the cumulative percentage of the word coverage becomes 74.2% in terms of types, which is taken as the basis in comparison of the vocabulary coverage of two books in the present study.

In Touchstone 3, the number decreases dramatically when it comes to the total of NGSL3 and academic words. As the third most common 1000 words are included in NGSL-3, its coverage is as low as 95types (7.8%) and the number of academic words is 30 (with coverage of 2.4%).

Another point to consider is the number and content of off-list words in the texts. In other words, they are used in the book, but these words are not among the most frequent 3000 words. Only a small number of infrequent words must be included in the textbook content. In Table 8, there are 187 (15.3%) off-list words that can be considered as low frequency lexical items. The number of off-list words is almost close to the number of NGSL-2 words.

Related Ratios and Indices	
Tokens	3686
Types	1217
Type-Token Ratio	**0.33**
Tokens per type	3.03
Lexical Density	0.57

Table 4.1b Type-token Ratio of Touchstone 3

In the output of VP-Compleat tool, type-token ratio (TTR) is displayed along with the lexical frequency figures. Table 9 shows the totals of types and tokens in the textbook content. Number of word types is divided by the total number of tokens to obtain TTR. The ratio allows us to identify the repetition rate of the text vocabulary, which is the answer to Research Question II.

TTR is displayed as a number that falls between 0 and 1.Whenthe total of types (1217) is divided by the total of tokens (3686), type-token ratio is 0.33 for Touchstone 3.

Result of type-token ratio is used to measure lexical diversity and recycling rate in a text. TTR was used to measure and compare the repetition of 20 randomly selected words in two textbooks in this research. A high TTR indicates increased lexical diversity whereas a low TTR shows that there are a great deal of repetitions.

4.2. Vocabulary Profile of the Traditionally-prepared Textbook, Headway Pre-intermediate

The aim of the first Research Question is to investigate the lexical load of two textbooks. The second textbook examined, Headway Pre-intermediate, is traditionally-prepared. Vocabulary profile analysis of Headway Pre-intermediate with VP-Compleat tool is presented in this part.

Frequency Level	TYPES (%)	TOKENS (%)	CUMULATIVE (%)
NGSL-1 WORDS	663 **(66.1%)**	3104 (86.6%)	66.1%
NGSL-2 WORDS	162 **(16.1%)**	225 (6.3%)	**82.2%**
NGSL-3 WORDS	73 (7.2%)	120 (3.3%)	89.4%
NAWL	12 (1.2%)	13 (0.4%)	90.6%
OFF-LIST	92 (9.1%)	122 (3.4%)	100
TOTAL	1003	3585	100

Table 4.2a VP-Compleat Output – Vocabulary Coverage of Headway Pre-intermediate

The statistics in Table 10 gives the total number of types, tokens and cumulative percentages based on frequency levels: NGSL-1, NGSL-2 and NGSL-3 in Headway Pre-intermediate Students' Book. Additionally, number of academic and off-list words is presented in the table.

The total number of types within NGSL-1 category is 663 (66.1%) out of 1003, which makes up most of the vocabulary items in the second book, Headway Pre-intermediate. The coverage of first 1000 most frequent words of English is higher here than the first textbook examined. The total of second 1000 most frequent words (NGSL-2) drops down to 162 (16.1%) as it can be seen in the table 10. The percentage of the types in NGSL-2 list is much lower when compared to NGSL-1. The cumulative percentage of the word coverage becomes 82.2% when the coverage of NGSL-2 words (16.1%) is added. While the proportion of NGSL-2 words

is lower in this book, the total of NGSL-1 and NGSL-2 (82.2%) cumulative is still higher than Touchstone 3 (74.2%). In the present research, total percentage of 2000 high frequency words (82.2%) is taken into consideration to compare the lexical load of two books.

In the second textbook, Headway Pre-intermediate, the number of types in NGSL-3 list is 73 (with coverage of 7.2%) and the number of academic words is 12 (with coverage of 1.2%). When it comes to off-list category, there are 92 types (9.1%) in Headway Pre-intermediate, which is lower than the low frequency words in Touchstone 3.

Related Ratios and Indices	
Tokens	3585
Types	1003
Type-Token Ratio	**0.28**
Tokens per typ	3.57
Lexical Density	0.50

Table 4.2b Type-token Ratio of Headway Pre-int.

As can be seen from Table 11 the number of types is 1003 and the total of tokens is 3585 in the second textbook (Headway). In order to answer Research Question II and see the recycling rate of the vocabulary items, types are divided by tokens to get type-token ratio. The ratio is 0.28 in Headway Pre-intermediate, which is lower than Touchstone 3.

4.3. Frequencies of Randomly Selected NGSL Words in two Textbooks

A set of 20 randomly selected words from 2000 corpus frequency band (NGSL)was examined along with their inflected forms in order to find their frequency values and to discover to what extent their frequency of occurrence reflect corpus-based frequency ordering. To increase the validity of the results, the words were chosen from both NGSL-1 and NGSL-2.

NGSL words & their corpus frequencies	Inflected forms	Their frequency in Headway	Their frequency in Touchstone
1. make (48)	**made, makes, making**	**5**	**10**
2. home (161)	**homed, homes, homing**	**5**	**2**
3. buy (242)	**bought, buying, buys**	**0**	**3**
4. reach (380)	**reached, reaches, reaching**	**1**	**2**
5. shop (437)	**shopped, shopping, shops**	**1**	**2**
6. listen (573)	**listened, listening, listens**	**0**	**0**
7. forget (651)	**forgets, forgetting, forgot, forgotten**	**1**	**1**
8. marry (775)	**married, marries, marrying**	**0**	**1**
9. suffer (837)	**suffered, suffering, suffers**	**0**	**0**
10. maintain (942)	**maintained, maintaining, maintains**	**0**	**0**
11. hang (1083)	**hanged, hanging, hangings, hangs, hung**	**0**	**0**
12. quick (1193)	**quicker, quickest**	**0**	**0**
13. trend (1279)	**trended, trending, trends**	**0**	**0**
14. usual (1383)	**-**	**1**	**0**
15. afford (1486)	**afforded, affording, affords**	**0**	**0**
16. unique (1592)	**-**	**0**	**0**
17. faith (1684)	**faiths**	**0**	**0**
18.multiple (1730)	**multiples**	**0**	**0**
19. bother (1842)	**bothered, bothering, bothers**	**0**	**0**
20. rural (1961)	**-**	**0**	**0**

Table 4.3 Frequency of Randomly Selected Words from 2000 High Frequency Word List

Table 12 shows the frequencies of randomly selected words with their inflected forms in the lists NGSL-1 and NGSL-2 and how many times they are repeated in total in each textbook.

Coverage and repetition of the words from NGSL-1 group is higher than NGSL-2 group in both books and the recycling rate of the words drops as their list standings decreases. For instance, out of 10 randomly selected NGSL-1 words, 5 of them are covered in Headway Pre-intermediate and 7 of them are represented in Touchstone 3. Also, out of 10 randomly selected NGSL-2 words, 1 one of them is covered in Headway Pre-intermediate and none of them are represented in Touchtone 3.

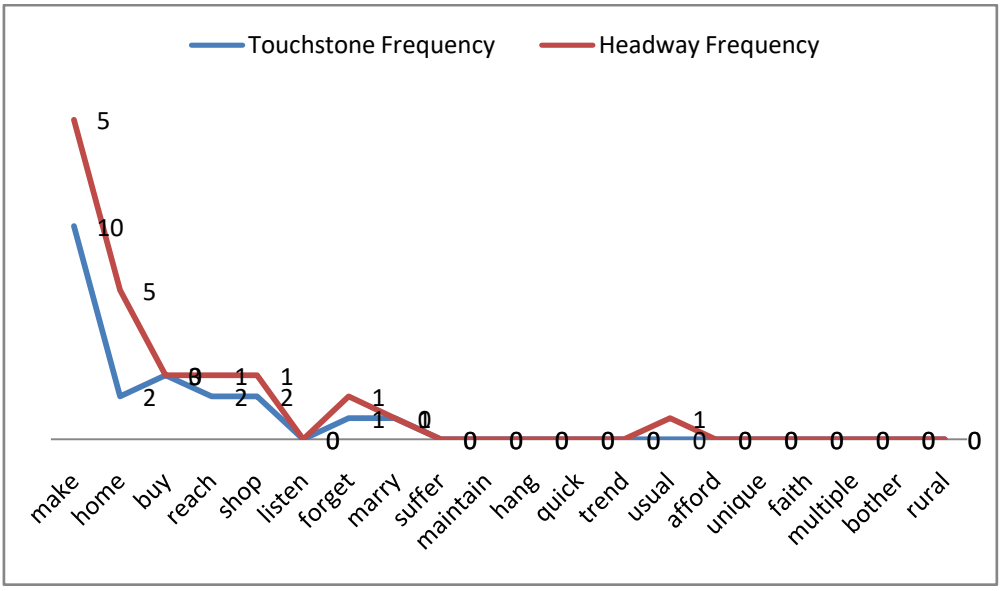

Figure 4.3 Frequencies of 20 NGSL Words in Two Textbooks

As may be observed in Figure 2, recycling rate of 20 random words from 2000 most frequent words drops corresponding to their standings further down on the NGSL list, which means that the item is less frequent.

Among randomly chosen 10 NGSL-1 words, almost half of them are covered in both textbooks and coverage of 10 random NGSL-2 words is not high enough to mention. Most of the randomly selected NGSL-2 words are not included in both Headway Pre-intermediate and Touchstone 3.

CHAPTER V

5. RESULTS, DISCUSSION AND RECOMMENDATIONS

5.1. Results and Discussion

5.1.1. Vocabulary Load of Two Textbooks

This corpus-based study investigated and compared load of high frequency lexical items in two EFL textbooks. To conduct the study, course books' content were analyzed by using the most frequent 2000 words of NGSL and NAWL as reference lists. Vocabulary Profiler Compleat tool on Lextutor website was used as the instrument for the analysis of the lexical profiles. The first research question addressed in this study is:

- What is the coverage level of the most frequent 2,000 words of NGSL in a corpus-informed book *(Touchstone Level 3 Student's Book)* and a conventional course book *(Headway Pre-intermediate Level Student's Book)?*

Frequency Level	Touchstone Types (%)	TOUCHSTONE CUMULATIVE (%)	Headway Types (%)	HEADWAY CUMULATIVE (%)
NGSL-1	702 **(57.6%)**	57.6 %	663 **(66.1%)**	66.1%
NGSL-2	202 **(16.6%)**	**74.2%**	162 **(16.1%)**	**82.2%**
NGSL-3	95 (7.8%)	82 %	73 (7.2%)	89.4%
NAWL	30 **(2.4%)**	84.4 %	12 **(1.2%)**	90.6%
OFF-LIST	187 **(15.3%)**	100	92 **(9.1%)**	100
TOTAL	1217	100	1003	100

Table 5.1.1 Comparison of Vocabulary Load in Two Textbooks

The analysis of vocabulary in the corpus-informed textbook, Touchstone 3, revealed that 1000 most frequent words (NGSL-1) occur with high frequency in the course book. More than half of the lexical items (57.6%) in the corpus-informed book, Touchstone 3, belong to NGSL-1 word list (See table 13). While the percentage of types in NGSL-1 list is 57.6% in the corpus-informed course book (Touchstone 3), it increases to 66.1% in the traditionally-prepared textbook (Headway Pre-intermediate).

The proportion of second most frequent 1000 words (NGSL-2) is 16.6% in Touchstone 3. The percentage decreases for NGSL-2 words in this corpus-informed textbook as they are used less frequently in most English contexts. Also because the textbook examined is at pre-intermediate level, only a reasonable number of NGSL-2 list words may be included in its content.

With the addition of NGSL-2 words (16.6%), vocabulary coverage cumulative for types rises to 74.2% in the corpus-informed book (See table 13). Two textbooks were compared based on this percentage as it gives the total coverage of most common 2000 thousand words of English. Using the frequency information as a basis in the comparison of textbooks can provide crucial information on their authenticity and efficiency. Schmitt and Schmitt (2012) maintain that the knowledge high frequency words is required in the first place in order to understand English to a large extend. Therefore, textbooks giving sufficient coverage of these frequent words may be considered more authentic.

The percentage showing the vocabulary load of 2000 most frequent words (74.2%) in Touchstone 3 is below the necessary level for understanding the texts well enough. Hirsch (2003) states that for good comprehension of a text 95% of the words should be understandable to the learner. Based on this fact, it can be said that a learner must also have the knowledge of NGSL-3 and academic words to some extent. That is, learners must also have the knowledge of less frequent words of English (words that are not among 2000 high frequency vocabulary). However, when the number of words that fall within NGSL-3 and Academic Word categories is high, it can cause difficulty in comprehension for learners.

In this textbook, Touchstone 3, the number of types within NGSL-3 list is 95 (with coverage of 7.8%) and there are 30 academic words (with coverage of 2.4%), which would not cause difficulty in understanding the texts. Also, because the number of words in NAWL list is quite low, they do not have to be removed from the book (See Table 14).

Another point to consider is the number and content of off-list words in the texts. The number of these low frequency words must be small enough to be ignored. In Table 13, there are 187 (15.3%) off-list words that are classified as low frequency lexical items, which is almost close to the number of NGSL-2 words and it is above the desired level for good comprehension. Therefore, the number of off-list words is expected to be more limited, especially for a textbook at Pre-intermediate level.

In Touchstone 3, the cumulative of the types apart from NGSL-1 and NGSL-2 lists is 25.8%. That is, approximately 26% of the types in the book are not among 2000 high frequency words of English.

In most of the previous research, studies analyzing the vocabulary profile of textbooks demonstrated very similar results in terms of the coverage of words from different frequency levels. It was identified that the higher the frequency of a word, the higher their coverage is in the textbook.

Similarly, Ünlü (2012) examined a traditional EFL textbook (Spot-On, Grade-8) in terms of its loading, distribution, and repetition patterns of the 2000 most frequent words of English. She concluded that most of the types (67.6%) included 1000 high frequency words. The percentage of second most frequent 1000 words was 15% and also the textbook included 6% academic words and 11% off-list words in her study.

Alhudithi (2017), likewise, carried out a corpus-based analysis of the vocabulary input in 12 textbooks used in Saudi Arabia. The researcher concluded that the highest number of lexical items covered in the book were from 1000 most frequent English words. The percentage of words in the other frequency levels were much lower than K-1 group, which shows that infrequent words were limited in the course books analyzed.

The statistics in Table 13 show that in Headway Pre-intermediate, the percentage of the types that fall within NGSL-1 group is 66.1%, which is higher than Touchstone 3 (57.6%). Since NGSL-1 includes the most frequently used 1000 words, it can be expected that this group of words are used more commonly than other frequency groups in the textbooks. That is, in Headway Pre-intermediate Students' Book, the first 1000 high-frequency words make up more than half of the lexical items similar to Touchstone 3.

Counterintuitively, as a conventional course book, Headway Pre-intermediate gives a better coverage of the most frequent vocabulary of English (NGSL-1) although it is a traditionally-prepared textbook.

The percentage of the second most frequent 1000 words (NGSL-2) is 16.1% in Headway Pre-intermediate (See table 13). The proportion of NGSL-2 group is much lower when compared to NGSL-1 list coverage (66.1%) as NGSL-2 group of words are not used as commonly as NGSL-1 list. Also because the inclusion of more of these words is more appropriate for textbooks at higher levels than pre-intermediate.

The proportion of NGSL-2 types in Headway Pre-intermediate (16.1%) is not much different from the proportion of NGSL-2 words in Touchstone 3 (16.6%). In fact, their percentage is almost the same in both textbooks.

Similar to Touchstone 3, the lexical coverage of NGSL-2 types decreases in Headway Pre-intermediate as they are less frequently used words of English. Also because in a pre-intermediate level textbook, only a limited number of second most frequent 1000 words (NGSL-2) can be covered.

In Headway Pre-intermediate, vocabulary coverage reaches 82.2% with the addition of NGSL-2 words (16.1%), which is closer to the desired level for a good comprehension of the texts than Touchstone 3.

For good comprehension of the text over 90% of the words should be understandable to the learners. Considering this fact, it can be assumed that learners must have knowledge of NGSL-3 and academic words to some extent. However, when the number of words that fall within NGSL-3 and academic word categories is not limited, learners may have difficulty in comprehension.

In the second textbook, Headway Pre-intermediate, the number of NGSL-3 words is 73 (with a coverage of 7.2%) and the number of academic words is 12 (with a coverage of 1.2%), which may be considered as appropriate for learners at this level. Furthermore, including a limited number of academic words at the pre-intermediate level could be a good preparation for the next level.

When it comes to off-list category, there are 92 words (9.1%) in Headway Pre-intermediate, which is lower than the proportion low frequency words in Touchstone 3. However, it can be still smaller for a good comprehension of the texts.

In Headway Pre-intermediate, the cumulative of the types apart from NGSL-1 and NGSL-2 lists is 17.8%. That is, approximately 18% of the types in the book are not among 2000 high frequency words of English.

Findings of vocabulary profile analysis of Headway Pre-intermediate are in line with the research by Milton and Alsaif (2012). Their study also found that most of the vocabulary input included in the textbook was from 2000 most frequent words of English. However, its vocabulary coverage rate was a bit higher in the study of Milton and Alsaif (2012) since it depends on the level of the book examined. In a similar study, Mukunkan and Aziz (2009) observed that most of the Malaysian English Language Textbooks include the most frequent 2000 words of English in their content to a large extend. Thus, it can be assumed that learners are expected to be familiar with high-frequency words of English in different contexts and parts of the world.

To summarize, vocabulary content of a corpus informed textbook (Touchstone 3) and a traditionally-prepared textbook (Headway Pre-intermediate) were analyzed in terms of their lexical frequency profile to answer the first Research Question in the present study. Lexical content of the course books were compared against 2000 high frequency vocabulary of English by using NGSL-1 and NGSL-2 base word lists as reference.

Frequency analysis of vocabulary load in texts provides us with invaluable information as to whether the content suits learners' lexical needs at given proficiency level. It also gives us the opportunity to check the efficiency of the book as regards its lexical content as it allows us to see whether there is a systematic and balanced vocabulary selection behind. According to Horst (2013), frequency-informed decisions must be the basis in vocabulary teaching and selection.

Furthermore, vocabulary load analysis provides crucial implications for checking the variety and complexity of words. Using corpus software to analyze lexical frequency is a practical and effective way of understanding how authentic the content is since high frequency is associated with more authentic English. It is argued by Reppen (2012) that "if a feature is very common, and is used by fluent native speakers of English, then we should certainly teach

that feature to learners" (p. 14). Nation (2001) also emphasizes that learning 2000 high-frequency words can be sufficient to comprehend most of the different types of texts in English.

Along with many studies, it has been demonstrated in this research that percentage of lexical coverage in the textbooks increases for the words with higher frequency. Alhudithi (2017) claims that 2000 high frequency words appear more commonly in English contexts and their lexical coverage make up the biggest amount of the textbooks' content.

In this study, although vocabulary content of Touchstone 3 and Headway Pre-intermediate textbooks showed similarity in terms of presenting 2000 most frequent words in most of their content, they differed in the proportion of the high frequency vocabulary covered. For the comparison, Headway was selected as a traditionally-prepared textbook and Touchstone 3 was chosen as a corpus-informed book.

Despite being a conventional textbook, Headway Pre-intermediate gives a higher coverage of 2000 high frequency words. To be more precise, the cumulative percentage of types in NGSL-1 and NGSL-2 lists is 74.2% in the corpus-informed course book (Touchstone 3) whereas it increases to 82.2% in the traditionally-prepared textbook (Headway Pre-intermediate). Since there is no study found comparing the lexical content of a conventional textbook with a corpus-informed textbook, results could not have been compared with a reference research in this regard.

Possible reasons for the difference in their coverage of high frequency words between two textbooks could be:

- Headway textbook series have been available for longer time and used more commonly than Touchstone series worldwide. Therefore, more feedback on its content may have been collected to update and adapt the book accordingly.
- Another reason could be that 5th edition of Headway was analyzed and it is a more recently published book than Touchstone 3. Thus, vocabulary content in the reading spots of the book might have been arranged according to more up-to-date topics.
- Apart from Touchstone series, there are hardly any books that claim to be corpus-informed worldwide. As the use of corpus for textbook preparation is a very rare practice among publishers, vocabulary content of them may not have been designed in a way that reflects authentic English sufficiently.

On the other hand, there is a big difference in the proportion of NGSL-1 and NGSL-2 words included in ELT textbooks depending on the level of the book. The proportion of lexical items from NGSL-2 list is much lower in the textbooks at elementary or pre-intermediate levels; however, its percentage increases as the level of the book gets higher (textbooks at intermediate or higher level include bigger amount of vocabulary).

Coverage of NGSL-3, academic and off-list words in a textbook at pre-intermediate level can be considered as its content of low frequency words. With that being given, approximately 26% of the types in Touchstone 3 and almost 18% of the types in Headway Pre-intermediate are among the low frequency vocabulary. Presenting fewer less frequent words may be viewed as strength of Headway in this regard. As the percentage of infrequent words is not low enough, the results reveal that vocabulary selection process in both books are not very systematic and the coverage of 2000 high frequency words may be increased in both textbooks.

5.1.2. Recycling Rate of High Frequency Words in Two Textbooks

This corpus-based study aimed to discover and compare recycling rate of 2000 most frequent words of English in a corpus-based and traditionally-prepared textbook with Research Question II. In order to obtain the recycling rate of vocabulary in two textbooks Vocabulary Profiler Compleat and Frequency tools on Lextutor website was utilized as the instrument. The second Research Question addressed in this study is:

2. What is the recurrence rate of high frequency words in a corpus-informed and a traditionally prepared course book?

Firstly, the output of VP-Compleat tool was used for obtaining recycling rate of vocabulary in textbooks as it displays the type-token ratio (TTR) for each textbook. The number of tokens and types in the textbook content was given in the output and the type-token ratio was calculated automatically by the tool. Total number of all the different words in a text (types) is divided by the total number of all the words (tokens) to obtain its type-token ratio. TTR is a number that falls within 0 and 1. It allows the researcher to see to what extent the words in a text are repeated. A low TTR shows that there is a great deal of repetitions whereas a high TTR indicates words are not recycled sufficiently.

Recycling is an important aspect of vocabulary learning. Folse (2004) states that multiple exposures are necessary for moving the new words from short term memory to long

term memory. Cobb (1999) supports this view stating that learners must encounter new words in a variety of contexts in order to recall and use them flexibly.

Related Ratios and Indices	Touchstone 3	Headway Pre-int.
Types	1217	1003
Tokens	3686	3585
Type-Token Ratio	**0.33**	**0.28**
Tokens per type	3.03	3.57
Lexical Density	0.57	0.50

Table 5.1.2a Comparison of Recycling Rates of Two Textbooks

When the total of types (1217) is divided by the total of tokens (3686), type-token ratio is 0.33 for the corpus-informed textbook, Touchstone 3. As can be seen from Table 5.1.2a the number of types is 1003 and the total of tokens is 3585 in the traditionally-prepared textbook, Headway Pre-intermediate. In order to answer Research Question II and see recycling rate of the vocabulary items, types are divided by tokens to get type-token ratio. The ratio is 0.28 in the traditionally-prepared textbook, Headway Pre-intermediate.

When the recycling rate is lower, it indicates higher repetition of the vocabulary. Type-token ratios of both textbooks are below 0.5 and close to 0. In other words, repetition of vocabulary in both course books is not very low. However, despite being a traditionally-prepared textbook, TTR of the vocabulary in Headway (0.28) is higher than TTR (0.33) of the words in Touchstone, which indicates that repetition of lexical items in Touchstone 3 is lower.

Counterintuitively, as a conventional course book, Headway Pre-intermediate gives a higher recycling rate than the corpus-informed textbook (Touchstone 3). Possible reason behind this result could be that the use of Headway for teaching English is more widespread globally when compared to Touchstone 3, which can lead to presenting better content after years of updating and improvement. The fact that there are hardly any books globally that claim to be corpus-informed may be another reason behind the lack of sufficient reflection of high frequency words in Touchstone 3. Since the use of corpus for textbook preparation is not a

common practice among publishers, lexical content of them might not have been designed in a way that reflects authentic English sufficiently.

Studies on recycling of vocabulary input in textbooks identified that learners are not given enough opportunities to repeat newly introduced words. For instance, Tomlinson (2008) and Dat (2008) maintain that repetition of lexical items is overlooked in spite of the amount of vocabulary activities provided in textbooks.

Nation (2001) emphasizes the importance of recycling by stating that repetition of words helps learners to deepen their knowledge and understanding of them along with increasing chance of learning the vocabulary: "repetition is essential for vocabulary learning because there is so much to know about each word that one meeting with it is not sufficient to gain this information, and because vocabulary items must not only be known, they must be known well so that they can be fluently accessed." (p. 76)

Alfotais (2012) also considers recycling crucial by saying that "the ability to better memorize a word comes hand in hand with the number of times this word has been recycled" (p. 20).

After checking the type-token ratios of both textbooks to examine the recurrence of the words, frequency values of all the types in the books were analyzed with the help of *Frequency* tool on Lextutor. The output of this tool gives all the types of in the textbook and how many times each type is repeated in the book.

Frequency Values	**Cumulative Coverage Touchstone 3**	**Cumulative Coverage Headway Pre-int.**
10 and above	4.3%	5.3%
5-10 occurrences	4.5%	6%
3-5 occurrences	10.8%	10.5%
2 occurrences	17.2%	16.1%
1 occurrence	62.8%	61.8%

Table 5.1.2b Comparison of Each Frequency Coverage in Two Textbooks

As can be seen in Table 15, for the corpus-informed textbook, Touchstone 3, result of the frequency values obtained from Frequency tool is as follows.

Out of 1217 types:

- 53 types (4.3%) are repeated 10 and more times.
- 57 types (4.5%) are recycled between 5-10 times.
- 132 types (10.8%) are repeated between 3-5 times.
- 210 types (17.2%) are recycled twice.
- 765 types (62.8%) have just 1 frequency. They are not repeated in the reading spots.

In conclusion, total proportion of types with 5 or more frequencies is 8.8% in the corpus-informed textbook.

The result of the frequency values obtained from Frequency tool for the conventional textbook, Headway Pre-intermediate, is as follows. Out of 1003 types:

- 54 types (5.3%) are repeated 10 and more times.
- 61 types (6%) are recycled between 5-10 times.
- 106 types (10.5%) are repeated between 3-5 times.
- 162 types (16.1%) are recycled twice.
- 620 types (61.8%) have only 1 frequency. They are not repeated in reading spots.

As a result, the total proportion of types with 5 or more frequencies is **11.3 %** in the conventional textbook.

It may be concluded from the frequency values in two textbooks that recycling of types in the traditionally-prepared textbook is slightly better than the corpus-informed textbook. To be more precise, in Touchstone, **62.8%** of the types are not repeated in the reading spots. The proportion of types with 1 frequency decreases to **61.8%** for Headway. That is, more than half of the types in both course books are not recycled. Nagy et al. (1987) states that accuracy cannot be over 5% in a multiple-choice question test if a word occurs with 1 frequency in a text. This result also suggests that both textbooks need improvement in terms of the recurrence rate of lexical items.

Recycling rate of the words can also be compared by considering the percentage of words with 5 or more frequencies in two textbooks because most exposure frequency studies suggest 5 repetitions as the lowest number required (Hwang &Nation, 1995; Saragi et al., 1978; Thornbury, 2002; Webb, 2007). For instance, in Touchstone **8.8%** of the types are recycled 5 or more times while in Headway the percentage of types with the same frequency is **11.3%**.

Nation (1990) claims that 5-16 exposures are necessary in order to learn new words while Özdem (2010) argues that even 5 repetitions are not enough to use the words effectively.

Proportion of the words recycled less than 5 times is 80% in Touchstone and 77.9% in Headway (see Table 15). Therefore, it may be asserted that learners are not presented with sufficient exposure opportunities in both textbooks to be able to acquire the words effectively.

In a study of vocabulary load and recycling in a conventional textbook, Ünlü (2012) demonstrated that 60% of the types were not repeated in the book and 9% of the lexical items were recycled 5 or more times. With results similar to this study the researcher concluded that the book contained "inadequate number of recycled words with insufficient frequency values" (p.82).

In another similar study Alhudithi (2017) observed that lexical items with 5 or more occurrences covered smallest percentage whereas words with fewer than 5 occurrences made up the largest proportion of the textbook. His research also indicated that "when more new words are introduced, there are fewer opportunities for repetition" (p.55).

The results of this study also demonstrated that lack of recycling may cause including more infrequent words in a textbook because learners would have a lot more words to acquire when the vocabulary is not repeatedly used. As a result, coverage of high frequency words in a context would decrease.

5.1.3. Frequencies of Randomly Selected Words in the Course Books

The aim of the third Research Question was exploring the repetition rate of 20 randomly chosen words in a corpus-informed and traditionally-prepared textbook and discovering to what extent their frequency of occurrence reflects corpus-based frequency ordering. The words were chosen from both NGSL-1 and NGSL-2 lists to increase the validity of the results.

The third question addressed in this study is:

3. What is the recycling rate of the randomly selected words in two textbooks?

3.1. To what extent does the frequency of occurrence of randomly selected words in two books reflect corpus-based frequency ordering?

NGSL words & their corpus frequencies	Their frequency in Headway Pre-int.	Their frequency in Touchstone 3
1. make (48)	**5**	**10**
2. home (161)	**5**	**2**
3. buy (242)	**0**	**3**
4. reach (380)	**1**	**2**
5. shop (437)	**1**	**2**
6. listen (573)	**0**	**0**
7. forget (651)	**1**	**1**
8. marry (775)	**0**	**1**
9. suffer (837)	**0**	**0**
10. maintain (942)	**0**	**0**
11. hang (1083)	**0**	**0**
12. quick (1193)	**0**	**0**
13. trend (1279)	**0**	**0**
14. usual (1383)	**1**	**0**
15. afford (1486)	**0**	**0**
16. unique (1592)	**0**	**0**
17. faith (1684)	**0**	**0**
18. multiple (1730)	**0**	**0**
19. bother (1842)	**0**	**0**
20. rural (1961)	**0**	**0**

Table 5.1.3a 20 Randomly Selected Words from the 2000 High Frequency Words

Table 16 shows the frequencies of randomly selected words with their inflected forms in NGSL-1 and NGSL-2 lists and their total repetition rate in each textbook.

In the traditionally-prepared textbook, Headway Pre-intermediate, 5 words are covered with different frequencies out of 10 NGSL-1 words. These are: make home, reach shop and forget. Only 1 word is included out of 10 NGSL-2 words. It is the word ‘usual’.

In total, out of 20 randomly selected words 14 of them are not included in Headway Pre-intermediate. It contains 6 random words with different frequencies and most of these are covered only once in this textbook. For instance, the words ‘reach’, ‘shop’, ‘forget’ and ‘usual’ occur once in the reading spots of the book. Also, the words ‘make’ and ‘home’ are repeated 5 times in total.

In the corpus-informed textbook Touchstone 3, 7 words are included with different frequencies out of 10 NGSL-1 words. These are: make, home, buy, reach, shop, forget and marry. None of the randomly selected 10 NGSL-2 words are covered in this book.

In total, out of 20 randomly selected words 13 of them are not included in Touchstone 3. It contains 7 random words with different frequencies. The words 'forget' and 'marry' are covered only once in the reading spots of the book. The rest of the words such as 'make', 'home', 'buy', 'reach' and 'shop' are repeated more than once.

Repetition and coverage and of 10 words from NGSL-2 group is lower than 10 words from NGSL-1 group in both books and the recycling rate of the words drops as their list standings decrease.

As can be seen in Table 16, the list standing of the word '*make'* is 48 on NGSL-1 and it is the first word picked. The word *'make'* is repeated 5 times in Headway Pre-intermediate and 10 times in Touchstone 3. Moreover, *'make'* is the word with highest frequency value in both textbooks.

The other word *'usual'* stands as the 1383th on NGSL-2. It occurs only once in Headway Pre-intermediate and it is not covered in Touchstone 3.

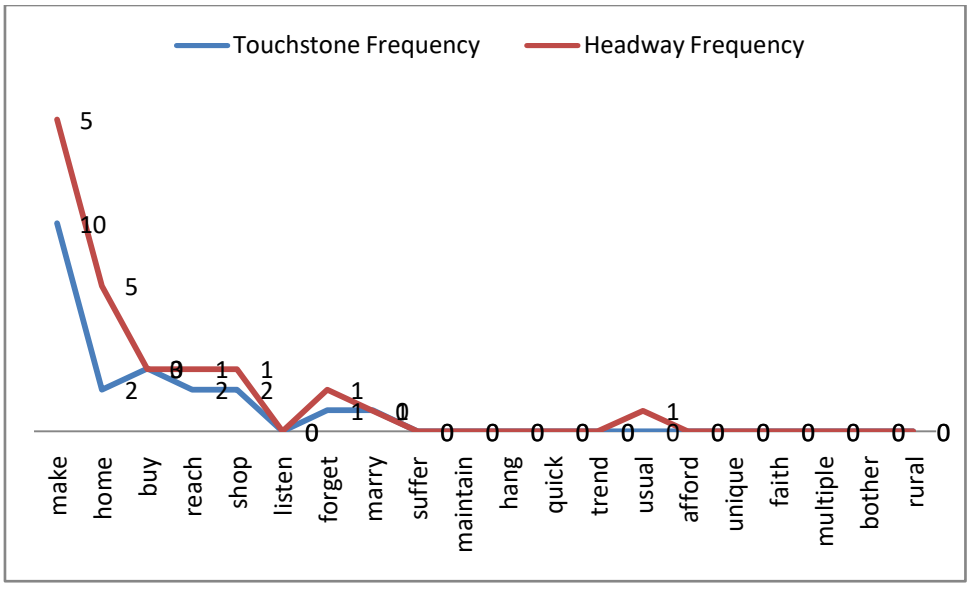

Figure 5.1.3b Repetition of Random NGSL Words in Two Textbooks

As may be observed in Figure 6, repetition rate of 20 random words from 2000 most frequent words drops as their standings in NGSL lists decrease.

10 words were chosen from the list NGSL-1, almost half of them are represented in both textbooks and the coverage of 10 random words from NGSL-2 list in the textbooks is too low to mention. Most of the randomly selected NGSL-2 words are not covered in both Headway Pre-intermediate and Touchstone 3.

In the corpus-informed book, Touchstone 3, approximately half of the randomly selected 20 words with their inflected words are recycled with varying values. In other words, their frequency is higher than 1 or 2 despite the fact that only reading spots were extracted from the whole book to discover the recycling of the randomly chosen words.

On the other hand, in the traditionally-prepared textbook, Headway Pre-intermediate, less than half of the randomly selected 20 words are covered just once within all the reading spots of the book and they are not recycled sufficiently. Most of the randomly chosen words with their inflected forms are covered only once in the traditionally-prepared textbook.

Third Research Question aimed to identify the repetition rate of the randomly selected words in a corpus-informed and traditionally-prepared textbook. Considering the results above, it may be concluded that there is a lack of systematic and sufficient recycling for the randomly selected 20 words in the traditionally-prepared textbook, Headway Pre-intermediate.

Nation (1990) states that new words must be recycled between 5-16 times in different contexts to provide learners with enough exposure for learning. It can be expected that some of the randomly chosen words may not be repeated more than 3 times in a limited context of reading spots of the course book, Headway Pre-intermediate. However, covering them only once doesn't give the learners enough opportunity for the required exposure frequency.

Corpus-informed textbook, Touchstone 3, seems to provide the learners with better opportunity in terms of recycling of 20 randomly chosen words. The number of words which are repeated more than once is higher in Touchstone 3 than Headway Pre-intermediate. However, repetition of the words may still be increased since recycling vocabulary more than 3 times is suggested in most studies. It is widely accepted that higher frequency values makes the recalling of words easier.

Second point examined in the Research Question III was to what extent randomly selected words' frequency of occurrence reflects corpus-based frequency ordering in a corpus-informed and traditionally-prepared textbook.

Widely used words of English are expected to have higher frequency values in both textbooks. Furthermore, their frequency of occurrence is expected to decrease corresponding to their standings on NGSL list.

As can be observed in Figure 6, repetition rate of 20 words from 2000 most frequent words decreases in line with their list standings on NGSL. That is, recycling of randomly selected words from 1000 high frequency are higher and most of the randomly chosen words from NGSL-2 do not appear in the textbooks. Moreover, the tendency of the random words' frequency drop according to their NGSL list standings is similar in both Touchstone 3 and Headway Pre-intermediate, which indicates that frequency values of randomly selected words in both texts are parallel to NGSL frequency of occurrence to a great extent.

5.2. Recommendations

The importance of lexis is unquestionable as an essential component of language learning and ELT textbooks' content. Despite this, different aspects of vocabulary input that learners and course books need have not been investigated thoroughly. Lack of research may have been caused by limited opportunities researchers had in the past to deal with vast number of words as well as other reasons.

Corpus and corpus-based studies made it possible to further analyze words in general and lexical content of language teaching materials and textbooks. These studies may be used both for evaluating the existing content and selecting materials to improve it. Corpus-informed materials such as grammar books, dictionaries and course books are being used by teachers and learners even if they do not always recognize that those materials are corpus-based (Chambers, 2019). Thus, it can be said that corpus has already became a part of the language learning and teaching.

However, researchers maintain that there are still many English teachers and textbook writers who are not familiar with the basic corpus linguistics terms and corpus use in language teaching. For instance, Şimşek (2020) concluded from her qualitative study with pre-service English teachers that many of the participants are not equipped with corpus literacy skills and the number of corpus-related courses is very scarce in teacher education (Callies, 2019; Frankenberg&Garcia, 2012). Corpus literacy is defined as "the ability to use the

technology of corpus linguistics (CL) to investigate language and enhance the language development of students" (Heather& Helt 2012, p. 417).

For this reason, this study highly suggests informing English teachers, course books writers and publishers about the existence of corpus and its use in the first place. In order to do this, corpus literacy courses could be integrated into education of language teachers. Textbook writers can be trained on the use of corpus, too. Once the teachers acquire corpus literacy skills, they can compare textbook content with the most frequent words or structures of English to identify its extend of authenticity and use the course books more flexibly. That is, professionals in the field of language teaching can evaluate textbooks more critically with better pedagogic and language awareness.

As a result, teachers may take more responsibility in material development and adaptation of the textbooks when their language awareness is raised.

It is essential because learners mostly rely on textbook content in their vocabulary learning process. Teachers make use of ELT textbooks to provide students with most of the input they get. There are studies which revealed that textbook writers follow a topic-based approach for content and vocabulary selection or sometimes they just decide on the lexical content intuitively. To illustrate, lexical content of six commonly-used course books were analyzed by Reda (2003) and their vocabulary content seemed to be topic-based. Similarly, the present research concluded that in textbooks coverage and recycling of most frequent words is not high enough, even in the corpus-informed course book.

Results of recycling rates in both course books showed that vocabulary repetition within the textbook require systematic analysis and methods to improve. In this regard, lexical frequency profiles of the words may be utilized to decide on which words should be selected and which ones should be recycled more often than the others.

This study analyzed the lexical content of two textbooks in comparison with high-frequency wordlists by using a corpus-based method. As the frequency of a word helps us to understand how often it is used by native speakers, it may be considered as an indicator of usefulness of the lexis (Nation and Waring, 1997; Richards, 1974). Having good command of the most frequent words would help learners comprehend most of the written or spoken contexts of English and use the language more with flexibility. Moreover, students can guess the meaning of some low frequency words when they know the most essential words in a context.

Based on the vocabulary profile results of the textbooks, this research shows the necessity of covering sufficient number of high-frequency words as a priority. However, course book writers and teachers must be informed that the use of most frequent words of English as reference is essential in presenting authentic content in materials and textbooks and investigating them more critically to check what weaknesses they have for teaching vocabulary.

With the help of vocabulary profiler tools, it is quite practical to use high-frequency lists as a reference to check the coverage of frequent and infrequent words in a text. Therefore, these tools can be used to avoid low frequency words which can be described as less useful part of the lexical content. This study also points out that it is crucial to use vocabulary coverage results as they reveal which frequent words are included and excluded in a context. This result can be used to cover more of high-frequency words and to remove some of the low frequency words in the content. It may be done in either in the design process of textbooks or during the selection of materials to supplement the course books. By profiling the content of the textbooks we can decide on the parts that are really indispensable for students' needs and remove the others instead of following the textbook content page by page. Additionally, selected parts of the course book could be supplemented with materials that reflect the most useful vocabulary and structures of English.

Nation (2000) and Ellis (2002) emphasized the impact of frequency and repetition on vocabulary learning. In addition to learning words adequately, learners can recognize and respond to these words in a shorter time when they occur with high frequency in a context. Therefore, frequency is perceived as an important factor contributing to the effective learning of vocabulary. Researchers note that words that occur less frequently in English is thought to be more difficult by learners (Chen & Truscott, 2010; Koirala, 2015). Furthermore, lower frequency of occurrence decreased the chance of delayed retention of the words.

High frequency is important also because learners may learn words implicitly by exposure. However, principled instruction of vocabulary started to be recommended too in language teaching field. For this purpose, making use of high frequency words may prove useful. Teachers may try to provide sufficient vocabulary repetition for learners with in-class activities during the lessons by taking these lists as reference although most of the textbooks are problematic in terms of recycling of the words. Since graded-readers are written based on the most frequent vocabulary, using them could be an effective and practical way of

presenting high-frequency words to the students and supplementing course books. These readers are useful both in terms of their coverage of useful vocabulary and high recycling rate. They would give the learners a chance to read extensively and learn the vocabulary implicitly. Also, follow-up activities of the readers could be a way of increasing repetition of the words and teaching explicitly.

Considering the points above, improving the vocabulary content of textbooks and supplementing them with materials that present more useful and authentic words is a must. Increasing the authenticity and selecting more useful vocabulary content would definitely be more beneficial and motivating for learners. Corpus-based materials can be reliable sources in this regard. More importantly, this process requires teachers to be better researchers, active content developers instead of just being passive consumers of the content provided for teaching. Schön (1987) emphasizes the same point by stating that teachers are expected to be more actively involved in teaching process as reflective practitioners with critical awareness. Therefore, education provided for teachers must encourage them to be more productive, too.

In future studies, both the number of books and sections investigated within the books may be considered to be extended in order to find distribution of high frequency words along with their coverage in the textbook.

Finally, the same type of comparative vocabulary profile analysis of textbooks can carried out by using different high-frequency word lists as the basis or different textbooks.

REFERENCES

Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon*. John Wiley & Sons.

Alfotais, A. (2012). *Investigating textbooks' input as a possible factor contributing to vocabulary knowledge failure among Saudi EFL learners at Taif University* (Unpublished master's thesis). University of Essex.

Alhudithi, E. (2017). *A Corpus-Based Analysis of English Vocabulary Input Provided in K-12 Textbooks Used in Saudi Arabia* [Doctoral dissertation, Colorado State University].

Alsaif, A., & Milton, J. (2012). Vocabulary input from school textbooks as a potential contributor to the small vocabulary uptake gained by English as a foreign language learners in Saudi Arabia. *The Language Learning Journal*, *40*(1), 21-33.

Anderson, J. R. 2010. (1980). *Cognitive psychology and its implications*. (7th ed.). Worth Publishers.

Arnaud, P. J., & Béjoint, H. (1992). *Vocabulary and applied linguistics*. Macmillan.

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

Behrens, H. (2009). Usage-based and emergentist approaches to language acquisition.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 711-733.

Callies, M. (2019). Integrating corpus literacy into language teacher education. *Learner Corpora and Language Teaching, 92, 245*. https://doi.org/10.1075/slcs.201.12cal

Cameron, L. (2001). *Teaching languages to young learners*. Cambridge University Press.

Carter, R., & McCarthy, M. (Eds.) (1988). *Vocabulary and language teaching*. Longman.

Chambers, A. (2019). Towards the corpus revolution: Bridging the research-practice gap. *Language Teaching*, 1-16. https://doi.org/10.1017/S0261444819000089

Chen, C., & Truscott, J. (2010). The effects of repetition and L1 lexicalization on incidental vocabulary acquisition. *Applied Linguistics, 31*(5), 693-713.

Coady, J., & Huckin, T. (Eds.). (1997). *Second language vocabulary acquisition.* Cambridge University Press.

Cobb. T. (2021, October 1). *Why & how to use frequency lists to learn words.* http://www.lextutor.ca/ResearchWeb/

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*(6), 671-684.

Criado, R. (2009). The distribution of the lexical component in ELT coursebooks and its suitability for vocabulary acquisition from a cognitive perspective: A case study. *International Journal of English Studies, 9* (3), 39–60.

Divjak, D. (2019). *Frequency in language: Memory, attention and learning*. Cambridge University Press.

Doughty, C. J., & Long, M. H. (2003). The scope of inquiry and goals of SLA. *The handbook of second language acquisition*, 3-16.

Ellis, N. (2013). Construction grammar and second language acquisition. *The Oxford handbook of construction grammar*.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*(2), 143-188.

Ellis, N. C. (2012). Frequency-based accounts of second language acquisition. *The Routledge handbook of second language acquisition*, 193-210.

Erman, B. (2009). Formulaic language from a learner perspective. *Formulaic Language, 2*, 323-346.

Folse, K. (2004). *Vocabulary myths: Applying second language research to classroom teaching.* The University of Michigan Press.

Frankenberg-Garcia, A. (2012). Raising teachers' awareness of corpora. *Language Teaching*, *45*, 475–489. https://doi.org/10.1017/S0261444810000480

Fujimori, J. (2005). The lexical composition of two oral communication textbooks. *The Language Teacher, 29*(7), 15–19.

Gass, S. M. (1988). Integrating research areas: A framework for second language studies. *Applied linguistics*, *9*(2), 198-217.

Hall, J. F. (1954). Learning as a function of word-frequency. *The American Journal of Psychology*, *67*(1), 138-140.

Harmer, J. (1991). *The practice of English teaching*. Longman.

Harwood, N. (2010). Issues in materials development and design. *English language teaching materials: Theory and practice* (pp. 3-30). Cambridge University Press.

Heather, J., & Helt, M. (2012). Evaluating corpus literacy training for pre-service language teachers: Six case studies. *Journal of Technology and Teacher Education*, *20*(4), 415-440.

Hirsch, E. D. (2003). Reading comprehension requires knowledge of words and the world. *American Educator*, *27*(1), 10-13.

Horst, M. (2013). Mainstreaming second language vocabulary acquisition. *The Canadian Journal of Applied Linguistics*, *16*(1), 171-188.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond A Clockwork Orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language. 11*, 207-223.

Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology*, *41*(6), 401.

Koirala, C. (2015, August). The word frequency effect on second language vocabulary learning. *Critical CALL–Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (pp. 318-323). Research-publishing net. http://dx.doi.org/10.14705/rpnet.2015.000352

Koosha, M., & Sharifi, M. (2004). Estimating the vocabulary size of Iranian EFL learners. *Journal of Science and Research*, *4*(1), 15-30.

Krashen, S. (1981). Second language acquisition. *Second Language Learning*, *3*(7), 19-39.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal*, *73*(4), 440-464.

Krashen, S. D., & Terrell, T. (1983). *Natural approach* (pp. 20-20). Pergamon.

Laufer, B. (1986). Possible changes in attitude towards vocabulary acquisition research. *International Review of Applied Linguistics*, *24*(1), 69-75.

Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. *Second language vocabulary acquisition*.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. *Language Learning*, *54*(3), 399-436.

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*(1), 1-26.

Laufer, B., & Sim, D. D. (1985). Measuring and explaining the reading threshold needed for English for academic purposes texts. *Foreign Language Annuals*, *18*(5), 405-411.

Lewis, M. (1993). *The lexical approach.* Language teaching publications.

MacKay, D. G. (1982). The problems of flexibility, fluency, and speed–accuracy trade-off in skilled behavior. *Psychological Review, 89*(5), 483.

Matsuoka, W. (2012). Searching for the right words: Creating word lists to inform EFL learning. In: D. Hirsh (Ed.), *Current perspectives in second language vocabulary research* (pp. 151– 177). Peter Lang.

McCarthy, M. (2008). Accessing and interpreting corpus information in the teacher education context. *Language Teaching*, *41*(4), 563–574. https://doi.org/10.1017/S0261444808005247

Melka, F. (1997). Receptive vs. productive aspects of vocabulary. *Vocabulary: Description, Acquisition and Pedagogy*, *33*(2), 84-102.

Michael, T. (2003). *Constructing a language: a usage-based theory of language acquisition*. Harvard University Press.

Moravcsik, H. Ouali & K. M. Wheatley (Eds.), *Formulaic language: Acquisition, loss, psychological reality, and functional explanations* (pp. 323–346). John Benjamins.

Mukundan, J., & Aziz, A. (2009). Loading and distribution of the 2000 high frequency words in Malaysian English language textbooks for Form 1 to Form 5. *Pertanika Journal of Social Sciences and Humanities*, *17*(2), 141-152.

Nagy, W., Scott, J., Kamil, M. L., Mosenthal, P. B., Pearson, P. D., & Barr, R. (2000). Vocabulary processes. *Handbook of reading research*, *3*.

Nation, I. (2006). How large a vocabulary is needed for reading and listening. *Canadian Modern Language Review, 63*(1), 59-82.

Nation, I. S. (2013). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P. (1990). *Teaching and learning vocabulary.* Heinle & Heinle Publishers.

Nation, I. S. P. (1993b). Vocabulary size, growth, and use. *The Bilingual Lexicon*, *6*, 115-134.

Nation, P. (1993a). Using dictionaries to estimate vocabulary size: Essential, but rarely followed, procedures. *Language Testing*, *10*(1), 27-40.

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

Nation, P. (2008). *Teaching vocabulary: Strategies and techniques*. Heinle Cengage Learning.

Nation, P., & Wang, K. (1999). Graded readers and vocabulary. *Reading in a Foreign Language, 12*(2), 355–380.

Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition. *Cognitive skills and their acquisition*.

Norberg, C., & Nordlund, M. (2018). A corpus-based study of lexis in L2 English textbooks. *Journal of Language Teaching and Research, (9*), 463–473.

Nordlund, M. (2016). EFL textbooks for young learners: A comparative analysis of vocabulary. *Education Inquiry, 7*(1), 27764.

Nordlund, M., & Norberg, C. (2020). Vocabulary in EFL teaching materials for young learners. *International Journal of Language Studies, 14*(1), 89-116.

Reda, G. (2003). English coursebooks: Prototype texts and basic vocabulary norms. *ELT Journal*, *57*(3), 260-268.

Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge University Press.

Reppen, R., Bunting, J., Diniz, L., Blass, L., Iannuzzi, S. & Savage, A. (2012). *Grammar and beyond*. Cambridge University Press.

Saragi, T., Nation, P. & Meister, G.F. (1978). Vocabulary learning and reading. *System. 6*, 72-78.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology: Human perception and performance*, *3*(1), 1.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge University

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329-363.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, *12*(3), 329–363.

Schmitt, N., & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *Studies in Second Language Acquisition*, *19*(1), 17-36.

Schmitt, N., Cobb, T., Horst, M., & Schmitt, D. (2017). How much vocabulary is needed to use English? Replication of van Zeeland & Schmitt (2012), Nation (2006) and Cobb (2007). *Language Teaching*, *50*(2), 212-226.

Schön, D. A. (1987). *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. Jossey-Bass.

Şimşek, T. (2020). *Corpora in Foreign Language Teacher Education: Introducing a Corpus Literacy Course to ELT Pre-service Teachers.* [Doctoral dissertation, Çukurova University].

Sinclair, J., & Sinclair, L. (1991). *Corpus, concordance, collocation*. Oxford University Press.

Thornbury, S. (2002). Using dictionaries. *How to teach vocabulary. Pearson Education.*

Tomasello, M., & Brooks, P. J. (1999). Early syntactic development: A construction grammar approach. *The Development of Language*, 161-190.

Tyler, A. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. Routledge.

Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, *22*(2), 217-234.

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, *15*(2), 130–163.

APPENDIX I: HEADWORDS OF THE FIRST 1,000 MOST FREQUENT ENGLISH WORD LIST FOUND IN TOUCHSTONE 3

ability	age	any
able	agency	anyone
about	agent	anything
above	ago	anyway
accept	agree	appear
access	agreement	application
accord	ahead	apply
account	aim	approach
achieve	air	area
across	all	argue
act	allow	argument
action	almost	arm
activity	alone	around
actually	along	arrive
add	already	art
addition	also	article
address	although	as
adult	always	ask
advance	among	associate
advantage	amount	assume
affect	analysis	at
after	and	attack
afternoon	animal	attempt
again	another	attend
against	answer	attention

author
available
average
avoid
away
baby
back
bad
balance
bank
bar
base
basic
basis
be
bear
beautiful
because
become
bed
before
begin
behavior
behind
believe
below
benefit
between
beyond
big
bill
bit
black
blue
board
body
book
both
box
boy
break
bring
brother
build
bus
business
but
buy
by
call
campaign
can
capital
car
card
care
career
carry
case
catch
cause
cell
cent
center
central
century
certain
certainly
challenge
chance
change
character
charge
check
child
choice
choose
church
city
claim
class

clean
clear
clearly
clock
close
clothes
club
cold
college
color
come
comment
common
communication
community
company
compare
competition
complete
completely
computer
concern
conclusion
condition
consider
contact
contain
continue
contract
control
conversation
cost
could
country
couple
course
court
cover
create
credit
cross
culture
current
customer
cut
daily
damage
dance
dark
data
date
daughter
day
dead
deal
death
debate
decide
decision
deep
define
degree
demand
department
depend
describe
design
despite
detail
determine
develop
development
die
difference
different
difficult
difficulty
direct
direction
director
discover

discuss
discussion
disease
do
doctor
document
dog
dollar
door
doubt
down
draw
dream
dress
drink
drive
drop
drug
due
during
each
early
earn
easily
easy
eat
economic
economy
education
effect
effort
either
election
else
employee
encourage
end
energy
enjoy
enough
enter
environment
especially
establish
estimate
even
evening
event
ever
every
everybody
everyone
everything
evidence
exactly
example
except
exercise
exist
expect
experience
explain
express
extend
eye
face
facility
fact
factor
fail
fall
family
far
fast
father
favorite
fear
feature
feel
female
few

field
fight
figure
file
fill
film
final
finally
financial
find
fine
finish
fire
firm
first
fish
fit
fix
floor
fly
focus
follow
food
foot
for
force
foreign
forget
form
former
forward
free
friend
from
front
full
fun
function
fund
further
future
gain
game
garden
general
generally
generation
get
girl
give
go
goal
good
government
great
green
ground
group
grow
growth
guess
guy
hair
half
hand
happen
happy
hard
have
he
head
health
hear
heart
help
here
herself
high
himself
history
hit

hold
holiday
home
hope
hospital
hot
hotel
hour
house
how
however
huge
human
husband
idea
identify
if
image
imagine
immediately
important
improve
in
include
income
increase
indeed
indicate
individual
industry
influence
information
inside
instance
instead
interest
international
interview
into
introduce
investment
involve
issue
it
item
itself
job
join
judge
just
keep
key
kid
kill
kind
know
knowledge
lack
land
language
large
last
late
laugh
law
lead
leader
learn
least
leave
legal
less
let
letter
level
lie
life
light
like
likely
limit

line
link
list
listen
little
live
local
long
look
lose
loss
lot
love
low
machine
magazine
main
maintain
major
make
male
man
manage
management
manager
many
mark
market
marry
match
material
matter
may
maybe
mean
measure
medium
meet
member
memory
mention
message
method
middle
might
mile
military
mind
mine
minute
miss
model
modern
moment
money
month
more
morning
most
mother
move
movement
movie
mr
mrs
much
music
must
myself
name
nation
national
natural
nature
near
nearly
necessary
need
network
never
new

news
newspaper
next
nice
night
no
nor
normal
not
note
nothing
notice
now
number
object
obviously
occur
of
off
offer
office
officer
official
often
oil
okay
old
on
once
one
only
open
operate
operation
opinion
opportunity
option
or
order
organization
organize
original
other
out
outside
over
own
page
paint
paper
parent
park
part
particular
particularly
party
pass
past
patient
pattern
pause
pay
people
per
percent
perform
performance
perhaps
period
person
personal
phone
pick
picture
piece
place
plan
plant
play
player
please

point
police
policy
political
poor
popular
population
position
positive
possibility
possible
post
potential
pound
power
practice
prefer
prepare
present
president
press
pressure
pretty
previous
price
private
probably
problem
process
produce
product
production
professional
profit
program
project
promise
property
proposal
propose
protect
prove
provide
public
pull
purchase
purpose
push
put
quality
quarter
question
quickly
quite
race
raise
range
rate
rather
reach
read
ready
real
realize
really
reason
receive
recent
recently
recognize
record
red
reduce
refer
reference
reflect
regard
region
relate
relation
relationship

release
remain
remember
replace
report
represent
require
research
resource
respect
response
responsibility
rest
restaurant
result
return
review
right
ring
rise
risk
road
rock
role
room
round
rule
run
sale
same
save
say
school
science
score
screen
sea
search
season
seat
second
section
security
see
seek
seem
sell
send
sense
separate
series
serious
serve
service
set
several
shall
shape
share
she
shop
short
should
show
side
sign
significant
similar
simple
simply
since
sing
single
sister
sit
site
situation
size
skill
sleep
small

smile
so
social
society
solution
some
somebody
someone
something
sometimes
son
song
soon
sorry
sort
sound
source
space
speak
special
specific
speech
spend
sport
staff
stage
stand
standard
star
start
state
statement
station
stay
step
still
stock
stop
store
story
strategy
street
strike
strong
structure
student
study
stuff
style
subject
success
successful
such
suddenly
suffer
suggest
summer
supply
support
suppose
sure
surprise
system
table
take
talk
target
task
tax
teach
teacher
team
technology
television
tell
tend
term
test
than
thank
that

the
themselves
then
theory
there
therefore
they
thing
think
this
though
through
throughout
throw
thus
time
to
today
together
tomorrow
too
top
total
touch
toward
town
track
trade
traditional
train
travel
treat
treatment
tree
trip
trouble
true
try
turn
type
unclear
under
understand
unit
university
until
up
upon
use
useful
usually
value
variety
various
version
very
view
village
visit
voice
vote
wait
walk
wall
want
war
watch
water
way
we
wear
week
weekend
weight
welcome
well
what
whatever
when
where
whether

which
while
white
who
whole
why
wide
wife
will
win
window
wish
with
within
without
woman
wonder
word
work
worker
world
worry
worth
would
write
wrong
yeah
year
yes
yesterday
yet
you
young
yourself

APPENDIX II: HEADWORDS OF THE SECOND 1,000 MOST FREQUENT ENGLISH WORD LIST FOUND IN TOUCHSTONE 3

abandon
abroad
absence
absolutely
abuse
academic
accident
accommodation
accompany
accuse
achievement
acknowledge
acquire
active
actor
actual
ad
additional
administration
admit
adopt
advertise
advertisement
advice
advise
affair
afford
afraid
aid
aircraft
alive
alternative
amaze
analyze
ancient
angry
announce
annual
anybody
anywhere
apart
apparently
appeal
appearance
appoint
appointment
appreciate
appropriate
approve
approximately
arise
army
arrange
arrangement
arrest
artist
aside
assess
assessment
asset
assistant
association
assumption
atmosphere
attach
attitude
attract
attractive
audience
award
aware
background
bag
ball
ban

band
basically
battle
beach
beat
beauty
bedroom
belief
belong
besides
bet
bike
bind
bird
birth
blame
block
blood
blow
boat
bomb
bond
bone
border
bore
borrow
boss
bother
bottle
bottom
brain
branch
brand
breakfast
breast
breath
breathe
bridge
brief
bright
broad
brown
budget
burn
busy
calculate
calm
camera
camp
cancer
capable
capacity
capture
careful
carefully
cash
cast
cat
category
celebrate
chain
chair
chairman
channel
chapter
characteristic
chart
cheap
chemical
chief
childhood
chip
circle
circumstance
citizen
civil
classic
climate
climb
closely
coach

coast

coat

code

coffee

collapse

colleague

collect

collection

column

combination

combine

comfort

comfortable

command

commercial

commission

commit

commitment

committee

communicate

comparison

compete

competitive

complain

complaint

complex

complicate

component

concentrate

concentration

concept

concert

conclude

conduct

confidence

confirm

conflict

confuse

connect

connection

consequence

conservative

considerable

consideration

consist

consistent

constant

construct

construction

consumer

content

context

contrast

contribute

contribution

convention

convince

cook

cool

cope

copy

core

corner

corporate

corporation

correct

council

count

county

crash

creation

crime

criminal

crisis

criterion

critic

critical

criticism

crop

crowd

crucial

cry

cultural

cup

currently

cycle

dad

danger

dangerous

dear

debt

decade

declare

decline

decrease

defeat

defend

defense

definitely

definition

delay

delight

deliver

delivery

democracy

demonstrate

deny

derive

description

deserve

desire

desk

destroy

device

diet

differ

dinner

directly

disappear

disappoint

discipline

discount

dish

disk

display

distance

distinguish

distribute

district

divide

division

domestic

double

dozen

draft

drama

dramatic

driver

dry

duty

ear

earth

east

edge

editor

educate

educational

effective

effectively

efficient

egg

elect

electronic

element

elsewhere

email

emerge

emergency

emotion

emotional

emphasize

employ

employer
employment
empty
enable
encounter
enemy
engage
engine
engineer
enhance
enormous
ensure
entire
entirely
entry
environmental
equal
equally
equipment
error
escape
essential
estate
eventually
everywhere
examination
examine
excellent
exception
exchange
excite
excuse
executive
exhibit
exhibition
existence
expand
expectation
expense
expensive
experiment
expert
explanation
explore
export
expose
expression
extent
extra
extreme
extremely
factory
failure
fair
fairly
faith
familiar
famous
fan
farm
farmer
fashion
fat
fault
favor
federal
fee
feed
fellow
finance
finger
flat
flight
flow
flower
folk
football
forest
formal
forth
frame

freedom
freeze
frequently
fresh
friendly
fruit
fuel
fully
fundamental
funny
furthermore
gap
gas
gather
gene
generate
gentleman
gift
glad
glass
global
gold
grade
graduate
grant
grateful
gray
guarantee
guard
guest
guide
gun
habit
hall
handle
hang
hardly
hat
hate
healthy
heat
heavy
hell
helpful
hi
hide
highlight
highly
hill
hire
historical
hole
honor
horse
host
household
hunt
hurt
ice
ideal
identity
ignore
ill
illness
illustrate
immediate
implement
implication
imply
import
importance
impose
impossible
impression
improvement
incident
increasingly
independent
index
industrial
inform

initial
initiative
injury
insist
install
institution
instruction
instrument
insurance
integrate
intellectual
intelligence
intend
intention
interaction
internal
interpretation
introduction
invest
investigate
investigation
investor
invite
island
joint
joke
journalist
journey
judgment
jump
justice
kick
king
kiss
kitchen
knock
label
labor
lady
lake
landscape
largely
latter
launch
lawyer
lay
leadership
lean
left
leg
legislation
length
lesson
liberal
library
license
lift
literature
load
loan
locate
location
lock
lovely
luck
lucky
lunch
mail
mainly
majority
manner
manufacture
manufacturer
map
marriage
mass
massive
master
mathematics
maximum
meal

meanwhile

meat

mechanism

medical

medicine

mental

merely

metal

milk

minimum

minister

minor

minority

mission

mistake

mix

mobile

monitor

moral

moreover

mostly

motion

mountain

mouse

mouth

multiple

murder

museum

musical

narrow

native

nearby

necessarily

neck

negative

negotiate

negotiation

neighbor

neighborhood

neither

net

nevertheless

nobody

noise

none

normally

north

northern

notion

novel

nuclear

nurse

objective

observation

observe

obvious

occasion

occupy

odd

online

onto

oppose

opposite

opposition

ordinary

origin

originally

otherwise

ought

ourselves

outcome

outline

output

overall

owner

pack

package

pain

pair

panel

participant

participate
partly
partner
passage
passenger
path
payment
peace
peak
percentage
perfect
perfectly
permanent
permit
personality
personally
perspective
phase
photo
photograph
phrase
physical
pilot
plain
plane
plastic
plate
pleasure
plenty
plot
plus
pocket
politician
politics
poll
pool
pop
possibly
powerful
practical
predict
premise
presence
presentation
preserve
presidential
prevent
previously
primarily
primary
prime
principal
principle
print
prior
priority
prison
procedure
proceed
producer
professor
progress
promote
proper
properly
proportion
prospect
protection
protein
protest
proud
provision
pursue
qualify
quick
quiet
quote
radio
rain
rank
rapidly

rare
reaction
reader
reality
reasonable
recall
recommend
recommendation
recover
reduction
reform
refuse
regional
register
regular
regulation
reject
relative
relatively
relax
relevant
relief
religion
religious
rely
remark
remind
remove
rent
repair
repeat
reply
reporter
representative
reputation
request
requirement
researcher
reserve
resident
resolution
resolve
respond
responsible
retain
retire
reveal
revenue
revolution
reward
rich
ride
river
roll
root
route
row
rural
rush
sad
safe
safety
salary
sample
satisfy
scale
scene
schedule
scheme
scientific
scientist
secondly
secret
secretary
sector
secure
seed
select
selection
self
senior

sentence
sequence
seriously
session
settle
settlement
severe
sex
sexual
shake
sharp
sheet
shift
ship
shirt
shock
shoe
shoot
shot
shoulder
shout
shut
sick
sight
signal
significantly
silence
skin
sky
slide
slightly
slip
slow
slowly
smell
smoke
snow
soft
software
soldier
solve
somehow
somewhat
somewhere
soul
south
southern
speaker
specifically
speed
spell
spirit
split
spot
spread
spring
square
stable
status
steal
stick
stone
storm
straight
strange
strength
stress
stretch
strongly
struggle
studio
substantial
succeed
sudden
sufficient
suggestion
suit
suitable
sum
sun
surely

surface
surround
survey
survive
suspect
sweet
swim
swing
switch
talent
tall
tape
taste
taxi
tea
tear
technical
technique
telephone
temperature
temporary
tennis
terrible
text
theater
theme
thick
thin
threat
threaten
ticket
tie
till
tiny
tip
tire
title
tone
tonight
tool
topic
totally
tough
tour
tourist
toy
tradition
traffic
transfer
transport
trend
trial
troop
truly
trust
truth
twice
typical
unable
unemployment
unfortunately
union
unique
unless
unlike
unusual
update
upper
urban
urge
user
usual
valuable
variable
vary
vast
vehicle
via
victim
victory
video

violence
vision
visitor
volume
volunteer
voter
wage
wake
warm
warn
wash
waste
wave
weak
weapon
weather
wed
west
wheel
whereas
widely
wild
wind
wine
winner
winter
witness
wonderful
wood
writer
yard
yellow
yield
youth

APPENDIX III: HEADWORDS OF THE ACADEMIC WORD LIST FOUND IN TOUCHSTONE 3

abdominal
absorb
absorption
accelerate
acceleration
accent
accumulate
accumulation
accuracy
accurately
acid
acidic
activate
actively
acute
adaptation
adaptive
adjacent
admission
adolescent
adverse
aerosol
aesthetic
affirm
afterward
aggregate
agriculture
airplane
algebra
algorithm
alien
alliance
allocate
allocation
altitude
aluminum
amino
amongst
amplitude
analogy
ancestor
anthropology
anti
antibiotic
antibody
antiquity
appendix
applause
apple
approximate
approximation
arbitrary
archaeology
architect
array
arrow
articulate
artifact
artificial
artistic
artwork
aspect
assembly
assert
assignment
athletic
atom
atomic
auction
audio
audit
authority
autonomy
availability
axiom
axis
backward
bacteria

bacterial
bang
bargain
barrel
basin
bat
beam
behavioral
bilingual
binary
biodiversity
biologist
biology
bizarre
blank
bleed
bodily
bonus
bound
bracket
breakdown
broadly
bubble
bucket
bulk
bullet
bundle

calcium
calculation
calculator
calculus
campus
candidate
capillary
capitalism
capitalist
carrier
cattle
censor
characterization
cheat
cheer
chemistry
chemotherapy
chess
chloride
chromosome
chronic
chunk
cinema
circa
circulate
circulation
civilization

clarify
classification
classify
classroom
clay
clever
click
client
clinic
clip
clone
closure
clue
coefficient
cognitive
coherent
collective
colonial
colony
comma
commentary
commodity
commonly
communicative
communist
compact
comparable

comparative
compensate
competence
competent
complement
complication
comply
composer
composite
comprehension
computation
conceive
conception
conceptual
conditional
conduction
cone
conference
configuration
confine
confound
congruent
connector
connotation
conscious
consciousness
consensus
consent
conservation
conserve
consonant
constitution
constrain
consultation
consumption
container
continent
continuity
contour
contradict
contradiction
contradictory
contrary
controversy
converge
convergence
coordinate
coordination
cord
coronary
corpus
correction
correctly
correlate
correlation
correspondence
corruption
cortex
credibility
criteria
critically
critique
crude
crystal
cue
cure
curriculum
cyclic
cylinder
damp
deadline
decay
deceive
defect
deficiency
definite
deflection
degrade
deliberately
delta
demonstrator

denominator
denote
dense
dependence
depict
derivative
descendent
descriptor
destination
detection
developmental
deviation
diagnose
diagnosis
diagnostic
diagram
dialect
diameter
diary
dictate
dictionary
differential
differentiate
differentiation
diffusion
dilemma
dilute
dimensional
dioxide
directive
disability
disadvantage
discharge
discourse
discrete
discrimination
displacement
dissection
dissertation
dissolve
distribution
disturbance
diverse
domain
dominance
dominant
domination
donor
dose
drain
drift
duration
dye
dynamic
ecological
ecology
economically
economist
effectiveness
elaborate
elastic
elasticity
electron
elementary
elevate
elevation
elimination
elite
embed
emergence
emission
emit
emperor
empirical
encode
enforcement
enormously
entity
entrant
enzyme
epidemic

epidemiology
equality
equilibrium
equivalence
erase
essence
essentially
estimation
ethical
ethics
evident
evolutionary
ex
execute
execution
exit
expertise
explicit
explicitly
exploit
exponential
fabric
facet
facilitate
factorial
faculty
fatigue
feedback
fertility
fetal
fever
fiber
fin
finite
flesh
flexibility
flip
fluid
flux
footnote
formally
formulation
forum
fossil
founds
fraction
fracture
freely
friction
fringe
fundamentally
fungus
fusion
gauge
generalization
generalize
genetically
genetics
ghost
globalization
goat
goodness
goods
gradient
gram
grammatical
graph
grasp
gravity
grid
gross
gut
habitat
halfway
handout
harvest
hawk
headquarter
hedge
helix
hepatitis

herbicide
hierarchy
hip
historically
homework
horizon
horizontal
hormone
hydrogen
identical
identification
ideology
illusion
immune
impact
implicit
importantly
impulse
incidence
incline
inclusion
incredible
incredibly
incumbent
independently
indicator
indifference
indigenous
indirect
individually
induce
induction
industrialization
industrialize
inequality
inevitably
infect
infectious
inference
inferior
infinite
infinity
influential
informal
inhibit
inhibition
initiate
initiation
inject
injection
innate
insect
insert
instability
instinct
integral
integration
intensity
intensive
interact
interestingly
interface
interfere
intermediate
interrupt
interval
intervene
interviewer
invade
invasion
inversion
invert
ion
irrelevant
irrigation
jazz
junior
justification
kidney
kilometer
lab

lateral
layout
leaf
lecturer
legend
legitimate
leisure
lever
lexical
liable
lifestyle
lifetime
likelihood
likewise
limb
linear
linguistic
linguistics
liter
liver
locally
locus
logical
longitudinal
loop
lump
lung

machinery
magnetic
magnitude
mall
manipulate
manipulation
manual
manuscript
marble
marginal
marker
marrow
maternal
mathematical
matrix
maximize
meaningful
mechanic
mechanical
media
mediate
membrane
memorize
mentor
mercury
merge
messenger

metabolism
metaphor
methodology
micro
mid
migrate
migration
millimeter
mineral
minimal
minimize
minus
missile
mobility
modification
mole
molecular
molecule
momentum
monetary
monkey
monopoly
morality
morphological
morphology
mortality
motif

motive
multi
multinational
multiply
mutation
myth
naked
namely
nasty
nationalism
neat
necessity
neo
nest
neural
neuron
neutral
nicely
niche
nitrogen
noble
node
noisy
nominal
non
nonetheless
nonlinear
norm
notation
novice
nucleus
null
numerical
nutrient
objection
obscure
observer
obtain
occupation
occurrence
offspring
onset
onwards
optical
optimal
optimum
oral
organ
organism
orient
orientation
oscillation
outer
outlet
overhead
overlap
overview
oxidize
oxygen
painful
par
paradigm
paradox
parameter
parcel
pardon
parental
parenthesis
partial
partially
particle
partition
pathway
peasant
periodic
peripheral
pest
pesticide
phenomenal
philosopher
philosophical

phonological
phosphate
photographic
physically
physician
physics
physiological
pi
planner
plantation
plausible
plug
plural
polar
pole
politically
portfolio
portray
portrayal
positively
poster
postgraduate
potassium
powder
practitioner
pragmatic
pre
precede
precipitate
precipitation
predator
prediction
predominantly
prejudice
preliminary
presume
prevalence
prey
primer
primitive
probe
problematic
processor
productive
productivity
profound
progression
progressive
projection
prominent
pronounce
proposition
protocol
proton
psychiatric
psychologist
psychology
publish
pulse
punch
punish
punishment
purely
puzzle
qualitative
quantitative
quantum
questionnaire
quiz
quotation
rack
radar
radiation
radius
rainfall
randomize
randomly
rational
rationality
ray
reactive

reactor
readily
realism
realistic
realm
receptor
recipe
reconstruct
regime
regression
rehabilitation
reinforce
rejection
reliability
render
repertoire
replacement
replicate
replication
reproduce
reproduction
republic
resemble
reservoir
residual
residue
resistant
respiratory
retrieve
revolutionary
rewrite
rhetoric
rhythm
ridiculous
ritual
robot
rope
rotate
rotation
rub
ruler
scatter
scenario
scholarship
scripture
scroll
secrete
sediment
selective
semantic
semester
semi
seminar
sensation
sensible
sensitivity
sensory
separately
separation
sexuality
shallow
shortly
shuttle
similarity
simplify
simulate
simulation
simultaneously
sin
singular
sketch
skip
slab
slash
slavery
slot
snake
sneeze
sniff
socialize
socially

sodium
solar
soluble
solute
solvent
sometime
sophisticate
span
spatial
specialty
specification
specimen
spectrum
sperm
sphere
splice
sponsorship
spontaneous
spray
stabilize
stack
stadium
stance
standardize
static
statistical
statistically
statistics
stereotype
stimulus
straightforward
strand
strategic
strictly
stripe
sub
subjective
subset
substitution
substrate
subtle
subtract
sufficiently
suicide
super
superior
supposedly
surgeon
surgical
surplus
susceptible
sustainable
swap
swell
sword
syllable
syllabus
symbolic
syndrome
syntactic
syntax
synthesis
synthetic
systematic
tech
technically
temporal
tempt
tense
terminal
terminology
terribly
textbook
theorem
theorist
thereby
thermal
thesis
thickness
thread
threshold

thumb

tolerance

ton

toxic

tract

traditionally

trait

trajectory

trans

transaction

transcribe

transcription

transformation

translation

transmission

transmit

transparency

transplant

trauma

treaty

tremendous

triangle

tribe

tricky

trivial

tropical

tumor

turbulent

tutor

ultimate

undergraduate

undermine

underneath

unemployed

unify

unintelligible

unity

unstable

uplift

upward

urine

usage

utility

utilize

utterance

vague

valid

validity

valve

variability

variance

variant

vector

vegetation

vein

velocity

verbal

vertical

viable

virtue

vitamin

vocabulary

volition

vowel

watershed

wavelength

weave

wheat

whereby

whichever

whoever

widespread

wisdom

workshop

yeast

APPENDIX IV: OFF-LIST WORDS OF TOUCHSTONE 3

academically
academy
advertisers
alpha
ambitions
amphibian
appetizer
archives
atm
avenue
aw
bare
beans
binoculars
blog
blogs
blur
born
brillant
butter
cab
calorie
chandeliers
chaperones
chickpeas
chopped
chute
chutney
clutter
cluttered
co
coaster
com
contributors
convenience
corn
costly
costume
crabs
daisies
dart
deadliest
delicious
discarded
discreetly
diving
dot
eligibility
escalator
etiquette
fabulous
fluently
flunking
footballers
fork
founder
fried
fries
frog
frustrated
garb
garlic
grooming
happily
humble
id
iguanas
imaginary
impeccable
ingredients
inherited
instant
interactive
intercepted
internet
jewelry
journalism
junk
jupiter
ketchup
lamb
lee
lentils
lighter
limber
lions
lo

mamba
mandarin
manicure
mayonnaise
melted
merry
metro
metropolitan
morel
motivational multimedia
nutritious
oh
ohmynews
onion
packets
pals
para
parrot
paste
pastries
pat
pedicure
penguins
photography
pianist
pilfer
pin
poison
poisonous
polite
proficient
ramblings
receipts
reconnect
reunion
reunions
reunite
ridiculously
rifle
robbing
rocky
roller
sanely
sauce
scuba
sears
semiconductor
sesame
shred
sibling
slouched
snack
snacks
snatch
snorkeled
songwriter
spices
spicy
spy
statues
straightened
stumbled
subscribers
subway
subways
supervises
supervisor
supportive
surfers
suspicious
suzanne
tag
tahini
talkative
teeming
teen
telescopes
theft
thermometer
thief
thieves
torn
tortoises
trash
treasure
turtles
unauthorized
undetected
unlimited
upperclassmen

vastly

vegetarian

versatile

vinegar

volcanic

webcasting

yelled

APPENDIX V: HEADWORDS OF THE FIRST 1,000 MOST FREQUENT ENGLISH WORD LIST FOUND IN HEADWAY PRE-INTERMEDIATE

a	age	anyone
ability	agency	anything
able	agent	anyway
about	ago	appear
above	agree	application
accept	agreement	apply
access	ahead	approach
accord	aim	area
account	air	argue
achieve	all	argument
across	allow	arm
act	almost	around
action	alone	arrive
activity	along	art
actually	already	article
add	also	as
addition	although	ask
address	always	associate
adult	among	assume
advance	amount	at
advantage	analysis	attack
affect	and	attempt
after	animal	attend
afternoon	another	attention
again	answer	author
against	any	available

average
avoid
away
baby
back
bad
balance
bank
bar
base
basic
basis
be
bear
beautiful
because
become
bed
before
begin
behavior
behind
believe
below
benefit
between
beyond
big
bill
bit
black
blue
board
body
book
both
box
boy
break
bring
brother
build
bus
business
but
buy
by
call
campaign
can
capital
car
card
care
career
carry
case
catch
cause
cell
cent
center
central
century
certain
certainly
challenge
chance
change
character
charge
check
child
choice
choose
church
city
claim
class
clean
clear

clearly
clock
close
clothes
club
cold
college
color
come
comment
common
communication
community
company
compare
competition
complete
completely
computer
concern
conclusion
condition
consider
contact
contain
continue
contract
control
conversation
cost
could
country
couple
course
court
cover
create
credit
cross
culture
current
customer
cut
daily
damage
dance
dark
data
date
daughter
day
dead
deal
death
debate
decide
decision
deep
define
degree
demand
department
depend
describe
design
despite
detail
determine
develop
development
die
difference
different
difficult
difficulty
direct
direction
director
discover
discuss
discussion

disease
do
doctor
document
dog
dollar
door
doubt
down
draw
dream
dress
drink
drive
drop
drug
due
during
each
early
earn
easily
easy
eat
economic
economy
education
effect
effort
either
election
else
employee
encourage
end
energy
enjoy
enough
enter
environment
especially
establish
estimate
even
evening
event
ever
every
everybody
everyone
everything
evidence
exactly
example
except
exercise
exist
expect
experience
explain
express
extend
eye
face
facility
fact
factor
fail
fall
family
far
fast
father
favorite
fear
feature
feel
female
few
field
fight

figure
file
fill
film
final
finally
financial
find
fine
finish
fire
firm
first
fish
fit
fix
floor
fly
focus
follow
food
foot
for
force
foreign
forget
form
former
forward
free
friend
from
front
full
fun
function
fund
further
future
gain
game
garden
general
generally
generation
get
girl
give
go
goal
good
government
great
green
ground
group
grow
growth
guess
guy
hair
half
hand
happen
happy
hard
have
he
head
health
hear
heart
help
here
herself
high
himself
history
hit
hold
holiday

home
hope
hospital
hot
hotel
hour
house
how
however
huge
human
husband
i
idea
identify
if
image
imagine
immediately
important
improve
in
include
income
increase
indeed
indicate
individual
industry
influence
information
inside
instance
instead
interest
international
interview
into
introduce
investment
involve
issue
it
item
itself
job
join
judge
just
keep
key
kid
kill
kind
know
knowledge
lack
land
language
large
last
late
laugh
law
lead
leader
learn
least
leave
legal
less
let
letter
level
lie
life
light
like
likely
limit
line

link
list
listen
little
live
local
long
look
lose
loss
lot
love
low
machine
magazine
main
maintain
major
make
male
man
manage
management
manager
many
mark
market
marry
match
material
matter
may
maybe
mean
measure
medium
meet
member
memory
mention
message
method
middle
might
mile
military
mind
mine
minute
miss
model
modern
moment
money
month
more
morning
most
mother
move
movement
movie
mr
mrs
much
music
must
myself
name
nation
national
natural
nature
near
nearly
necessary
need
network
never
new
news

newspaper
next
nice
night
no
nor
normal
not
note
nothing
notice
now
number
object
obviously
occur
of
off
offer
office
officer
official
often
oil
okay
old
on
once
one
only
open
operate
operation
opinion
opportunity
option
or
order
organization
organize
original
other
out
outside
over
own
page
paint
paper
parent
park
part
particular
particularly
party
pass
past
patient
pattern
pause
pay
people
per
percent
perform
performance
perhaps
period
person
personal
phone
pick
picture
piece
place
plan
plant
play
player
please
point

police
policy
political
poor
popular
population
position
positive
possibility
possible
post
potential
pound
power
practice
prefer
prepare
present
president
press
pressure
pretty
previous
price
private
probably
problem
process
produce
product
production
professional
profit
program
project
promise
property
proposal
propose
protect
prove
provide
public
pull
purchase
purpose
push
put
quality
quarter
question
quickly
quite
race
raise
range
rate
rather
reach
read
ready
real
realize
really
reason
receive
recent
recently
recognize
record
red
reduce
refer
reference
reflect
regard
region
relate
relation
relationship
release

remain
remember
replace
report
represent
require
research
resource
respect
response
responsibility
rest
restaurant
result
return
review
right
ring
rise
risk
road
rock
role
room
round
rule
run
sale
same
save
say
school
science
score
screen
sea
search
season
seat
second
section
security
see
seek
seem
sell
send
sense
separate
series
serious
serve
service
set
several
shall
shape
share
she
shop
short
should
show
side
sign
significant
similar
simple
simply
since
sing
single
sister
sit
site
situation
size
skill
sleep
small
smile

so
social
society
solution
some
somebody
someone
something
sometimes
son
song
soon
sorry
sort
sound
source
space
speak
special
specific
speech
spend
sport
staff
stage
stand
standard
star
start
state
statement
station
stay
step
still
stock
stop
store
story
strategy
street
strike
strong
structure
student
study
stuff
style
subject
success
successful
such
suddenly
suffer
suggest
summer
supply
support
suppose
sure
surprise
system
table
take
talk
target
task
tax
teach
teacher
team
technology
television
tell
tend
term
test
than
thank
that
the

themselves
then
theory
there
therefore
they
thing
think
this
though
through
throughout
throw
thus
time
to
today
together
tomorrow
too
top
total
touch
toward
town
track
trade
traditional
train
travel
treat
treatment
tree
trip
trouble
true
try
turn
type
unclear
under
understand
unit
university
until
up
upon
use
useful
usually
value
variety
various
version
very
view
village
visit
voice
vote
wait
walk
wall
want
war
watch
water
way
we
wear
week
weekend
weight
welcome
well
what
whatever
when
where
whether
which

while	with	would
white	within	write
who	without	wrong
whole	woman	yeah
why	wonder	year
wide	word	yes
wife	work	yesterday
will	worker	yet
win	world	you
window	worry	young
wish	worth	yourself

APPENDIX VI: HEADWORDS OF THE SECOND 1,000 MOST FREQUENT ENGLISH WORD LIST FOUND IN HEADWAY PRE-INTERMEDIATE

abandon
abroad
absence
absolutely
abuse
academic
accident
accommodation
accompany
accuse
achievement
acknowledge
acquire
active
actor
actual
ad
additional
administration
admit
adopt
advertise
advertisement
advice
advise
affair
afford
afraid
aid
aircraft
alive
alternative
amaze
analyze
ancient
angry
announce
annual
anybody
anywhere
apart
apparently
appeal
appearance
appoint
appointment
appreciate
appropriate
approve
approximately
arise
army
arrange
arrangement
arrest
artist
aside
assess
assessment
asset
assistant
association
assumption
atmosphere
attach
attitude
attract
attractive
audience
award
aware
background
bag
ball
ban
band
basically
battle

beach
beat
beauty
bedroom
belief
belong
besides
bet
bike
bind
bird
birth
blame
block
blood
blow
boat
bomb
bond
bone
border
bore
borrow
boss
bother
bottle
bottom
brain
branch
brand
breakfast
breast
breath
breathe
bridge
brief
bright
broad
brown
budget
burn
busy
calculate
calm
camera
camp
cancer
capable
capacity
capture
careful
carefully
cash
cast
cat
category
celebrate
chain
chair
chairman
channel
chapter
characteristic
chart
cheap
chemical
chief
childhood
chip
circle
circumstance
citizen
civil
classic
climate
climb
closely
coach
coast
coat
code

coffee
collapse
colleague
collect
collection
column
combination
combine
comfort
comfortable
command
commercial
commission
commit
commitment
committee
communicate
comparison
compete
competitive
complain
complaint
complex
complicate
component
concentrate
concentration
concept
concert
conclude
conduct
confidence
confirm
conflict
confuse
connect
connection
consequence
conservative
considerable
consideration
consist
consistent
constant
construct
construction
consumer
content
context
contrast
contribute
contribution
convention
convince
cook
cool
cope
copy
core
corner
corporate
corporation
correct
council
count
county
crash
creation
crime
criminal
crisis
criterion
critic
critical
criticism
crop
crowd
crucial
cry
cultural
cup

currently
cycle
dad
danger
dangerous
dear
debt
decade
declare
decline
decrease
defeat
defend
defense
definitely
definition
delay
delight
deliver
delivery
democracy
demonstrate
deny
derive
description
deserve
desire
desk
destroy
device
diet
differ
dinner
directly
disappear
disappoint
discipline
discount
dish
disk
display
distance
distinguish
distribute
district
divide
division
domestic
double
dozen
draft
drama
dramatic
driver
dry
duty
ear
earth
east
edge
editor
educate
educational
effective
effectively
efficient
egg
elect
electronic
element
elsewhere
email
emerge
emergency
emotion
emotional
emphasize
employ
employer
employment
empty

enable
encounter
enemy
engage
engine
engineer
enhance
enormous
ensure
entire
entirely
entry
environmental
equal
equally
equipment
error
escape
essential
estate
eventually
everywhere
examination
examine
excellent
exception
exchange
excite
excuse
executive
exhibit
exhibition
existence
expand
expectation
expense
expensive
experiment
expert
explanation
explore
export
expose
expression
extent
extra
extreme
extremely
factory
failure
fair
fairly
faith
familiar
famous
fan
farm
farmer
fashion
fat
fault
favor
federal
fee
feed
fellow
finance
finger
flat
flight
flow
flower
folk
football
forest
formal
forth
frame
freedom
freeze
frequently

fresh
friendly
fruit
fuel
fully
fundamental
funny
furthermore
gap
gas
gather
gene
generate
gentleman
gift
glad
glass
global
gold
grade
graduate
grant
grateful
gray
guarantee
guard
guest
guide
gun
habit
hall
handle
hang
hardly
hat
hate
healthy
heat
heavy
hell
helpful
hi
hide
highlight
highly
hill
hire
historical
hole
honor
horse
host
household
hunt
hurt
ice
ideal
identity
ignore
ill
illness
illustrate
immediate
implement
implication
imply
import
importance
impose
impossible
impression
improvement
incident
increasingly
independent
index
industrial
inform
initial
initiative
injury

insist
install
institution
instruction
instrument
insurance
integrate
intellectual
intelligence
intend
intention
interaction
internal
interpretation
introduction
invest
investigate
investigation
investor
invite
island
joint
joke
journalist
journey
judgment
jump
justice
kick
king
kiss
kitchen
knock
label
labor
lady
lake
landscape
largely
latter
launch
lawyer
lay
leadership
lean
left
leg
legislation
length
lesson
liberal
library
license
lift
literature
load
loan
locate
location
lock
lovely
luck
lucky
lunch
mail
mainly
majority
manner
manufacture
manufacturer
map
marriage
mass
massive
master
mathematics
maximum
meal
meanwhile
meat
mechanism

medical
medicine
mental
merely
metal
milk
minimum
minister
minor
minority
mission
mistake
mix
mobile
monitor
moral
moreover
mostly
motion
mountain
mouse
mouth
multiple
murder
museum
musical
narrow
native
nearby
necessarily
neck
negative
negotiate
negotiation
neighbor
neighborhood
neither
net
nevertheless
nobody
noise
none
normally
north
northern
notion
novel
nuclear
nurse
objective
observation
observe
obvious
occasion
occupy
odd
online
onto
oppose
opposite
opposition
ordinary
origin
originally
otherwise
ought
ourselves
outcome
outline
output
overall
owner
pack
package
pain
pair
panel
participant
participate
partly
partner

passage
passenger
path
payment
peace
peak
percentage
perfect
perfectly
permanent
permit
personality
personally
perspective
phase
photo
photograph
phrase
physical
pilot
plain
plane
plastic
plate
pleasure
plenty
plot
plus
pocket
politician
politics
poll
pool
pop
possibly
powerful
practical
predict
premise
presence
presentation
preserve
presidential
prevent
previously
primarily
primary
prime
principal
principle
print
prior
priority
prison
procedure
proceed
producer
professor
progress
promote
proper
properly
proportion
prospect
protection
protein
protest
proud
provision
pursue
qualify
quick
quiet
quote
radio
rain
rank
rapidly
rare
reaction
reader

reality
reasonable
recall
recommend
recommendation
recover
reduction
reform
refuse
regional
register
regular
regulation
reject
relative
relatively
relax
relevant
relief
religion
religious
rely
remark
remind
remove
rent
repair
repeat
reply
reporter
representative
reputation
request
requirement
researcher
reserve
resident
resolution
resolve
respond
responsible
retain
retire
reveal
revenue
revolution
reward
rich
ride
river
roll
root
route
row
rural
rush
sad
safe
safety
salary
sample
satisfy
scale
scene
schedule
scheme
scientific
scientist
secondly
secret
secretary
sector
secure
seed
select
selection
self
senior
sentence
sequence
seriously

session
settle
settlement
severe
sex
sexual
shake
sharp
sheet
shift
ship
shirt
shock
shoe
shoot
shot
shoulder
shout
shut
sick
sight
signal
significantly
silence
skin
sky
slide
slightly
slip
slow
slowly
smell
smoke
snow
soft
software
soldier
solve
somehow
somewhat
somewhere
soul
south
southern
speaker
specifically
speed
spell
spirit
split
spot
spread
spring
square
stable
status
steal
stick
stone
storm
straight
strange
strength
stress
stretch
strongly
struggle
studio
substantial
succeed
sudden
sufficient
suggestion
suit
suitable
sum
sun
surely
surface
surround
survey

survive
suspect
sweet
swim
swing
switch
talent
tall
tape
taste
taxi
tea
tear
technical
technique
telephone
temperature
temporary
tennis
terrible
text
theater
theme
thick
thin
threat
threaten
ticket
tie
till
tiny
tip
tire
title
tone
tonight
tool
topic
totally
tough
tour
tourist
toy
tradition
traffic
transfer
transport
trend
trial
troop
truly
trust
truth
twice
typical
unable
unemployment
unfortunately
union
unique
unless
unlike
unusual
update
upper
urban
urge
user
usual
valuable
variable
vary
vast
vehicle
via
victim
victory
video
violence
vision
visitor

volume
volunteer
voter
wage
wake
warm
warn
wash
waste
wave
weak
weapon
weather
wed
west
wheel
whereas
widely
wild
wind
wine
winner
winter
witness
wonderful
wood
writer
yard
yellow
yield
youth

APPENDIX VII: HEADWORDS OF THE ACADEMIC WORD LIST FOUND IN HEADWAY PRE-INTERMEDIATE

abdominal
absorb
absorption
accelerate
acceleration
accent
accumulate
accumulation
accuracy
accurately
acid
acidic
activate
actively
acute
adaptation
adaptive
adjacent
admission
adolescent
adverse
aerosol
aesthetic
affirm
afterward
aggregate
agriculture
airplane
algebra
algorithm
alien
alliance
allocate
allocation
altitude
aluminum
amino
amongst
amplitude
analogy
ancestor
anthropology
anti
antibiotic
antibody
antiquity
appendix
applause
apple
approximate
approximation
arbitrary
archaeology
architect
array
arrow
articulate
artifact
artificial
artistic
artwork
aspect
assembly
assert
assignment
athletic
atom
atomic
auction
audio
audit
authority
autonomy
availability
axiom

axis

backward

bacteria

bacterial

bang

bargain

barrel

basin

bat

beam

behavioral

bilingual

binary

biodiversity

biologist

biology

bizarre

blank

bleed

bodily

bonus

bound

bracket

breakdown

broadly

bubble

bucket

bulk

bullet

bundle

calcium

calculation

calculator

calculus

campus

candidate

capillary

capitalism

capitalist

carrier

cattle

censor

characterization

cheat

cheer

chemistry

chemotherapy

chess

chloride

chromosome

chronic

chunk

cinema

circa

circulate

circulation

civilization

clarify

classification

classify

classroom

clay

clever

click

client

clinic

clip

clone

closure

clue

coefficient

cognitive

coherent

collective

colonial

colony

comma

commentary

commodity

commonly

communicative

communist
compact
comparable
comparative
compensate
competence
competent
complement
complication
comply
composer
composite
comprehension
computation
conceive
conception
conceptual
conditional
conduction
cone
conference
configuration
confine
confound
congruent
connector
connotation
conscious
consciousness
consensus
consent
conservation
conserve
consonant
constitution
constrain
consultation
consumption
container
continent
continuity
contour
contradict
contradiction
contradictory
contrary
controversy
converge
convergence
coordinate
coordination
cord
coronary
corpus
correction
correctly
correlate
correlation
correspondence
corruption
cortex
credibility
criteria
critically
critique
crude
crystal
cue
cure
curriculum
cyclic
cylinder
damp
deadline
decay
deceive
defect
deficiency
definite
deflection
degrade

deliberately
delta
demonstrator
denominator
denote
dense
dependence
depict
derivative
descendent
descriptor
destination
detection
developmental
deviation
diagnose
diagnosis
diagnostic
diagram
dialect
diameter
diary
dictate
dictionary
differential
differentiate
differentiation
diffusion
dilemma
dilute
dimensional
dioxide
directive
disability
disadvantage
discharge
discourse
discrete
discrimination
displacement
dissection
dissertation
dissolve
distribution
disturbance
diverse
domain
dominance
dominant
domination
donor
dose
drain
drift
duration
dye
dynamic
ecological
ecology
economically
economist
effectiveness
elaborate
elastic
elasticity
electron
elementary
elevate
elevation
elimination
elite
embed
emergence
emission
emit
emperor
empirical
encode
enforcement
enormously
entity

entrant
enzyme
epidemic
epidemiology
equality
equilibrium
equivalence
erase
essence
essentially
estimation
ethical
ethics
evident
evolutionary
ex
execute
execution
exit
expertise
explicit
explicitly
exploit
exponential
fabric
facet
facilitate
factorial
faculty
fatigue
feedback
fertility
fetal
fever
fiber
fin
finite
flesh
flexibility
flip
fluid
flux
footnote
formally
formulation
forum
fossil
founds
fraction
fracture
freely
friction
fringe
fundamentally
fungus
fusion
gauge
generalization
generalize
genetically
genetics
ghost
globalization
goat
goodness
goods
gradient
gram
grammatical
graph
grasp
gravity
grid
gross
gut
habitat
halfway
handout
harvest
hawk
headquarter

hedge
helix
hepatitis
herbicide
hierarchy
hip
historically
homework
horizon
horizontal
hormone
hydrogen
identical
identification
ideology
illusion
immune
impact
implicit
importantly
impulse
incidence
incline
inclusion
incredible
incredibly
incumbent
independently
indicator
indifference
indigenous
indirect
individually
induce
induction
industrialization
industrialize
inequality
inevitably
infect
infectious
inference
inferior
infinite
infinity
influential
informal
inhibit
inhibition
initiate
initiation
inject
injection
innate
insect
insert
instability
instinct
integral
integration
intensity
intensive
interact
interestingly
interface
interfere
intermediate
interrupt
interval
intervene
interviewer
invade
invasion
inversion
invert
ion
irrelevant
irrigation
jazz
junior
justification

kidney
kilometer
lab
lateral
layout
leaf
lecturer
legend
legitimate
leisure
lever
lexical
liable
lifestyle
lifetime
likelihood
likewise
limb
linear
linguistic
linguistics
liter
liver
locally
locus
logical
longitudinal
loop
lump
lung
machinery
magnetic
magnitude
mall
manipulate
manipulation
manual
manuscript
marble
marginal
marker
marrow
maternal
mathematical
matrix
maximize
meaningful
mechanic
mechanical
media
mediate
membrane
memorize
mentor
mercury
merge
messenger
metabolism
metaphor
methodology
micro
mid
migrate
migration
millimeter
mineral
minimal
minimize
minus
missile
mobility
modification
mole
molecular
molecule
momentum
monetary
monkey
monopoly
morality
morphological

morphology
mortality
motif
motive
multi
multinational
multiply
mutation
myth
naked
namely
nasty
nationalism
neat
necessity
neo
nest
neural
neuron
neutral
nicely
niche
nitrogen
noble
node
noisy
nominal
non
nonetheless
nonlinear
norm
notation
novice
nucleus
null
numerical
nutrient
objection
obscure
observer
obtain
occupation
occurrence
offspring
onset
onwards
optical
optimal
optimum
oral
organ
organism
orient
orientation
oscillation
outer
outlet
overhead
overlap
overview
oxidize
oxygen
painful
par
paradigm
paradox
parameter
parcel
pardon
parental
parenthesis
partial
partially
particle
partition
pathway
peasant
periodic
peripheral
pest
pesticide

phenomenal
philosopher
philosophical
phonological
phosphate
photographic
physically
physician
physics
physiological
pi
planner
plantation
plausible
plug
plural
polar
pole
politically
portfolio
portray
portrayal
positively
poster
postgraduate
potassium
powder
practitioner
pragmatic
pre
precede
precipitate
precipitation
predator
prediction
predominantly
prejudice
preliminary
presume
prevalence
prey
primer
primitive
probe
problematic
processor
productive
productivity
profound
progression
progressive
projection
prominent
pronounce
proposition
protocol
proton
psychiatric
psychologist
psychology
publish
pulse
punch
punish
punishment
purely
puzzle
qualitative
quantitative
quantum
questionnaire
quiz
quotation
rack
radar
radiation
radius
rainfall
randomize
randomly
rational

rationality
ray
reactive
reactor
readily
realism
realistic
realm
receptor
recipe
reconstruct
regime
regression
rehabilitation
reinforce
rejection
reliability
render
repertoire
replacement
replicate
replication
reproduce
reproduction
republic
resemble
reservoir
residual
residue
resistant
respiratory
retrieve
revolutionary
rewrite
rhetoric
rhythm
ridiculous
ritual
robot
rope
rotate
rotation
rub
ruler
scatter
scenario
scholarship
scripture
scroll
secrete
sediment
selective
semantic
semester
semi
seminar
sensation
sensible
sensitivity
sensory
separately
separation
sexuality
shallow
shortly
shuttle
similarity
simplify
simulate
simulation
simultaneously
sin
singular
sketch
skip
slab
slash
slavery
slot
snake
sneeze

sniff
socialize
socially
sodium
solar
soluble
solute
solvent
sometime
sophisticate
span
spatial
specialty
specification
specimen
spectrum
sperm
sphere
splice
sponsorship
spontaneous
spray
stabilize
stack
stadium
stance
standardize
static
statistical
statistically
statistics
stereotype
stimulus
straightforward
strand
strategic
strictly
stripe
sub
subjective
subset
substitution
substrate
subtle
subtract
sufficiently
suicide
super
superior
supposedly
surgeon
surgical
surplus
susceptible
sustainable
swap
swell
sword
syllable
syllabus
symbolic
syndrome
syntactic
syntax
synthesis
synthetic
systematic
tech
technically
temporal
tempt
tense
terminal
terminology
terribly
textbook
theorem
theorist
thereby
thermal
thesis

thickness
thread
threshold
thumb
tolerance
ton
toxic
tract
traditionally
trait
trajectory
trans
transaction
transcribe
transcription
transformation
translation
transmission
transmit
transparency
transplant
trauma
treaty
tremendous
triangle
tribe
tricky
trivial
tropical
tumor
turbulent
tutor
ultimate
undergraduate
undermine
underneath
unemployed
unify
unintelligible
unity
unstable
uplift
upward
urine
usage
utility
utilize
utterance
vague
valid
validity
valve
variability
variance
variant
vector
vegetation
vein
velocity
verbal
vertical
viable
virtue
vitamin
vocabulary
volition
vowel
watershed
wavelength
weave
wheat
whereby
whichever
whoever
widespread
wisdom
workshop
yeast

APPENDIX VIII: OFF-LIST WORDS OF HEADWAY PRE-INTERMEDIATE

actresses
adore
agers
ambulance
annoyed
babes
beard
bitterly
bolt
born
bullied
cab
carp
carriage
celebrities
chaos
chatty
chef
commemorative
cookies
crept
crisps
cursed
cyber
delicious
doll
engagement
etc
euro
fairy
fierce
fighters
flavoured
foreigners
gymnastics
hacker
hackers
hq
hugely
hunter
intelligent
jumper
jungle
km
knight
licked
lightning
lollies
lraq
martial
memorabilia
memorabİ
mudd
mysterious
nicknamed
notorious
oh
organizer
patent
pillow
porch
prince
protestors
queen
reminders
renewed
risky
robe
ross
slum
soccer
sporty
stunt
stunts
stuntwoman
stuntwomen
teen
terrified
thumping
thunder
tiger
trampolining
trophies
unexpected
urgently
veggie
viewer
ward
wink
woolly

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı Soyadı : Mürüvvet SÖĞÜT

Doğum Yeri :

Eğitim Durumu

Lisans Öğrenimi : 2011 - Orta Doğu Teknik Üniversitesi, İngilizce Öğretmenliği

Yüksek Lisans Öğrenimi: 2022 - Akdeniz Üniversitesi, İngiliz Dili Eğitimi

Bildiği Yabancı Diller : İngilizce, Almanca

İş Deneyimi

2011- 2012 Öğretim Görevlisi, Ankara Üniversitesi, Yabancı Diller Yüksekokulu

2012- 2014 Öğretim Görevlisi, İstanbul Kemerburgaz Üniversitesi, Yabancı Diller Yüksekokulu

2015-2016 Öğretim Görevlisi, Gazi Üniversitesi, Yabancı Diller Yüksekokulu

2016-2020 Öğretim Görevlisi, Antalya Bilim Üniversitesi, Yabancı Diller Yüksekokulu

2020 - Serbest Çalışan İngilizce Öğretmeni

İletişim

E-Posta Adresi :

Tarih :

Doküman Görüntüleyici

Turnitin Orijinallik Raporu

İşleme kondu: 06-Haz-2022 14:42 +03
NUMARA: 1851502942
Kelime Sayısı: 17713
Gönderildi: 1

M S SADE Muruvvet Sogut
tarafından

Benzerlik Endeksi	Kaynağa göre Benzerlik	
%6	Internet Sources:	%5
	Yayınlar:	%3
	Öğrenci Ödevleri:	%2

alıntıları çıkar bibliyografyayı çıkar küçük eşleşmeleri çıkar mod:
raporu hızlı görüntüle (klasik) Change mode yazdır yenile İndir

<1% match (10-Şub-2022 tarihli internet)
https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/339786/yokAcikBilim_10072889.pdf?isAllowed=y&sequence=-1

<1% match (09-Şub-2022 tarihli internet)
https://acikbilim.yok.gov.tr/bitstream/handle/20.500.12812/249605/yokAcikBilim_10258369.pdf?isAllowed=y&sequence=-1

<1% match (04-Oca-2021 tarihli internet)
https://dialnet.unirioja.es/descarga/articulo/4125131.pdf

<1% match (07-Eki-2021 tarihli internet)
https://www.kafkas.edu.tr/dosyalar/sobedergi/file/22/02.pdf

<1% match ()
López-Jiménez, María Dolores. "THE TREATMENT OF VOCABULARY IN EFL TEXTBOOKS", Universidad de Sevilla // UNED, 2017

<1% match (29-Ağu-2017 tarihli internet)
http://uir.unisa.ac.za

<1% match (28-Nis-2015 tarihli öğrenci ödevleri)
Submitted to University of Wales Swansea on 2015-04-28

<1% match (14-Ara-2016 tarihli öğrenci ödevleri)
Submitted to University of Wales Swansea on 2016-12-14

<1% match (02-Oca-2020 tarihli öğrenci ödevleri)
Submitted to University of Wales Swansea on 2020-01-02

<1% match (09-Eyl-2021 tarihli internet)
http://dspace.univ-setif2.dz

<1% match (12-Kas-2016 tarihli internet)
https://www.coursehero.com/file/12398567/67-1986-Laufer-IRAL/

<1% match (yayınlar)
Cathrine Norberg, Marie Nordlund. "A Corpus-based Study of Lexis in L2 English Textbooks", Journal of Language Teaching and Research, 2018