T. C.
ULUDAĞ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
YABANCI DİLLER EĞİTİMİ ANABİLİM DALI
İNGİLİZ DİLİ EĞİTİMİ BİLİM DALI

VOCABULARY LOAD OF THE ENGLISH LANGUAGE COURSEBOOK FOR GRADE 10

(YÜKSEK LİSANS TEZİ)

MELİKE AGAN

BURSA 2008

U.Ü.S.B.E.
YABANCI DİLLER E.
ANA BİLİM DALI
İNG. DİLİ EĞİTİMİ
BİLİM DALI

VOCABULARY LOAD OF THE ENGLISH LANGUAGE COURSEBOOK FOR GRADE 10
(YÜKSEK LİSANS TEZİ)

Melike
AGAN

BURSA
2008

T. C.
ULUDAĞ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
YABANCI DİLLER EĞİTİMİ ANABİLİM DALI
İNGİLİZ DİLİ EĞİTİMİ BİLİM DALI

(YÜKSEK LİSANS TEZİ)

MELİKE AGAN

BURSA 2008

T. C.
ULUDAĞ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
YABANCI DİLLER EĞİTİMİ ANABİLİM DALI
İNGİLİZ DİLİ EĞİTİMİ BİLİM DALI

VOCABULARY LOAD OF THE ENGLISH LANGUAGE COURSEBOOK FOR GRADE 10

(YÜKSEK LİSANS TEZİ)

Melike AGAN

Danışman

Yard. Doç. Dr. Meral ÖZTÜRK

T. C.
ULUDAĞ ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜ'NE

Uludağ Üniversitesi Yabancı Diller Anabilim Dalı, İngiliz Dili Eğitimi Bilim Dalı'nda U2004502 numaralı Melike AGAN'ın hazırladığı "Lise 2. Sınıf İngilizce Ders Kitabında Yer Alan Sözcüklerin Kullanım Sıklığı Bakımından İncelenmesi" konulu Yüksek Lisans Tezi ile ilgili tez savunma sınavı, 21/ 03/ 2008 günü 10.00 – 11.00 saatleri arasında yapılmış, sorulan sorulara alınan cevaplar sonunda adayın tezinin başarılı olduğuna oybirliği ile karar verilmiştir.

Üye (Tez Danışmanı ve Sınav Komisyonu Başkanı)

Üye

Üye

...../...../ 20.....

ÖZET

LISE 2. SINIF İNGILIZCE DERS KITABINDA YER ALAN SÖZCÜKLERIN KULLANIM SIKLIĞI BAKIMINDAN İNCELENMESİ

AGAN, Melike

Yüksek Lisans Tezi, Yabancı Diller Eğitimi Ana Bilim Dalı

Tez Danışmanı: Yrd. Doç. Dr. Meral ÖZTÜRK

Şubat 2008, 78 sayfa

Bu çalışma Türkiye'de tüm devlet okullarının 10. sınıf öğrencilerine İngilizce öğretiminde kullanılması zorunlu olan An English Course For Turks Advanced 1 kitabının kelime yükü analizini içermektedir. Bu çalışmanın başlıca amacı, bu kitabın kelime yükü ile bu kitabı İngilizce derslerinde kullanmak zorunda olan öğrencilerin kelime bilgilerinin karşılaştırılmasıdır. Bu tez konusu, kitabın anlaşılabilirliği ile ilgili, uygulamada ortaya çıkan bir problemden yola çıkılarak oluşturulmuştur, çünkü bir İngilizce öğretimi ders kitabı ancak öğrenciler tarafından anlaşılabilir olmasıyla yabancı dilin kazanılmasında etkili olabilir. Kitabın anlaşılabilirliği hakkında çıkarımlar yapmak için, kitabın kelime yükü ve öğrencilerin kelime bilgisi düzeyleri karşılaştırılmalıdır. Bu karşılaştırmayı yaparak kitabın kelime yükünün öğrencilere uygun olup olmadığını ortaya çıkarmak için, kitabın yaklaşık olarak yarısı tarayıcıdan geçirilerek, sayfaları bilgisayar ortamında kullanılabilecek elektronik belgeler haline getirilmiştir. Tarama işleminden sonra kitabın kelime yükü Range32 programıyla analiz edilmiş ve öğrencilere standart bir kelime düzeyi saptama testi uygulanmıştır. Kitabın kelime yükünün öğrencilerin kelime düzeylerine uygunluğu konusunda yorumlarda bulunabilmek için kelime yükü analizi sonuçları ile öğrencilere uygulanan kelime düzeyi testlerinin sonuçları karşılaştırılmış ve öğrencilerin bu kitabın kullanımı esnasında yaşayacakları zorluk düzeyi saptanmıştır.

ABSTRACT

VOCABULARY LOAD OF THE ENGLISH LANGUAGE COURSEBOOK FOR GRADE 10

AGAN, Melike

M.A. Thesis, Department of English Language Teaching

Supervisor: Assist. Prof. Dr. Meral ÖZTÜRK

February, 2008, 78 pages

This study covers vocabulary load analysis of An English Course For Turks Advanced 1, which is compulsorily used in all state high schools in Turkey for teaching English to 10th grade students. The main aim of this study is to compare vocabulary load of the course book and vocabulary sizes of the students who have to use the book as an EFL textbook. This thesis was constructed from a practical problem concerning the comprehensibility of the book, because an EFL course book can be efficient for acquisition of a foreign language if it is understandable for learners. To make inferences about comprehensibility of the course book, vocabulary load of the book and vocabulary sizes of students are to be compared. By making this comparison to reveal whether the vocabulary level of this book is appropriate for the intended students, about half of the book was scanned and pages of the book were turned into electronic documents that can be used at the computer environment. After scanning, vocabulary load analysis of the book was done using a word analysis program Range32 and the intended students were implemented a standardized word level test. To comment on suitability of vocabulary load of the book with vocabulary sizes of students, results of the vocabulary load analysis and word level tests which were implemented to the students were compared and the level of difficulty that students can experience respecting vocabulary level while using the book was established.

PREFACE

In today's world, after the effects of globalization on all countries in the world, knowing a foreign language has become a necessity instead of being a privilege. Learning a new language broadens our cultural and intellectual horizons and provides us to have a better understanding of our own language. It is also an encouragement for our acceptance of people who are different from us. Learning foreign languages can be a personal satisfaction by helping to find a better job and increase your income, as well as opening you the gates of an endless knowledge repertory.

Among the world languages, especially English, gained a great importance because of its being used widely in the whole world as an international language of communication. Because of the wide use and currency it is a compulsory school subject from primary schools to universities in Turkey. Students are introduced to English by two classes at the 4th grade of primary education, then class time increases at secondary schools and high schools. Almost each year, till graduation from high school, Turkish students learn English but what proportion of them can speak English and understand what they hear? The answer of this question is known by everyone. Almost none….. So what can be the reason for Turkish students' being unsuccessful to gain fluency in English? Many reasons can be listed in relation to this question concerning the English teaching system in our state high schools but to me, the most striking reason is the outdated teaching materials (course book) which were prepared decades ago without considering systematicity, and levels and background knowledge of the intended students in terms of vocabulary load. The books which are made compulsory to use in all state high schools by The Ministry of National Education, have a fairly high vocabulary burden, which causes that intended students can not cope with the unknown words in the textbook. While they try to acquire language structures and skills, they have to struggle with the long unknown wordlists which makes language learning process a torture for them instead of being enjoyable and motivating. Starting off these ideas, I examined in this study, whether the vocabulary of the 10th grade EFL text book, An English Course For Turks Advanced 1, is appropriate for the intended students' vocabulary sizes. From the results, I draw conclusions about the comprehensibility and manageability of the course book for the students who have to learn English from this course book.

I would like to express my gratitude to all those who gave me the possibility to complete this thesis by providing me with useful and helpful assistance. Firstly, I would like to deeply thank my honorable advisor Assist. Prof. Meral ÖZTÜRK for her interest, patience, encouragement and stimulating suggestions. Without her care and consideration, this study would likely not have matured. Second, I would like to thank my beloved husband Mehmet TARAKÇI for his sincere encouragement, care and support during my study. His patient love enabled me to complete this work. Third, I would like to express my gratitude to my father and mother for their endless encouragement and support. As in all my endeavors, they supported me with all their hearts. February 2008 Melike AGAN

CONTENTS

Page

LIST OF TABLES

Page

ABBREVIATIONS

EFL: English as a foreign language

Vol: Volume

i.e : id est

e.g exempli gratia

e.t.c. : adecetera

Sts : Students

SECTION ONE
INTRODUCTION

1.1 INTRODUCTION

Comprehensibility is a must in order for every kind of learning to take place. Two factors can be mentioned in relation to the comprehensibility of a text in a foreign or second language; one of them is the difficulty of the grammatical structures in the input and the other is the vocabulary load in the input. That's the reason why it is necessary to use a course book for language teaching which contains comprehensible input that Krashen (1985) called "i+1". Krashen's hypothesis was developed with grammatical acquisition in mind. Comprehensible input is seen as a major causative factor in L2 acquisition by Krashen. His Input Hypothesis is one of the most influential theoretical positions in the literature. According to Stephen Krashen, L2 acquisition takes place when a learner understands input that contains grammatical forms that are at i+1 (i.e. a little more advanced than the current state of learner's interlanguage). Krashen maintains that exposure to comprehensible input is both necessary and sufficient for second language learning to take place. The hypothesis states that "Humans acquire language in only one way; by understanding messages,or by receiving comprehensible input." In other words, understanding input that contains structures a little bit beyond their current level of competence brings progress for learners along the natural order. We move from "i" our current level, to i+1, the next level along the natural order, by understanding input containing i+1. If input is understood, and there is enough of it, the necessary grammar is automatically provided. But, If the context level is not accessible according to the current level of the students,with the help of simplification, contextual and extraliguistic clues, the level of the input should be made appropriate for the current level of students.

Krashen proposed three stages in turning input into intake:

1.Understanding an L2 i+1 form,

2.Noticing a gap between the L2 i+1 form and the interlanguage rule which the learner currently controls,

3. The reappearance of the i+1 form with minimal frequency.

Shortly, Krashen claims that L2 acquisition depends on comprehensible input. When the second factor, vocabulary load, is examined, it is necessary that students know at least 95% (Laufer,1989) of the words in a text to provide comprehensible input for effective language learning, and for guessing from context students need to know at least 98% of the words in the input (Nation 2001). Hu and Nation did a research on the text coverages. Text coverage is a term which can be related to i+1 input. We determine i+1 input related with vocabulary load by text coverage (How many of the running words does a student know in a text?). They compared the effect of four text coverages on reading comprehension of a fiction text. In the 100% text coverage, no words were unknown. In the 95% text coverage version, one word was unknown, on average, in every twenty. In the 90% text coverage level one word in every ten running words was unknown, and in the 80% text coverage there was one unknown word in every five running words. The results of the research of Hu and Nation indicate that, there is a predictable relationship between text coverage and comprehension; increase in the number of the known words improved comprehension. Some learners attained adequate comprehension at the 95% coverage level, but most did not. At the 90% coverage level, a smaller number gained adequate comprehension and the 80% level none of the students did. Hu and Nation concluded that for largely unassisted reading for pleasure, learners would need to know around 98% of running words in the text. With this coverage for almost all learners, there will be a chance to gain adequate comprehension. Instead of adequate comprehension, 95% coverage is needed for a standard of a minimally acceptable comprehension.

To summarize, If text coverage is related to the strands of learning from meaning-focused input and fluency development, then learners would need to have 95% coverage for learning from meaning-focused input, and 98-100% coverage for fluency development. This means that learners need to have simplified learning material of various levels in order to learn from meaning-focused input and to develop fluency in reading if they are to learn from these strands at all stages of their second language development.

As English teachers, we expect that a course book fulfill our teaching goals. To provide this, students have to know at least 95% of the words in the course book. ”An English Course For Turks Intermediate” is compulsory to be used by all English teachers teaching 10th grade students at all state high schools and the common opinion of teachers using this book is that the vocabulary load of the book is not suited to the students’ current vocabulary size.

From the comments of teachers who used the book as a course book, it can be understood that students need to prepare long wordlists to be able to follow English lessons. And It is observed that, students can’t focus on gaining fluency in language and acquiring structures aimed to teach in every unit while they struggle with the high number of unknown words. This evokes the idea that vocabulary load of the book makes it difficult for students to learn the aimed subjects of the lessons and, because of the difficulty in comprehending the book, they lose their motivation in learning English. The more unknown words they list, the less they learn new words. ”An English Course For Turks” has to be a comprehensible course book which was prepared considering vocabulary levels of the intended students, since all teachers teaching 10th grade students at all state high schools are forced to use this book. All these practical problems that English teachers experience while using this book, constructed the basis of this study. This research aims to investigate vocabulary load of the course book, An English Course For Turks (Advanced 1) for Grade 10, and to see whether the difficulty level of the words used in the book is suitable for students’ vocabulary level to benefit from this book in learning English efficiently.

SECTION TWO
LITERATURE REVIEW

In this chapter terms related with the vocabulary count will be explained. Then the classification of vocabulary into high and low frequency words, academic words and technical words will be discussed. This chapter ends with the vocabulary coverage studies.

2.1 Vocabulary Count

For research and teaching purposes, words are to be counted. Words are counted as *tokens, types, lemmas or word families.*

The term *token* is used if every word form in a spoken or written text is counted, even if the same word form occurs more than once. The number of tokens in the sentence "It is not easy to say it correctly." is eight even though the same word (i.e. it) is used twice. This type of counting words is used when the number of words on a page or in a line are counted or the length of a book is told.

The term *type* is used when each word is counted once, even if the word occurs more than once in a sentence or a text. If the types of the sentence "It is not easy to say it correctly." are counted, the number will be seven, since there are seven different words in the sentence. This type of counting is used when the subject matter is "How large a person's vocabulary is" or "The number of words needed to know to read a book".

When the headword and some of its inflected and reduced forms are counted as a single word in a text, "*Lemmas*" are counted. That means, plural, third person singular, past tense, past participle, -ing, comparative, superlative inflections of words are counted under the single group of basewords. The same form is counted as different lemmas when the form is used as a different part of speech (such as *walk* as a noun or *walk* as a verb) as different lemmas.

Lemmas are used as the unit of counting because of the idea of learning burden. The amount of effort required to learn a word is explained as learning burden. When students learn the baseword and how to use the inflectional system, the learning burden on the students reduces. If a learner knows *mend*, it is easier for him/her to understand "mends". Counting lemmas is a step in trying to explain the learning burden on the students.

Nation defines word families as a baseword, its inflected forms and derived forms (Nation,2001). In other words, word family is a stem plus all closely related affixed forms as defined by Level 6 of Bauer and Nation's (1993) scale. The Level 6 definition of affix includes all inflections and the most frequent, productive and regular prefixes and suffixes. (page 255.) It includes only affixes that can be added to stems that can stand as free forms (For example, *specify* and *special* are not in the same word family because *spec* is not a free form.)

From the point of view of reading, 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. So, *sing, sings, sang*, and sing*ing* may all be members of the same word family for a learner with a command of the inflectional suffixes of English. Also the word family of the headword *sing* includes the derivations such as *singer.* The size of a learner's word family increases, as a his/her affixation knowledge develops. The important principle behind the idea of a word family is that once the base word or even a derived word is known, the recognition of other members of the family requires little or no extra effort. Clearly, the meaning of the base in the derived word must be closely related to the meaning of the base when it stands alone or occurs in other derived forms, for example, *hard* and *hardly* would not be members of the same word family.(Bauer and Nation,1993, Word Families) In the following example, how new members of a word family require little learning effort can be seen. Many readers will not be familiar with the lexeme *marmelize* which is a British dialected word, that means *"to beat up"*. After having met the word form marmelize, such readers will have no difficulty in understanding the sentence:

She marmelized him.

The meaning of *marmelized* is understood, although they have never met before. In the same way, the sentence:

He fully deserved her marmelization.

Is easily understood even though it contains the new word-form *marmelization.*

Counting words as word families provides a more comprehensible evaluation than lemmas. Lemmas are the words in the same category but word families consist of all kinds of uses of the word: noun, adjective verb etc... Nation and Bauer state that If a learner knows a baseword, he/she is thought to know its transparent derivations. Transparent derivations are the derived forms of the words. That is, a learner who knows *solve* can guess the meaning of *solution*. Since comprehending regularly inflected or derived members of a family does not require much more effort by learners if they know the baseword and if they have control of basic word-building processes, the learning burden on the learner's reduces and this facilitates their learning.

To decide what can be included in a word family have significant effects on teaching and learning. These decisions may be used to find answers to some questions that have been asked relating to teaching and learning. The questions include:

What are feasible goals for a second language vocabulary development programme?

What are the contributions of the vocabulary growth of the various sources of vocabulary learning, such as inferring from context, morphological generalization and direct teaching (Whysocki and Jenkins 1987; White, Power and White 1989)?

How much attention should a teacher give to the word-building systems such as affixation and compounding?

Is a particular reading text suitable for a particular learner? (Bauer and Nation, 1993)

In the present research, Michael West's General Service List is used for word count. West developed this list around 2000 high frequency words that would be most useful for a learner of English as a second or foreign language. He called his list "a general service list" because the words it contained would be needed (of service) in a wide range of situations (general), genres and uses. This list is the indispensable basis of all uses of English (Hirsh and Nation,1992).

Table 2.1 and 2.2 show sample word families from the General Service List:

Table 2.1 Word families from the most frequent 1000 words.

A	ACCEPT
AN	ACCEPTABILITY
ABLE	ACCEPTABLE
ABLER	UNACCEPTABLE
ABLEST	ACCEPTANCE
ABLY	ACCEPTING
ABILITIES	ACCEPTED
UNABLE	
INABILITY	
ABOUT	
ABOVE	

Table 2.2 Word families from the most frequent 2000 words:

ABROAD	ACCIDENTAL
ABSENCE	ACCIDENT
ABSENCES	ACCIDENTS
ABSENT	ACCIDENTALLY
	ACCUSTOM
ABSOLUTE	ACCUSTOMS
ABSOLUTELY	ACCUSTOMING
ACCUSE	ACCUSTOMED
ACCUSING	
ACCUSES	

2.2 Types of Vocabulary

In Nation 2001, vocabulary is classified as *high frequency, academic,technical and low frequency words*.

High frequency words are a small group of words which cover a large proportion of running words in spoken and written texts and occur in all kinds of uses of the language. Michael West's General Service List presents the highest frequent 2000 words in English. This list consists of about 165 function words such as *a, some, two*,

because, to. Other words in the list are content words such as nouns, verbs, adjectives and adverbs. Obviously, the high-frequency words have great importance for learning general English, that's why, both teachers and students should spend time on these words, by direct teaching, direct learning, incidental learning and planned meeting.

Academic and *Technical Words* are specialized vocabulary which occur in certain types of texts. There can be special vocabularies for speaking or reading academic texts, for reading newspapers, for reading stories or writing letters. Academic vocabulary mostly occurs in the texts of academic study in English. Academic Word List (Coxhead, 1998) consists of 570 word families which are not among the most frequent 2000 words of English.(See table 2.3) Since the academic words contain formal vocabulary instead of technical words, they are sometimes called sub-technical vocabulary.

Table 2.3 Sample head words from the Academic Word List

abandon	accumulate	aggregate
abstract	acquire	aid
academy	adapt	albeit
access	adequate	allocate
accommodate	adjacent	alter
accompany	adjust	alternative
accurate	administrate	ambiguous
achieve	adult	amend
acknowledge	advocate	analogy
acquire	affect	analyze

Technical words are groups of words that occur in only one subject area. (e.g. cabotage, amortization) .

Low frequency words are very infrequent and they cover only a small proportion of any text. Low frequency words also cover the moderate frequency words that did not manage to get into the high frequency list. Proper names are also among the low frequency words. 4% of the running words in the Brown Corpus consist of proper names. According to the jobs, interests and specializations, technical vocabulary may be low frequency or high frequency words. Some words are low frequency words for

each user of the language. For example eponymous, gibbous, plummet. These kinds of words may mean a rarely expressed idea or there may be other higher frequency words that have the same meanings. They may also be old-fashioned, formal or foreign words. Another possibility is that they belong to a particular dialect or vulgar.

Table 2.4 shows a sample text where 80% of the running words are high frequency words. The academic and low frequency words are shown underlined and the technical words are shown in italics.

Table 2.4 A sample text showing high-low frequency words, technical words and academic words

Sustained-yield management ought to be long term government policy in indigenous forests *zoned* for production. The adoption of such policy would represent a *breakthrough* – the boundary between a *pioneering*, extractive phase and an *era* in which the *timber* industry adjusted

to living with forests in *perpetuity.* A forest sustained is a forest in which harvesting and *mortality* combined do not exceed *regeneration.* Naturally enough, faster-growing forests produce more *timber*, which is why attention would tend to swing from *podocarps* to *beech* forests regardless of the state of the *podocarp* resource. the colonists cannot be blamed for *plunging* in without thought to whether the resource had limits. They brought from *Britain* little experience or understanding of how to maintain forest structure and a *timber* supply for all time. Under *German* management it might have been different here. The *Germans* have practiced the sustained approach since the seventeenth century when they faced a *timber, rimu,* prompts a similar response- no more contraction of the *indigenous* forest and a balancing of yield with *increment* in selected areas.

In order to define a desirable vocabulary size for learners, word families are used to count the words. Although there is not a definite vocabulary size for a successful language user, to continue increasing the vocabulary size is very important for learners. For low frequency, it is not necessary to spend classroom time to directly teach them since words in this group may be learnt in time by using language. Instead learners may be taught strategies to deal with these low frequency words more easily. Guessing from the context, using word parts to remember, using vocabulary cards and dictionary are among the helpful vocabulary learning strategies. For the learners who know the high frequency words, teachers should concentrate on expanding and refining the learners' control of vocabulary they know, instead of spending the limited classroom time by dealing with the low frequency words.

2.3 Vocabulary Coverage Studies

Vocabulary is seen as just one of the numerous elements that contribute to the learners' overall performance in L2. From the results of interviews which are related to how learners deal with and comprehend the texts they read in an English course, Laufer and Sim concluded that the most pressing need of foreign language learners was vocabulary, then subject matter knowledge and then syntactic structures (Nation 2001). There is a growing recognition of the importance of vocabulary development for L2 learners since many of them are severely hampered in reading comprehension and other skills by a simple lack of word knowledge. It is clear that limited vocabulary knowledge is not adequate to meet the demand of students (Read,1988).

If having enough vocabulary knowledge is of a great importance for the students to learn a foreign language in the most efficient way, teachers should deal with the vocabulary sizes of students to select a material that can best suit the level and needs of the students. The reason for this claim is that the level of the material which will be used for teaching English has to be one level beyond the current levels of students. This idea was expressed by Krashen as Comprehensibility Hypothesis. According to Krashen's Comprehensibility (Input) Hypothesis, a language is acquired when a learner understands input that contains grammatical forms that are a little more advanced than the current level of the learner. Students have to be taught by i+1 materials to make the course understandable and efficient. This input is called "Comprehensible Input" and Krashen suggests that learning a language depends on this Comprehensible Input (Ellis,1997). If the level is far below or above their current level, the course will be useless or too complex for the students to cope. Although Krashen's Comprehensibility Hypothesis is related with the grammatical forms, it can be implemented in the vocabulary teaching area under the light of the research on vocabulary.

Control of the reading skill can be a major factor in vocabulary development for both native and non-native speakers. Research on L1 reading shows that vocabulary knowledge and reading comprehension are very closely related to each other. This is not a direct relationship. Vocabulary knowledge can help reading and reading can contribute to vocabulary growth (Nation,2001). Related with the vocabulary size of a

learner and the learner's reading comprehension, the term *text coverage* is used. Text coverage is the amount of comprehension that is provided by the current vocabulary knowledge. In other words, texts coverage gives the percentage of how much of a text is understood by a person. Table 2.5 shows the vocabulary size and text coverage in the Brown corpus:

Table 2.5 The vocabulary sizes and text coverage in the Brown Corpus

Vocabulary Size	Text coverage
1000	72.0%
2000	79.7%
3000	84.0%
4000	86.8%
5000	88.7%
6000	89.9%
15,851	97.8%

The figures in the table refer to written texts and are from Francis and Kucera (1982) in which over 1,000,000 running words are used from 500 texts of around 2000 running words long. The figures in the last line of the table are from Kucera (1982). The COBUILD Dictionary claims that 15,000 words cover 95% of the running words of their corpus. The figures in the table are for lemmas and not word families. Word families would give higher coverage. The figures in the table assume that high frequency words are known before lower frequency words and shows that knowing about 2,000 word families gives near to 80% coverage of written text. The same number of words gives greater coverage of informal spoken text - around 96% (Schonell, Meddleton and Shaw,1956).

Current studies show that there is a threshold for successful reading and for reaching that threshold the text coverage should be around 95% of the words in the text. In other words, successful comprehension is not simply decoding the vocabulary in a text. Lack of more than 5% of the running words in a text can make reading a formidable task, so the input becomes incomprehensible. According to Hu and Nation (in press) for largely unassisted reading for pleasure, around 98% of the running words is to be known. In Hu and Nation's study, it is suggested that, for a fiction text, the all-or-nothing threshold is 80% of vocabulary coverage. Learners with this coverage could not achieve adequate comprehension. The probabilistic threshold for successful comprehension is 98%. With this coverage, almost all learners have a chance of gaining adequate comprehension. A minimally acceptable comprehension is gained by 95% coverage. (Nation,2001)

From the aspect of meaning focused input, the needed coverage for learners is 95%. This amount of coverage is needed for learning vocabulary from meaning-focused input. For fluency development, the necessary coverage becomes 98%. For learners who can not reach this level, simplified reading material of various levels should be used in order to learn from meaning focused input and to develop fluency in reading. The figure of 95% coverage is also needed for a learner to use his/her reading skills that were acquired in the first language (Laufer 1988,1992,1997;Hirsh& Nation,1992).

Research on L1 reading manifested that there is a close relationship between vocabulary knowledge and reading comprehension. This is an indirect relationship. Vocabulary knowledge can facilitate reading and make it enjoyable because of the increase in the percentage of text coverage and reading can contribute to vocabulary growth. For learning vocabulary, learners should have 95% coverage from the reading texts, which means the percentage of unknown vocabulary should not be above 5% of the tokens in the text and preferably no more than 2% to make sure that comprehension and guessing can occur and no less than 1-2% to make sure that there is new vocabulary to learn. Qualities of texts are also important for text coverage and learning vocabulary. Texts such as continuing the same topic, for example, provide favorable conditions because of the repetition of the unknown vocabulary. Repetition provides a good opportunity for unknown words to be known, so learners may

encounter less unknown words as they progress in the book, by the way, their text coverage increases. It should be kept in mind that, unknown vocabulary that decreases the percentage of text coverage causes learners' reading to slow down and makes difficult to gain smoothness and flow needed for pleasurable reading. To provide learners input at the right level pleasurable reading, that is the level just beyond their present vocabulary knowledge, graded readers can be used.

Elementary learners can read graded reader where 95% of the vocabulary is already familiar. Remaining words can be learnt by guessing from context or dictionary use. (Nation,2001)

Some more researches about the optimum degree of comprehension can be listed as below:

Hill and Thomas (1988:45) suggest that the number of unknown words in a text should not be more than one in every ten running words. Hirsh and Nation (1992) suggest that the unknown words should not be more than two in every 100 running words. West considered that the ratio of unknown words to known should be one in fifty running words (West 1955:21) Nation and Wang (1999) after examining a corpus of forty-two graded readers, concluded that the ideal percentage coverage of text by newly introduced words is 4% or less to make comprehension and guessing from the context easier. It seems that there is a consensus among the results of the researches on account of knowing the 95% percent of the running words in a text to benefit from the reading material in language learning.

In teaching English, to provide the necessary text coverage for the students in order to follow the lesson, it is important that teachers match learners' level and reading material to suit the various goals of learning vocabulary through reading, developing fluency in reading, reading with adequate comprehension and reading for pleasure. The several ways for doing this include letting learners select the material that they want to read on a trial and error basis or to test learners' vocabulary knowledge using receptive vocabulary measures like the Vocabulary Levels Test and then advise learners.

Testing learners' vocabulary at the beginning of the course is of great importance for teachers to have a chance of choosing the suitable level material and

setting goals for vocabulary learning. In some ways, testing vocabulary may be easier than testing structures of a language because the testing is done for separate items. Before extending the topic, testing vocabulary sizes of the learners, how much vocabulary is needed to know by a learner should be discussed. Hirsh and Nation (1992) examined unsimplified texts which provide the most favorable conditions for reducing the vocabulary burden in reading. They concluded that 2000-word general service vocabulary is not enough to read for pleasure. To read such kind of unsimplified material, learners should have a vocabulary knowledge of around 5000 word families.

In the present study students' vocabulary sizes are compared with the vocabulary load of the book. For making this comparison, students' vocabulary levels are to be defined by using a vocabulary test. In such vocabulary tests, different kinds of vocabulary items can be used for different testing purposes.

Some tests are prepared using True/False items. These kind of tests are used to test the breadth of learners vocabulary knowledge and their results can be used for placement and researches. (Nation, 1993a) Items in the tests can be as in the examples:

Write T if a sentence is true. Write N if it is not true. Write X if you do not understand the sentence.

1. *We cut time into minutes, hour and days.* ____
2. *Some children call their mother mama.* ____
3. *All the world is under water.* ____
4. *When you keep asking, you ask once.* ____

Like the test items above learners can be tested by YES/NO or checklist tests. The results can also be used for placement, and deciding what course level a learner should be placed in.

Example for YES/NO checklist:

Thick the words you know:

*adviser*____	*moisten*____
ghastly ____	*pitiful*_____
concord ____	*profess*____
implore ____	*stourge*____
*morlorn*____	*discard* ____

Some tests are prepared using definition completion items as the following. (Read, 1995)

Choose one word from the list on the right to complete the sentence. Do not use the same word twice.

1. *A journey straight to a place is_____*
2. *An illness that is very serious is _____*
3. *A river that is very wide is_____*
4. *Part of your body that is not covered by any clothes is_____*

alien
direct
acute
bare
faint
common

Some tests give more sensitive results for evaluation purposes. (A sensitive multiple choice test, Joe,1994)

chronic means *a. lasting for a long time*
b. dissatisfied
c. to greatly decrease
d. effective and harmless
e. don't know

2.4 Research Questions

In the present study, the English language course book "An English course For Turks Advanced 1" which is used for teaching English at grade 10 in Turkish state schools is examined and answers for the following questions are searched for:

1. How much vocabulary is needed to understand the course book?

2. How much of the course book can an average learner understand with the vocabulary he/she knows?

3. How much vocabulary is needed to understand the different sections of the book?

   a. How many words does each section require to be known in order to be comprehensible? (i.e. tokens)

   b. Which sections of the book require to be known more and less number of different words (i.e. types)?

   c. How many word families does each section require to be known in order to be comprehensible?

   d. Which sections of the book require to know more and less word families?

4. How much of the different sections of the course book can an average learner understand with the vocabulary he/she knows?

SECTION THREE
METHODOLOGY

Section 3.1 introduces the course book (An English Course For Turks Advanced 1), that is analyzed in terms of vocabulary load. Parts of the book are examined in detail, by using extracts from the book. Methods for selecting the sample is also presented. Section 3.2 presents how the vocabulary load analysis is done and how text coverage is examined. In section 3.3 participants of the study is presented.

3.1. The Vocabulary Load Analysis

In section 3.1.1 some information about qualities of the course book is given. Units of the book, lessons of each unit and parts of each lesson (dialogues, reading passages, sample sentences and exercises) are presented. In section 3.2 how word family analysis is done and tools for this analysis are presented.

3.1.1. The Course Book

In the present research, An English Course For Turks Advanced 1 Grade 2 was analyzed. This book was written by Semahat DİKMEN, Nevzat GÜRMAN, Ülkü ÖZGÜLER and M. Lütfi SALMAN. The book was published by National Education Press, in 1992. The first edition of the book was printed in 1965 and this was the 15th printing of the course book. Apart from the cover of the book, there is no difference between the first edition and the current printings. Using this book is made compulsory by the National Education Ministry, in all of the general state high school's 10th grades in our country.

The book consists of five units. Units one, two and three have two sections: The First Lesson and The Second Lesson. Units four and five have three sections: The First Lesson, The Second Lesson and The Third Lesson (See table 3.1). In each lesson, there are three parts: Structure, Reading and Exercises. The structure part is handled as two different sections: Dialogues and Sample sentences.(See table 3.2)

Table 3.1 Units and lessons in the book

UNIT ONE	UNIT TWO	UNIT THREE	UNIT FOUR	UNIT FIVE
Lesson 1	Lesson 1	Lesson 1	Lesson 1	Lesson 1
Lesson 2	Lesson 2	Lesson2	Lesson 2	Lesson 2
			Lesson 3	Lesson 3

Table 3.2 Parts of a lesson

Part One	Part Two	Part Three
Structure	Reading Passage	Exercises
a.dialogues		
b.sample sentences		

Part one is Structure. This part of the book presents the grammatical structures which will be taught in this unit in the form of dialogues and sample sentences. (See table 3.3 and 3.4)

Table 3.3 Textbook Extract: Unit Three, The First Lesson, Dialogue

UNIT THREE

The First Lesson

PART I
STRUCTRE

Mr Stone :What good weather we're having for May?

Mr Hicks :Yes, it's really very nice indeed. I wish I was on holiday...

Mr Stone: Can you take a few days off?

Mr Hicks: Not really. Besides, if I took a few days off where would I go?

Mr Stone: Well, you could go to the coast. I know what I would do if if I had a few days off.

Mr Hicks:What?

Mr Stone: I'd go to the Lake District and do some fishing.

Mr Hicks: No, it's too crowded. I think I'd just sit in the garden and do some reading.

do s

Table 3.3 presents a dialogue whose purpose is to sample the usage of "If Clause Type Two".

Table 3.4 Textbook Extract: Sample Sentences (Unit One, The First Lesson)

Table 3.4 presents the sentences used to show the usage of affirmative and interrogative forms of "Future Continuous Tense".

I'll be	having dinner leaving for Istanbul Painting the bathroom taking a finishing these letters doing an experiment travelling by bus	this time tomorrow. tomorrow night. then. in the afternoon. next lesson. all night.

Will you be going anywhere near the	post office? chemist's? electrician's? bookshop? butcher's? greengrocer's? supermarket? record shop? grocer's?

The second part in each unit is reading passages (See table 3.5). In this part, a passages about 500 words are given. Some passages are supported with comprehension questions after them.

Table 3.5 Textbook Extract: Unit One, The First Lesson, Reading Passage

PART II

READING

What sort of cities will people be living in a hundred years from now? Will they be living in cities under the sea or will they be living in huge tower cities, several thousand meters high? Architects, builders, and city planners are already thinking about the cities of the future. One committee of architects and designers in England says that future cities can be built at sea. They have even built a model of such a city. It is shaped like a harbor with huge, outer walls. Flats and shops can be built into these walls.

Why are people in the twentieth century so concerned about the cities of the twenty-first century. The population of the world is increasing very rapidly indeed. In many countries more and more people are moving from the countryside to towns and cities. It is so in Turkey, too. Take Istanbul and Ankara, for example. They are getting bigger and bigger day by day. Until only a short time ago, say two or three decades, the population was well below one million. It has doubled, tripled and even quadrupled. This, in turn, has caused several problems: overcrowding,air and water pollution, transport, crime and poverty. These are the problems that need to be solved urgently. Experts and those who are responsible have been thinking hard to solve them i.e widen and improve the roads, and clean the water..: etc. A lot is being done. Take the Golden Horn, for example. Istanbul's population has exploded from one million to seven million. That is why a lot of slums, factories and warehouses mushroomed along the shores of the Golden Horn. More and more industrial wastes and city's raw sewage poured into the horn. Now many of the buildings that contributed directly to pollution have already been demolished and vacant land is being converted into parks and playgrounds.

The third part in each unit is exercises. (See table 3.6, 3.7, 3.8) In this section students are given grammar exercises in different forms (Building questions from statements, building new sentences, filling in the blanks, combining sentences, matching...)

In table 3.6 "building questions from statements" exercises are presented.

Table 3.6 Textbook Extract: Unit One, the First Lesson, Exercises

PART III

EXERCISES

A. "Look at the example and then do the same..

A: I'm going to watch T.V. now.

B: Will you still be watching it when I get back?

1.A: I'm going to do my homework now.

B:.....

2.A: I'm going to type these letters now.

B:.....

3.A; I'm going to clean the windows now.

B:

4.A: I'm going to draw a picture of a volcano now.

B:.....

In table 3.7 a controlled exercise for sentence formation is shown.

Table 3.7 textbook Extract: Unit Two, The Second Lesson, Exercises

D. Look at the example and do the same.

Example:

John is in the library/in the laboratory. John is either in the library or in the laboratory.

1.Ali wants to study: physics/chemistry.

2.They have decided to spend their holiday: Turkey/Cyprus.

3.The director will: telephone/send a letter.

4.If you like, we can: watch television /go to the cinema.

5.Most people's eyes are: brown/blue.

6.They will be leaving for Bodrum: this weekend/next weekend.

7.By the end of this month they will have been in Ankara for: eight/nine months.

8.The speaker will be introduced: by the president/by the secretary.

In table 3.8 matching exercises for practicing “If Clause Type 2” are, presented.

Table 3.8 Textbook Extract: Unit Three, The Second Lesson, Exercises

PART III EXERCISES

A. Choose appropriate clauses from column B to go with the 'if clauses in column A.

Example:

A.

1. If that man took a taxi, he'd be able to catch his train.
1. If that man took a taxi,
2. If I were rich,
3. If Mrs Short ate less,
4. If I went to England,
5. If you wrote more letters,
6. If that teacher didn't talk so much,
7. If she bought an alarm clock,
8. If he brushed his hair,

B.

you would receive more.

she wouldn't be so fat.

he'd be able to catch his train.

I wouldn't eat any English food.

those students would work harder. he'd probably get a job.

they'd be able to help immediately.

she'd get to the office earlier.

I'd buy a jet plane.

There is a word list at the end of the book, which shows the words which have been taught in the book in an alphabetical order with their pronunciations. Nevertheless, the pronunciations are given in the form of phonetic alphabet and the syllabus does not contain lessons aiming to teach the alphabet. So, students are not able to benefit from the word list while they study pronouncing the words by themselves for they are not able to read the phonetic alphabet. The list was not prepared according to the units, instead, words in the book are listed randomly.

If the 1st and the 15th printings are examined, it can clearly be seen that the only change in the book is the book's cover. There is no difference between the 1st and 15th printings respecting the topics, the reading passages, the syllabus, the pictures and even misspellings. Because the book was not updated, the topics do not attract student's attention and they find it boring. For example, in a text about sports, the book mentions about Yaşar Doğu, İsmail Akçay and Metin Oktay as famous people in different branches of sports. But today's teenagers do not know these people because they were famous decades ago. Also, it seems that the methodology of the book may not support the contemporary English Teaching methods, whose goal is developing fluency in language rather than teaching the grammatical structures without using it in real life. In other words, it can be said that the book is outdated in terms of content and methodology.

3.1.2 The Analysis

The vocabulary load analysis of the book in terms of frequency in language provides important implications related with the efficiency of the book in teaching English and suitability of the book with the intended learners, because load (i.e. difficulty) is equated with frequency. That is, words in the book are analyzed in terms of their frequency and more frequent words are accepted to be easier than less frequent words because they occur about 80% of the running words in all kinds of texts. Once they are learned, they facilitate comprehension, thus, they facilitate learning. So, the proportion of the most frequent words in an English course book gains importance because of the priority to teach and learn the most frequent words to make the learning process more efficient and easier for the learners. In other words, if EFL students are taught by a course book which has a great proportion of higher frequency words, it means that the valuable classroom time is spent on the most necessary words to comprehend and use the new language, instead of the infrequent words which are not worth spending time during the lesson. On the other hand, if the EFL course book, covers a great proportion of lower frequency words, there will be much more words to learn for the students since the words they have to learn are not repeatedly used because of their infrequency, so learning new words will be more difficult. The less words learners know from teaching material, the less comprehensible the material becomes. This condition results in learners' getting higher level input than i+1 (just one level

beyond their current knowledge) as Krashen states in his Comprehensibility Hypothesis. (Krashen, 1985). As a result, the whole learning process becomes an infailure for both teachers and learners.

To start the vocabulary load analysis of the book, one lesson from each unit was sampled with all of its sections (structure, reading, exercises). About 45% of the book was sampled for analysis because of the time constraints. This amount of sampling was thought to be representative enough. The sampling was not done only from the first 50 pages or the first units of the book. It was done systematically from each unit of the book in order to obtain a representative sample and to see whether the vocabulary level is the same in all of the units. Table 3.9 shows which lessons are chosen from each unit. The First Lesson from Unit One, The Second Lesson from Unit Two, The First Lesson from Unit Three, The second Lesson from Unit Four and The Third Lesson from Unit Five are sampled for the analysis. All the sections in a lesson, ie. Dialogues, Sample Sentences, Exercises and Reading Passages were used in the analysis. (See Appendix I for the whole sample lesson.)

Table 3.9 Sample lessons randomly chosen from each unit for the analysis

Unit 1	Unit 2	Unit 3	Unit 4	Unit 5
Lesson 1		Lesson 1		
	Lesson 2		Lesson 2	
				Lesson 3

The lessons were turned into electronic documents by scanning, in order to be used at the computer environment. Since the scanned documents were in the BMP or JPEG (picture file) format, they couldn't be modified. So, a PDF transformer program, Abby Fine Reader version 7.0 was used to transform the image files to word documents.

3.1.2.1 How was the analysis done?

After the data had been turned into word format, the frequency analysis was done by using a special computer program RANGE32. It is a program used to compare a text against vocabulary lists to see what words in the text are and are not in the frequency lists, and to see what percentage of the items in the text are covered by the lists. It can also be used to compare the vocabulary of two texts to see how much of the same vocabulary they use and where their vocabulary differs. RANGE and FREQUENCY

were programmed by Alex Heatley. The lists were designed by Paul Nation and Averil Coxhead. The program can be reached from the following website: http://www.vuw.ac.nz/lals/staff/Paul_Nation

The program has two different parts RANGE and FREQUENCY. FREQUENCY runs on a text file to make a frequency list of all the words in a single text. The output is an alphabetical list, or a frequency ordered list. It gives the rank order of the words, their raw frequency and the cumulative percentage frequency.

As it has been explained before, to know the frequency of the the words in the book will provide information about the difficulty of the book and its suitability for the intended learners. To get this information, words in the book are compared with the wordlists of 1000, 2000 most frequent words and Academic Words in English. This comparison, which is important in determining the difficulty of the text, is done using the RANGE32 program and the vocabulary load of the book is revealed.

RANGE is used to compare the vocabulary of up to 32 different texts at the same time. For each word in the texts, it provides a range or distribution figure (how many texts the word occurs in), a headword frequency figure (the total number of times the actual headword type appears in all the texts), a family frequency figure (the total number of times the word and its family members occur in all the texts), and a frequency figure for each of the texts the word occurs in. It can be used to find the coverage of a text by certain word lists, create word lists based on frequency and range, and to discover shared and unique vocabulary in several pieces of writing. RANGE can be used to compare a text against vocabulary lists to see what words in the text are and are not in the lists, and to see what percentage of the items in the text are covered by the lists. It can also be used to compare the vocabulary of two texts to see how much of the same vocabulary they use and where their vocabulary differs.

It is useful, for example, for seeing what low frequency words are in an exam question paper, a technical information note or a text aimed at foreign readers. It may also be used to check the vocabulary of simplified reading texts or language course books to see how many of the words in the texts are among the high frequency words of English. It may also be used to see how much learning the vocabulary of one text provides in comparison with the words in a different text.

In combination with the three base lists that are available with it, it has been used, by various researchers, to answer the following questions.

What common vocabulary is found in all these texts?

How large a vocabulary is needed to read this text?

If a learner has a vocabulary of 2,000 words, how much of the vocabulary in the text will be familiar to the learner?

What are the words in the text which the learner is not likely to know?

How well does the course book prepare learners for the vocabulary in newspapers?

How rich a vocabulary do second language learners use in their free writing?

3.1.2.2 Using the Baseword Lists

RANGE can be used with an unlimited number of word lists. These allow it to classify some of the words in the input files into word families. The program will give different figures depending on whether the base word lists which have been added in the range programme by the programmer himself are used or not. If the base word lists are used, the figures will represent a mixture of families and types. All the words in the base word lists are counted as families and the remainder are counted as types. If the base word lists are not used, then all the words are counted as types, because it is the base word lists that are used to make families.

Three ready made baseword lists are available. The first (BASEWRD1.txt) includes the most frequent 1000 words of English. The second (BASEWRD2.txt) includes the 2nd 1000 most frequent words, and the third (BASEWRD3.txt) includes words not in the first 2000 words of English but which are frequent in upper secondary school and university texts from a wide range of subjects. All of these base lists include the base forms of words and derived forms. The first 1000 words thus consists of around 4000 forms or types. The sources of these lists are A General Service List of English Words by Michael West (Longman, London 1953) for the first 2000 words, and The Academic Word List by Coxhead (1998, 2000) containing 570 word families. The first thousand words of A General Service List of English Words are usually those in the list with a frequency higher than 332 occurrences per 5 million words, plus months, days of the week, numbers, titles (Mr, Mrs, Miss, Ms, Mister), and frequent greetings (Hello, Hi

etc). The word forms in the base lists are grouped into word families under a headword. For example, the headword *aid* has the following family members *aided*, *aiding*, *aids*, and *unaided*. In the base lists the family members have a Tab in front of them. The headword occurs just before the family members and has no Tab.

3.1.2.3 Preparing Texts For The Analysis

As it was said before, the sample parts chosen from the book were transferred into the computer environment as image files by scanning and then they were transformed into word files by using a PDF transformer program ABBY FINEREADER. To start the scanning, Abby Finereader 7.0 was set to the computer and run. The scanner (Canon SmartBase MPC190) was turned on. Using the scan button in the program, the pages of each lesson were transferred into batch files (the newly scanned documents which are prepared for the transformation). Before starting to scan the documents, the image quality of the scanning should be increased to at least 300 dpi for getting more successfully read texts by the program. After the scanning action, the next step was reading the image files. The program reads the image files to transform into word format. This action lasts about 10 seconds or more per page, according to the density of the text and images on the page. The last step is saving the documents in word format.

After the transforming had finished, the texts were prepared for the analysis program (Range32). The texts should be modified in order to get reliable results from the program. First, the letters indicating different parts such as A.,B.,C.s were deleted from the texts.(See table 3.10, 3.11)

Table 3.10 The original extract from the book

UNIT TWO The Second Lesson

PART I STRUCTURE

A.

Murat: Is Yilmaz coming to the wedding tomorrow?

Necdet : Yes, he is. And he might bring his sister too.

Murat: Oh, is she back from Istanbul then?

Necdet: Yes, she got back last night.

Table 3.11 The modified text which has been prepared for the analysis

UNIT TWO The Second Lesson PART I STRUCTURE ..... Murat: Is Yilmaz coming to the wedding tomorrow? Necdet : Yes, he is. And he might bring his sister too. Murat: Oh, is she back from Istanbul then? Necdet: Yes, she got back last night

Proper names were deleted from the texts. (See table 3.12, 3.13)

Table 3.12 The original extract from the book

John: Have you seen my umbrella? I can't find it anywhere. Mary: Have you looked in the car? It might be there. John: No, it's not there. I've looked. Mary: Did you take it to work yesterday? John: Yes, I think so. Mary: Well, you might have left it in the office. John: No, I couldn't have done that. It was raining when I left.

Table 3.13 The modified text which has been prepared for the analysis

..... Have you seen my umbrella? I can't find it anywhere. ..... Have you looked in the car? It might be there. ..... No, it's not there. I've looked. ..... Did you take it to work yesterday? ..... Yes, I think so. ..... Well, you might have left it in the office. ..... No, I couldn't have done that. It was raining when I left.

Punctuation marks were deleted. (See 3.14, 3.15)

Table 3.14 The original extract: a reading text

Crime is on the increase everywhere in large cities and towns as well as in small villages. The most common crimes are probably housebreaking or burglary, and robbery of one kind or another. In most cases the police are successful in catching criminals. It might take a long time but usually the burglar or housebreaker is caught and is eventually sent to prison. The general public can help the police to catch criminals in many ways. In some Countries you often see notices on the walls of stations, public buildings or government offices with. 'Help Prevent Crime' written in large letters. Of course the public can help to prevent crime, particularly crimes such as burglaries and theft. They can report suspicious-looking people to the police, and they can take many different precautions in their own homes. They can have double locks on their doors; they can have burglar alarms; or they can simply keep a fierce dog. Occasionally however the general public helps the criminal. Take the Topraks for example. They had been away on holiday for a fortnight. When they returned home, they found that their house had been burgled. A lot of valuable things had been stolen, so naturally they called the police. After the police had thoroughly examined the house, they began to ask questions. When did you go away on holiday?" asked the senior police officer."At the beginning of the month," replied Mr. Toprak "Yes, a fortnight ago", added Mrs. Toprak. "The burglars must have known that we were away."

Table 3.15 The modified text which has been prepared for the analysis. (Dashes are used to show the erased punctuation marks.)

READING

Crime is on the increase everywhere in large cities and towns as well as in small villages - The most common crimes are probably housebreaking or burglary-and robbery of one kind or another In most cases the police are successful in catching criminals - It might take a long time but usually the burglar or housebreaker is caught and is eventually sent to prison-

The general public can help the police to catch criminals in many ways - In some countries you often see notices on the walls of stations public buildings or government offices with -Help Prevent Crime- written in large letters - Of course the public can help to prevent crime particularly crimes such as burglaries and theft - They can report suspicious looking people to the police and they can take many different precautions in their own homes - They can have double locks on their doors - they can have burglar alarms or they can simply keep a fierce dog-

Occasionally however the general public helps the criminal Take the Topraks for example - They had been away on holiday for a fortnight - When they returned home- they found that their house had been burgled. A lot of valuable things had been stolen, so naturally they called the police. was burgled - After the police had thoroughly examined the house - they began to ask questions -

- When did you go away on holiday - - asked the senior police officer -

-At the beginning of the month-- replied Mr- Toprak

-Yes- a fortnight ago-- added Mrs- Toprak- -The burglars must have known that we were away--

Short forms were turned into long forms. (See table 3.16, 3.17)

Table 3.16 The original extract from the book

Zafer : I'm leaving for Germany tomorrow. Ay§e : How are you going? . Zafer : I'm flying from Esenboga. Ay§e : Why don't you go by bus instead? It's much more interesting. Zafer : I wish I could, but I have to be in Germany the day after tomorrow. Ay§e : What a pity! If you went by bus, you'd be able to see something of other countries too.

Table 3.17 The modified text which has been prepared for the analysis.

Zafer : I am leaving for Germany tomorrow. Ay§e : How are you going? . Zafer : I am flying from Esenboga. Ay§e : Why do not you go by bus instead? It is much more interesting. Zafer : I wish I could, but I have to be in Germany the day after tomorrow. Ay§e : What a pity! If you went by bus, you would be able to see something of other countries too.

The modified word texts were saved as rich text format, which is the only suitable file format for the program. If the texts had not been modified, the results for vocabulary load of the book would have been found higher than its current level since the program counts each separate item as a word and put it in the results pages and percentages. The passages were modified in the word files and they were transferred to the rich word format files by copying and pasting.

3.1.2.4 Output of the Computer Analysis

After the texts had been prepared, the analysis was done for each part in each unit (sample sentences, dialogue, passage, exercises). RANGE provides a table which shows how much coverage of a text each of the three base lists provides. Here are the results of the analysis for each part of the lesson (See table 18).

Table 3.18 The Results of The Analysis For Unit Three Lesson One Dialogue

Processing file: D:\Documents and Settings\Melike Agan\Desktop\tez\19 Nisan\unit 3 lesson 1 dialogue.txt

Number of lines: 31

Number of words: 236

Reading: D:\Documents and Settings\Melike Agan\Desktop\tez\22.02.2006\RANGE32[1][1].vocab.in text.Nation\BASEWRD1.txt

Reading: D:\Documents and Settings\Melike Agan\Desktop\tez\22.02.2006\RANGE32[1][1].vocab.in text.Nation\BASEWRD2.txt

Reading: D:\Documents and Settings\Melike Agan\Desktop\tez\22.02.2006\RANGE32[1][1].vocab.in text.Nation\BASEWRD3.txt

WORD LIST	TOKENS/%	TYPES/%	FAMILIES
one	213/90.25	107/87.70	92
two	12/ 5.08	9/ 7.38	9
three	1/ 0.42	1/ 0.82	1
not in the lists	10/ 4.24	5/ 4.10	?????

Number of BASEWRD1.txt types: 4119 Number of BASEWRD1.txt families: 998

Number of BASEWRD2.txt types: 3708 Number of BASEWRD2.txt families: 988

Number of BASEWRD3.txt types: 3107 Number of BASEWRD3.txt families: 570

Table of Ranges: Types

122 Words appear in 1 input files

Table of Ranges: Families

102 Words appear in 1 input files

Total	236	122	102

Types Found In Base List One

TYPE	RANGE	FREQ
A	1	31
ABLE	1	2
ACTUALLY	1	1
ADDRESS	1	1
AFTER	1	1
ALL	1	9
ALWAYS	1	2
AMOUNTS	1	1
AN	1	1
AND	1	18

Table 3.18 shows the results of the analysis for dialogue part of Unit 3 Lesson 1. The analysis starts with the number of lines and words in the text. The number of lines is 31 and the number of is 236. Then, which lists has been used is presented. There are three ready made base lists available for the analysis. The first list (BASEWRD1.txt) includes the most frequent 1000 words of English. The second (BASEWRD2.txt) includes the 2nd 1000 most frequent words, and the third (BASEWRD3.txt) includes Academic Words in English. All of these base lists include the base forms of words, their with their derivations and inflections.

Numerals under the heading "TOKENS" show how many tokens in the text belong to the basewordlist one, two or three. The number of the tokens includes the total counting of the words, no matter how many times they are used. For example, If article "a" is used 10 times, it is counted as 10 tokens. Numerals under the heading "TYPES" show how many types there are in the text belonging the basewordlist one, two and three. In this part of the analysis, each word is counted once. The heading "FAMILIES" shows the number of word families, that is, the words in the text are counted according to their family groups. Under this heading the baseword, its derivations and inflections are counted as only one item. For example, the words play, playing, plays, played are counted as only one word. Percentages, next to the numerals of types, tokens and word families from each basewordlist show the proportion of each counting concerning the distribution of the items in the frequency lits. For instance, the number of list one tokens is 213, which constitutes the 90.25% of the total number of tokens in the text.

Numerals below the heading "RANGE" show how many passages the word has been used in. Since only one passage was analyzed, range is 1 for all of the words. The numerals below the heading frequency show how many times the same word has been used in the passage.

After analyzing each part of the lesson, the whole lesson and at last the whole sample of the book were analyzed (See appendix 2 for more results pages).

3.2 Participants

Participants of this study are students from Zübeyde Hanım Vocational Girls' High School. Ninety students from two different classes participated in this study. The students were administered Nation's 1000 and 2000 word level tests(Nation,1993a) to reveal their vocabulary sizes. These are standardized tests for determining testees' vocabulary knowledge in the most frequent 1000 and 2nd 1000 words of English. The tests were implemented at the end of the 9th grade as they are going to use An English Course For Turks Advanced 1 the following year. Results of the word level tests are used to see whether the vocabulary load of the book is appropriate for the vocabulary sizes of the intended students.

The 1000 word level test comprises 40 sentences. Testees are asked to read and comprehend the sentences, then according to the meaning mark the sentence as "true" or "not true". If they think that a sentence is true, they are to write "T" in the space at the end of the sentence. If the fact which is stated by the sentence is not true, testees are to write "N" in the same place. But, if the sentence is not understood, testees are asked to write "X" instead of T or N. (See table 3.19 for the sample items from the test. See appendix 3 to see the whole test.)

Table 3.19 Sample items from Nation's 1000 word level test

We can stop time	.....
You must look when you want to find the way	.....
When someone asks "What are you called?", you should say your name	.....
There are many ways to get money.	.....
All the world is under water.	.....
When you keep asking, you ask once.	.....

The 2000 word level test was made by Norbert Schmitt, Diane Schmitt and C. Clapham (Nation, 2001). The test constitutes ten group of words with their descriptions in English. In each group there are six words and descriptions of three of them. Testees are asked to find which of the words in the group match the three descriptions and write the numbers of words next

to the descriptions. Words in the groups are of different kinds (nouns, verbs, adjectives...).(See table 3.20 for the sample items from the test, see appendix 3 for the whole test.)

Table 3.20 Sample items from the 2000 word level test

1.copy	
2.event	
3.motor	----end or highest point
4.pity	----this moves a car
5.profit	---- thing made to be like another
6.tip	

SECTION FOUR

RESULTS AND DISCUSSION

This section will be developed according to the research questions stated in chapter 2. The results of the vocabulary analysis of the coursebook and the Vocabulary Levels Test implemented to the intended learners of this course will be examined comparatively. In section 4.1 the vocabulary load of the course book will be presented. This section tries to answer the first research question. In section 4.2 there will be a comparison between the vocabulary load of the book and the size of the learners' vocabularies which is related with the second research question and in section 4.3 the vocabulary load of the book will be examined according to each part of the book which are dialogues, sample sentences, exercises and reading passages.

4.1 Vocabulary Load of the Coursebook

Section 4.1 attempts to search for answers for the first research question:

Research Question 1. How much vocabulary is needed to understand the course book?

Table 4.1 shows the vocabulary level of the book. In table 4.1 the total number of types, tokens and families that attained from the analysis of the whole sample from the coursebook is shown. In the table, the numerals show the number of types, tokens and word families, considering their existence in each wordlist. The first column, shows the frequency wordlists (the first 1000, the second 1000 and the academic wordlists). The percentage values of each numeral is also shown below the numbers. The values of the distribution of the words in these three lists are given according to the vocabulary analysis of the sample randomly chosen from the course book. The results indicate that there are 909 different word families overall in the sample. 597 word families, which constitute 59.7% of the 909 word families existing in the book are among the list of the most frequent 1000 wordlist, 242 word families, which constitute 24% of the total number, are among the most frequent 2000 word list and 70 word families, which constitute about 12% of the total word families are from the academic word list (the academic word list contains 570 words so, the proportion is 12%).

Table 4.1 The Overall Load of The Course book

WORDLIST	TOKENS	TYPES	FAMILIES
ONE	7209 (69,38%)	1003 (60.35%)	597 (59,7%)
TWO	1070 (9.39 %)	368 (22.14%)	242 (24%)
THREE	642 (5.64%)	100 (6%)	70 (12%)
Not in the lists	1176 (15,59%)	191 (11,5%)	
Total	11390	1662	909

1662 types are used in this book. 1003 of the types, which is about 60% of the types exist in list one, 368 types, which is about 22% of the types used in the book are in list two and 100 types which is about 6% of the types used in the book exist in the academic word list. There are 191 types, 11.5% of the total number, which are not in these three lists. That is, the 11.5% of the types in the book are not among the most frequent 1000 and 2000 words or academic word list for they are lower frequency words.

When the tokens of the book are examined, it is seen that the total number of the tokens is 11390. About 70 % of the tokens (=7209) in the sample, are among the words from the most frequent 1000 word list, about 10% of the tokens (=1070), are among the most frequent 2000 word list, about 5,5% (=642), of the words are among the academic wordlist. About 15% of the tokens (=1176) are lower frequency words which are not in the most frequent 1000 and 2000 wordlists and the academic wordlist. It is important to note here that, only 70% of the words (=7209), are among the most frequent 1000 and 2000 wordlists. From these findings, it can be concluded that the sample used for analysis covers 11390 tokens which are counted considering how many times each word is used in the book and at least 21% of the words in the sample are lower frequency, difficult words which do not exist in the most frequent 1000 and 2000 wordlists. (As it was said before, 5.5% of the words are in the Academic Wordlist, and 15% of the words are higher level words which are not in the 1000 and 2000 wordlists and the academic wordlist. The count of these two findings is 21% of the words in the sample.)

4.2 Vocabulary Levels of the Students and the Coursebook

It is necessary to know the vocabulary size of the intended learners of the course book to make comparisons between the vocabulary level of the book and current levels of students. In this section the current vocabulary levels of students, which are tested by Nation's Vocabulary Levels Tests for the most frequent 1000 and 2000 words will be revealed and the answers of the second research question "How much of the course can does an average learner understand with the vocabulary he/she knows?" will be discussed.

Table 4.2 shows the vocabulary levels of the intended students. The word knowledge of the students is given using the results that have been attained from the implementation of Nations Vocabulary Levels Tests. As it is seen in table 4.2.1 the students know, on average, 391 word families of the first list (that is, the students know 39.1% of the most frequent 1000 words) and they know 196 word families of the second list (that is the students know 19.6% of the second most frequent 1000 words.) Their academic vocabulary knowledge was not tested since the students were seen to have a very low vocabulary level from the most frequent 1000 and 2000 words and it was not necessary to test their vocabulary knowledge from the academic vocabulary. From these two surveys, it is estimated that the students may know 587 words from 2000 words (39.1% of the first list, 19.6% of the second list.)

Table 4.2 The Vocabulary Levels of The Students

List One (The most frequent 1000 words)	391 word families
List Two (The most frequent 2000 words)	196 word families
List Three(Academic word list- 570 words)	(not tested)
TOTAL	587

As can be seen in table 4.2 the students know 391 word families on average from the first 1000 words. Only 597 word families from the first 1000 words (i.e. list one) appear in the sample. In order to identify the number of list one words familiar to the learners in the sample, 39.1% of the total number of the used words from list one have to be calculated. Table 4.3 shows the vocabulary load of the book and the vocabulary level of the students comparatively. As it is seen in the table, if learners know 39.1% of the 1000 words they may know 233 (.39x597) word families from list one in the

sample. As we implemented Nation's Vocabulary Levels Tests and these tests contain a standard number of sample words from the 1000 wordlist it is impossible to identify exactly which of the words the students know from this list. By the way, only the proportion of their knowledge from these 1000 words may be meaningful while comparing the load of the book with the levels of the students. The same calculation is necessary to get the proportion of the students' knowledge for the words of list two. Calculating 19.6% of 242 words gives the result that students may know 46 word families from the 242 word families of list two which were used in the course book. So, the total number of words which may be known by the students from the word families in the book is 279 and the proportion of this vocabulary size to the number of the total word families in the book (909 word families) is about 30.6%. This calculation shows that students may only understand less than one third of the words in the book.

Table 4.3 Vocabulary Load of the Coursebook and Vocabulary Size of Students

	Vocabulary Load of The Book	Vocabulary Level of the Sts
List One	597 word families (65.6%)	233 word families (39.1% of 597)
List Two	242 word families (26.6%)	46 word families (19.6% of 242)
List Three	70 word families (8.1%)	(not tested)
TOTAL	909 word families	233+46 =279 (30.5% of 909 WF)

By evaluating these findings under the light of Krashen's comprehensibility hypothesis and Hu and Nations's study about text coverage (in press, Vocabulary density and comprehension ,mentioned in Chapter 2), it can be understood that, knowing about 30.5% of the word families in the book, students will have difficulty in using this course book to learn English since the input presented to them is not on the same level or a few levels beyond their current level. The vocabulary level of the book is too high to manage for them and it is too challenging to accomplish understanding of the input in order to use it for learning the aimed language skills.

4.3 Different Sections

The aim of this section is to find answers for the research question three:

Research Question 3: How much vocabulary is needed to understand the different sections of the book?

In this section, each different portion of the course book will be handled with respect to their vocabulary loads and the compatibility of these vocabulary loads with the students' current vocabulary size.

4.3.1. Vocabulary Loads of the Different Sections of Book as Tokens

This section tries to answer the Research Question 3.a:How many words does each section require to be known in order to be comprehensible? (i.e. Tokens).Table 4.4 shows the number of tokens in each section. The number of tokens are examined here since the number of unfamiliar tokens in the text directly affects comprehension. That is, whenever an unknown word repeats, the student can not understand the same word and the overall comprehension decreases. That's why token coverage shows the comprehensibility rate of the text better. As it is seen in the table, with the number of 3406, the highest number of the tokens are in Reading Passages sections and with the number of 2078, the lowest number of tokens are in Sample Sentences sections. Concerning these figures, it is possible to say that, the longest part of the book is Reading Passages and the shortest part of the book is Sample Sentences. The second longest part of the book is Exercises, with 3206 tokens. The Dialogue part comes third, with 2700 tokens, in the case of length.

Table 4.4 Vocabulary Loads of the Different Sections in the Book as Tokens

WORD LIST	DIALOGUES	SAMPLE SENT.	EXERCISES	PASSAGES
ONE	1934 71%	1339 64.44%	2235 69.71%	2394 70.29%
TWO	211 7.81%	183 8.81%	366 11.42%	310 9.10%
THREE	138 5.11%	141 6.79%	171 5.33%	192 5.64%
Not in the lists	417 15.44%	415 19.97%	434 13.54%	510 14.97%
TOTAL	2700	2078	3206	3406

When the table is examined with respect to the distribution of the tokens in the three wordlists, It can be said that Dialogues are the easiest sections to comprehend for the students since 71% of the tokens are from the most frequent 1000 words, 7.81% of the words are from the most frequent 2000 words, 5.11% of the words are from the

Academic Wordlist. Since there are more words from list one (71%) and less words from list 2 and list 3 (total about 12%) than the other sections, Dialogues are concluded to be the most comprehensible section among the other sections of the book. Sample Sentences is the most difficult part for students to understand since it has the least percentage of list one words (64.44%) and the highest percentage of difficult words which are unlikely to be known by these learners, and this amounts to 25% (6% from list three and 19.97% from not in the lists) of unknown words. With the lowest number of higher frequency tokens and highest number of lower frequency tokens Sample sentences is the most difficult part. Although the aim of Sample Sentences sections is to give examples for the usage of the grammatical rules to help students to understand more easily and it does have to be understandable and accessible for the students, this matter was not taken into consideration by the writers of the course book. So, this made the Sample Sentences Parts less useful for the students and less effective as a teaching tool for the teachers that have to use this book in English classrooms.

From the respect of the words in the "not in the lists" line, it is seen that Sample Sentences Parts have the highest percentage of lower frequency words with 19.97% and exercises have the lowest percentage of lower frequency words with 13.54%.These words are almost certain to be unknown to the learners. So, the percentage of unknown words differs between 19.97% and 13.54%. That means, even for a learner who knows all of the words in the most frequent 1000, 2000 wordlists and the academic wordlist (2570 word knowledge), it is impossible for him/her to understand about 15% of the words in the book. If the intended students' level is remembered, an average student knows only 39% of the words in list one and 19% of the words in list two. The students can not know any words from list three since their level in list two is very low. Under the light of previous studies of vocabulary size and comprehension (Laufer, 1989), it is known that a learner is supposed to know 95%-98% of the words in a text to say that it is suitable for the learner's level and it is comprehensible for the learner and facilitative of his/her learning via reaching the sufficient degree of comprehension. If even a student who knows all of the words in these three lists cannot reach the necessary comprehension level, how can a student with a vocabulary size below 50% in these lists be expected to comprehend and benefit from this course book to learn English? How can such an inaccessible course book be used effectively by a teacher?

4.3.2. Vocabulary Loads of Different Sections in the Book as Types

In this part answers will be searched for the following questions:

Research Question 3: How much vocabulary is needed to understand the different sections of the book?

Research Question 3.a: Which sections of the book require to know more and less number of different words (i.e. types)?

Table 4.5 Vocabulary Load of the Different Sections in the Book as Types

WORD LIST	DIALOGUES	SAMPLE SENT.	EXERCISES	PASSAGES
ONE	438 71.45%	263 66.58%	519 67.58%	674 64.19%
TWO	112 18.22%	73 18.48%	164 21.35%	177 16.86%
THREE	21 3.43%	20 5.06%	32 4.17%	64 6.10%
Not in the lists	42 6.85%	39 9.87%	53 6.90%	135 12.86%
TOTAL	613	395	768	1050

In this part, different sections of the book will be compared as types. By examining the distribution of words in different sections as types it is possible to make inferences about the difficulty of the different parts of the book for the students. Types are examined, because the number of tokens in the book may not show clearly how many words students have to know in order to understand the section. That's because of the repetition of the unknown tokens in the book. That means, if an unknown token repeats many times, the number of unknown tokens increases by making the text more difficult for the students. So, the difficulty of the text in which the unknown token exist once and another token in which the unknown token exist ten times is different but the number of unknown words that students have to know is the same. In both of the conditions, students have to learn one word. By examining the number and percentages of types in the book,the vocabulary load of the book as types will be appeared and to estimate which section requires students to know more types than the other sections and which section has the highest vocabulary load will be possible. Table 4.5 shows the number of types from each section according to their distribution in wordlists.

When the table is examined, it is seen that with 1050 types, the highest number of different types are used in the Passages sections of the book. With 768 types, Exercises are the second and with 613 types Dialogues are the third in the order. With 395 types Sample Sentences have the least number of different types. So, Reading Passages require the highest vocabulary size from students to understand since the words in this section are quite varied. This result is not surprising since understanding 100% of a reading passage is not always necessary for a student to follow and benefit from a reading lesson. The unknown words may be ignored, guessed or tolerated by a student. Also, if one of the aims of a reading lesson is considered to be learning new words, the high number of different types in reading passages should not be surprising. Nevertheless, when the results are compared with students' vocabulary sizes (39% from list one, 19% from list two and none from list three) the reading texts in the book are not accessible for students. That's why students have to prepare long wordlists to understand the important unknown vocabulary, which are on the focus of the meaning. Since, the vocabulary sizes of the intended students are too low to understand the texts, they are unable to understand the texts. Their inability to cope with the level of difficulty may also have a negative effect on student motivation and interest.

The breadth of the required vocabulary size to understand the sample sentences sections is 395 types, so this section of the book has the least vocabulary load. The number of types that students require to know is low when compared with the other sections. But when the percentages of the words in Not in the Lists line are compared with the percentages in table 4.2, Not in the Lists line, a striking point appears. Whereas the percentage of the types from Sample Sentences in Not in the Lists line is 9.87%, the percentage of tokens in the same line is 19.97%. This shows that, while the percentage of unknown types is 9.87%, because of the repetition of the types throughout the section, they have a higher token coverage in the text. That is, they constitute the 19.97% of the text in this section. For the aim of this part of the book is to give examples of new grammatical points for students, the 100% understanding is necessary to avoid the shift of students' attention from grammar to vocabulary. While the students are struggling with the words that are not familiar, their focus changes and this part of the book becomes less useful and also causes to decrease student motivation.

The same point is correct for the dialogues. In the dialogues sections, the percentage of types in Not in the Lists line is 6.85%, but as tokens, the percentage of Not in the Lists tokens increases to 15.44%. In the same way, while the number of unknown types is not high, the repetition of the types (as tokens) causes a student to see the same unknown word again and again, so the comprehension level decreases.

This repetition may have another result. The repeated unknown words may be learnt in time and the level of comprehension may be a little bit higher for the Sample Sentences and Dialogues. But for the reading passages, there is little or no chance for the types to repeat in the text because the percentages of the words as tokens and types are not very different from each other. As can be remembered, the percentage of the tokens from passages in list three is 5,64%, and the percentage of types in list three is 6,10%. Also, the number of lower frequency tokens in reading passages is 14,97% and the percentage of lower frequency types is 12,86%. This suggests that most words are used only once. So, the unknown words in the reading passages will probably remain unknown.

4.3.3 Vocabulary Load of The Different Sections in the Book As Word Families

In this part answers will be searched for the following questions:

Research Question 3c: How many word families does each section require to be known in order to be comprehensible?

Research Question 3d: Which sections of the book require to know more and less word families?

Table 4.6 shows the number of words in each section of the book as word families. By examining the word families in the book, it is possible to get results about the number of words that students have to know to comprehend the book as the results attained by examining the types in the book. But this time,the description of word changes. While types are being counted, the inflected forms of a word are counted as separate items, but for the word family count the basewords, their inflections and transparent derivations are counted as only one item. Related with this count of word families, it is assumed that a student who knows the baseword, may understand its inflected and derived forms. In this respect, the word family analysis gives more meaningful results for teaching.

Table 4.6 Vocabulary Load of the Different Sections in the Book As Families

WORD LIST	DIALOGUES	SAMPLE SENT.	EXERCISES	PASSAGES
ONE	308 77%	189 72.6%	347 70.6%	423 73.31%
TWO	76 19%	54 20.7%	117 23.8%	113 19.5%
THREE	15 3%	17 6.5%	27 5.4%	41 7.1%
TOTAL	399	260	491	577

When the total numbers of word families are examined, it is seen that, with 577 word families, Reading Passages sections of the book have the highest number of word families the same as the evaluation of tokens and types in the book and Sample Sentences sections have the least number of word families again the same as the number of tokens and types in the Sample Sentences. The second most difficult sections are Dialogues and the third most difficult sections are Exercises in the order of vocabulary load as word families. So, for the students, the most difficult parts to understand are Reading Passages and the most comprehensible parts are Sample Sentences. The number of word families according to the wordlists shows that, the most comprehensible part for the students would be Dialogues because they have the most words from list one (308 word families from 399 are in list one) that's because students know more words from list one. In the same way, the least comprehensible sections of the book are Exercises which have 347 of its word families from list one (347 of 491 total number of word families). That means students will have great of difficulty in mastering the new forms of language because of their challenge with the unknown words in these sections. Reading Passages require to know the most number of word families as the total number of different word families are higher than the other three sections.

4.3.4 Vocabulary Loads of Different Sections in the Book and Vocabulary Size of Intended Students

In this part, answers for the following question will be searched:

Research Question 4: How much of the different sections of the course book can an average learner understand with the vocabulary he/she knows?

Table 4.7 shows the Comparison of the Vocabulary Load of Each Section and the Vocabulary size of the students in terms of the number of word families. The comparison is done as word families because the Levels Test, which was implemented to the students to determine their vocabulary size, tests word family knowledge.

Table 4.7 Vocabulary Size of Students With Respect to the Vocabulary Load of the Course Book

	DIALOGUES		SAMPLE SENTENCES		EXERCISES		PASSAGES	
WORD LIST	Wfs in the sample	Wfs known to the students	Wfs in the sample	Wfs known to the students	Wfs in the sample	Wfs known to the students	Wfs in the sample	Wfs known to the students
ONE	308 (77%)	120 (39.1% of 308)	189 (72.6%)	54 (39.1% of 189)	347 (70.6%)	135 (39.1% of 347)	423 (73.31%)	165 (39.1% of 423)
TWO	76 (19%)	15 (19.6% of 76)	54 (20.7%)	10 (19.6% of 54)	117 (23.8%)	22 (19.6 of 117)	113 (19.5%)	22 (19.6% of 113)
THREE	15 (3%)	Not tested.	17 (6.5%)	Not tested.	27 (5.4%)	Not tested.	41 (7.1%)	Not tested
TOTAL	399	135 (33% of 399)	260	64 (24.5% of 260)	491	157 (31.9% of 491)	577	187 (32.5% of 577)

Table 4.7 shows the vocabulary load of each section and the vocabulary size of the intended students and the coursebook comparatively. In the table, how many word families there are in each section according to their distribution in the wordlists are shown in the "Wfs in the sample" columns (word families in the sample) and students' knowledge of word families with their percentages are shown in the "Wfs known to the students" columns. The numerals in the "Wfs in the sample" columns were attained from the analysis of the sample chosen from the coursebook and they show how many word families there are in each section according to the frequency lists. The percentages below the numerals show the percentages of the word families from each list in each section. For example, the total number of word families in the Dialogues is 399. 308 of them (77%) are from list one, 76 of them (19%) are from list two and 15 of them (3%) are from list three. The numerals in the "Wfs known to students" columns show the number of word families students may know in each section. They were attained from calculating the proportion of students' vocabulary sizes in each list with the number of word families in different sections. For example, in Dialogues, there are 308 word families from list one and if we assume that the intended learners know 39% of the words in list one, the learners' vocabulary knowledge of the list one words in Dialogues becomes 120 words.(308/100=3,08 / 3,08x39=120) In the same way, as there are 76 word families from list two in Dialogues part and the students' percentage of known words is 19%, it can be assumed that the students may know 15 words of the list two words in Dialogues section. List three was not tested, since the students' level in the first 1000 and 2000 words is very low. In the table, the words which are not in the lists cannot be shown because they can not be classified and counted. Apart from these Not in the Lists word families, the learners' total vocabulary knowledge is given in the Total line. The numerals and percentages in this line were attained by the addition of the learners known words from list one and two. For instance, from the Dialogues section, a student may understand 120 word families from list one and 15 word families from list two. They do not know any words from list three. So, the learners' word knowledge from Dialogues is 135 word families which is the 33% of the total number of word families in dialogues sections.

By examining the numerals in the table, it is possible to say that the most comprehensible section of the book for the intended students is Dialogues part since the 77% of the words in the section belongs to wordlist one while the percentage of list one

words in other three sections are close to each other(Passages have 73% of its word families from list one, Sample Sentences 72.6% of its word families and Exercises have 70.6% of its word families from list one) . The reason for this inference is that students' vocabulary size is higher in the first list (39.1% of the words in List One.) than the other two lists. The higher the number of List One words, the more chance for a learner to know the words in a section. As the list three words are examined, it is seen that again the Dialogues sections become the most comprehensible since it has the least words in list three which the intended students are impossible to know. Dialogues sections have 3% of its words from list three which means that students have less risk to encounter unfamiliar words. In terms of total percentages of known high frequency vocabulary, these sections are also the most comprehensible. An average student may understand 120 of 308 word families from list one, 15 of 76 word families from list two and no word families from list three since his/her level in list two is fairly low. So, the learner may understand 33% of this part. That means, lower frequency vocabulary causes the main reason of difficulty rather than the unknown higher frequency words.

With the 24.5% learner comprehension of the high frequency vocabulary, the least comprehensible sections for students are Sample Sentences. A learner may understand 54 of 189 word families from list one words and 10 of 54 words from list two words in Sample Sentences.

When the total percentages of known words are examined, students know 32.5% of the word families in Reading Passages. Their knowledge of words are not very different from Dialogues and Exercises. Just the percentage of their word knowledge decreases to 24,5% in Sample Sentences and this makes the Sample Sentences sections less comprehensible than the other three sections according to the word family results. From the total results of word families, the difficulty of reading passages are seen approximately the same as Dialogues and Exercises. Nevertheless, for the Reading Passages, the problem is not high frequency words. As it was said before in section 4.3.2, the real problem for the Reading Passages is lower frequency words which are in list three and which are not in these three lists since they are lower frequency words than list 1, list 2 and list three words. Since these lower frequency words can not be counted as word families, they can not be shown in this table and the difficulty of Reading Passages is artificially seen close to the load of exercises and dialogues.

According to the number of list three word families, students are impossible to know 15 of 399 (3%) words from Dialogues, 17 of 260 (6.5%) from Sample Sentences, 27 of 491 (5.4%) from Exercises and 41 of 577 (7.1%) from Reading Passages. It can be estimated even only by the results of the word families from list three that students will not be able reach the threshold of 95% text coverage which is necessary for successful reading comprehension (Nation,2001) to follow the coursebook and use it for learning English. If the percentages of unknown words from list one and list two are considered alongside with list three, it is easy to see clearly that by knowing 39.1% of the list one WFs, 19.6% of the list two WFs and knowing no list three WFs, the intended students of the coursebook can not reach the necessary threshold of text coverage, which means the input they receive becomes incomprehensible for them (Krashen, 1985).

SECTION FIVE
CONCLUSION

5.1. Conclusions

This study showed that the vocabulary load of the course book "An English Course For Turks Advanced 1" is rather high for the level of the 10th grade students. The results that were attained from the analysis of 50% of the book were generalized to the whole book. From the results of this analysis, it is possible to see that the percentage of the low frequency and academic words are quite high. The level of these words is too high for the students to manage.

An average 10th grade student cannot comprehend even half of the words from list one, which constitutes about 70% of the words in the book. Apart from them, there are words from list 3 and lower frequency words which are not in these three lists. So, the students can understand some words from the list one (about 39%) and list two (about 19.1%) and they cannot understand the rest of the words.

When the sections of the book are compared, reading passages sections, with the highest percentage of unknown words and by the way, having the highest vocabulary load, are the least accessible sections of the book for the students. According to the percentage of low frequency words and their word coverage Exercises and Dialogues are equally more accessible for the level of the students than Reading Passages. But they are still above the level of the intended students of this course book. The load of the Sample Sentences causes difficulty for the students if their function is considered. Since their aim is to present new structures to the students, they have to be comprehensible for the students in order to keep the focus of the students' attention on the structures and make the students gain accuracy and fluency on the new forms of language. The vocabulary load of Sample Sentences sections isn't as high as the load of other sections because a fewer number of words (types or word families) are needed to be known, but the percentage of the repetition of unknown words is higher. That's why the students encounter the unknown words constantly and the book becomes inaccessible and inefficient for the students.

From the findings of this study, it is clear that the vocabulary load of the book is very high. What can be the causes of this high vocabulary load of the book?

This course book may be following the previous book of the set. This book may have its own vocabulary teaching aim and this aim may be starting from the first book of the set. Therefore, the vocabulary load may actually be appropriate for the intended users of the book. However, for at least the last ten years, the English books used in primary and secondary schools have been replaced and the previous books in an English Course For Turks series are not used. Even if the primary and secondary school degree English books of the series were simple as far as the vocabulary load is concerned, students who start the second year of their high school education would not have formed the basis for reaching the vocabulary level of this book and have difficulty in understanding the words in the book. Another reason for this difficulty may result from the English curriculum which is designed by The Ministry of National Education. The primary and secondary school English curriculum may be too heavy for the students to learn. So, the words which cannot be learned from each level may accumulate and at the 10^{th} grade the vocabulary load of the book becomes insurmountable for the students.

Also the reason for this difficulty may be the unsystematic treatment of the words in the book. If the book did not follow a syllabus, it could not build the vocabulary knowledge of students gradually with enough recycling and practice of learned items or proper control of vocabulary load of the units. Since, there is not any information about the syllabus and aims of the book in the teachers' book, whether the book has a syllabus can not be known. So, the randomly chosen and uncontrolled word usage may be the reason of the high vocabulary load of the book.

5.2 Further Study

The following items may be researched for further studies.

For getting a clearer result about the compatibility of the vocabulary load of the book with the vocabulary sizes of the intended students, words in all of the pages of the book may be researched. This study covers randomly chosen sample which is about 50% of the book.

The course book may be analyzed comparatively with another course book in the same level. To make the comparison more meaningful, the selected course book for the comparison should be a popular and reputable course book that has a systematic vocabulary syllabus, i.e., that has a control over the vocabulary load. If a course book with these qualities is selected, the comparison will be done with a norm and this will make drawing conclusions easier. The comparative analysis may contain an analysis of the frequency of words as well as the cycling of vocabulary and repetition of the words in the text. If repetition of the words is not enough, the chance to learn the unknown words becomes lower.

It is known that the words in "Not in the Lists" are 3000 wordlist words and lower frequency words than 3000 wordlists. The computer program used in this study (i.e. the Range Program) did not do a detailed analysis of the words beyond the 2000 word level. To make the level of the book clearer, the new version of the Range Program and the up-to-date BNC Lists in the website "Lextutor" , which cover from 1000 words to 20000 words may be used. During the analysis phase of the study, these lists were not available, that is why the Range Program was used for the analysis. For further research, the more detailed lists in this site may be used instead of Nation's Range Program.

Lastly and the most importantly, for determining the level of the students, a standardized test , i.e. the Vocabulary Levels Test by Nation (2001) with a limited number of words was used. So, the percentages of students knowledge from each level is known but which words they know cannot exactly be known. That means, for instance the percentage of list one words in dialogues is 77% and the students' vocabulary size from list one is 39%. But does this 77% cover this 39%? It is impossible to know this exactly from that kind of standardized vocabulary levels test. The ideal thing to do is to test all of the words in the book to get an absolute conclusion

about the level of the students according to the book. But this means to test hundreds even thousands of words. Because of the time constraint, and impracticality of this kind of testing, it could not be used for this study.

5.3 Recommendations

As the Ministry of National Education makes use of this book compulsory in teaching English to state high school students, the following may be helpful for both the students and the teachers.

This book may be revised and the words used in the book may be simplified by the Ministry of National Education, considering the vocabulary load of the intended 10th grade students.

Teachers may omit the reading passages in the book and use their own passages chosen from other resources suitable for the level of the students. But that is not a practical way because of the difficulty of choosing.

Teachers may train students in the learning strategies and build a large vocabulary so that the load of the book will no longer be heavy. For initial learning of words using word cards may be taught. This is an effective strategy since they are efficient in terms of time and effort from students. While using the word cards, learners focus deeply on acquiring new vocabulary in an enjoyable way which cannot be provided from reading or dictionary use. Also, learners can have a control over their repetition and vocabulary learning process. Ellis,1995 states that while a student is learning the meaning of a word, it is best to link the word with its meaning. But the students should be informed about how they can prepare the word cards, how long and how often they will study. These word cards should be prepared by writing the word on one side of the paper and writing its meaning on the other side. Lado,Baldwin and Lobo,1967, Mishima,1967 and Laufer and Shmueli,1997 found that if the L1 meaning is written, learning will generally be better. Using pictures can also be effective for learning words when the word cards are prepared. Kellog and Howe,1971 concluded that using pictures provides learning to take place significantly faster than the written words. After the preparation, the word cards can be used to study recalling words.

Teachers may teach memory techniques to facilitate remembering the new words. For instance, the students can be taught to use physical actions when learning a word or say a new word aloud when studying. Connecting word to a previous personal experience, connecting the word with its synonyms and antonyms, using semantic maps, grouping words together to study them and using the keyword method are among the other memory techniques that can be taught to students in order to facilitate their learning and memorizing the new words.

Students may be taught to guess some unknown words from context or they may ignore some words while they are reading the texts. These techniques may lighten the burden to learn too many words together by looking at the dictionary and making long wordlists.

Dictionary skills may also be developed. First of all, students should be given information about the best dictionary to suit their level and purposes. Then they should be thought the abbreviations used in the dictionary and the phonetic alphabet to understand the pronunciation of the words they look up. Also, they should be encouraged to use the dictionary properly. Students should be informed that looking up each unknown word is not necessary, instead they should be encouraged to look up the words which are on the focus of the meaning. Besides, they should be given information about how to choose the appropriate meaning of the word that supports the overall meaning of the text is.

REFERENCES

Bauer, L. - Nation, I.S.P.

1993 "Word Families." International Journal of Lexicography, 6, 253-279.

Chen, Qi – Ge, Guanc-chun

2007 "A Corpus-based Lexical Study on Frequency and Distribution of Coxhead's AWL Word Families in Medical research Articles (RAs), English For Specific Purposes 26, 502-514

Coxhead, Averil

2000 "A new Academic Word List", Tesol Quarterly, Vol.34, No.2, pp 213-238.

Ellis, Rod

2000 "Second Language Acquisition", Oxford University Press, New York.

Francis,W.N. And Kucera,H.

1982 "Frequency Analysis of English Usage.", Boston: Houghton Mifflin Company

Gardner, Dee

2007 "Validating the Construct of Word in Applied Corpus-Based Vocabulary Research: A Critical Survey, Applied Linguistics, 28/2: 241-265

Hill, D. R. - Thomas, H. R.

1988a "Survey Review: Graded Readers (Part:1), ELT Journal, 42, 44-52.

Hirsh, David – Nation, Paul

1992 "What Vocabulary Size is Needed to Read Unsimplified Texts For Pleasure?", Reading in a Foreign Language, 8(2).

Hu, M.- Nation, I.S.P.

(in press) "Vocabulary Density and Reading Comprehension.", Reading in a Foreign Language.

Joe, A.

1994 "Generative Use and Vocabulary Learning." Unpublished MA Thesis, Victoria university of Wellington.

Krashen, S. D.

1985 "The Input Hypothesis.", London, Longman.

Kucera,H.

1982 "The Mathematics of Language.", The American Heritage Dictionary, Boston:Houghton Mifflin Company 2nd ed.

Laufer, B.

1988 "The Concept of Synforms (similar lexical itams) in Vocabulary

Acquisition.", Language and Education, 12, 113-132.

Laufer, B.

1992a "Corpus-based Versus Lexicographer Examples in Comprehension and Production of New Words." Euralex '92- Proceedings, 71-76.

Nation, I.S.P.

2001 "Learning Vocabulary in Another Language", Cambridge University Press, Cambridge.

Nation, Paul – Wang , K

1999 "Graded Readers and Vocabulary", Reading in a Foreign Language, 12, 355-380

Read, John

1988 "Measuring the Vocabulary Knowledge of Second Language Learners", RELC Journal, 19; 12

Read, John

1995 "Refining the Word Associates Format as a Measure of Depth of Vocabulary Knowledge." Studies of Applied Linguistics, 1, 1-17 New Zealand

Schonell, F.J. - Meddleton, I.G. - Shaw, B.A.

1956 "A Study of Oral Vocabulary of Adults, Brisbane: University of Queensland Press.

Shirato, Junko – Stapleton, Paul

2007 "Comparing English Vocabulary In a Spoken Learner Corpus: Pedagogical Implications Arising From An Empiricla Study In Japan", Language Teaching Research, 11/4, 393-412

Sutarsyah, Cucu – Nation, Paul – Kennedy Graeme

1994 "How Useful Is EAP Vocabulary For ESP? a Corpus Based Case Study", RELC Journal, 25;34

Ward, Jeremy

1999 "How Large A Vocabulary Do EAP Engineering Students Need?, Reading in A Foreign Language, 12(2), 1999

White,T.G – Power M.A. -White, S.

1983 "Morphological Analysis: Implications for Teaching and Understanding Vocabulary Growth." Reading Research Quarterly, 24, 283-304

Whysocki, K – Jenkins, J.R.
1987 “Deriving Word Meanings Through Morphological Generalization.” Reading Research Quarterly, 22, 66-81.

APPENDIXES

Appendix:1 A sample lesson from "An English Course For Turks Advanced I"

UNIT FIVE

The Third Lesson

PART 1

STRUCTURE

A.

Student : Could I have the key to the dark room please?

Teacher : Why?

Student : I want to develop the film I took last week.

Teacher : No, I'm sorry. I'm not letting anybody use the dark room this week. The last time I let a student use it by himself, some of the equipment got damaged.

Student : If I break anything, I promise I'll have it repaired, or have it replaced.

B.

Mother : There's something wrong with the iron again.

Son : Let me have a look at it. I can probably mend it myself.

Mother : No, thank you. The last time you tried to mend it you almost caused a fire.

Father : That's right. We'd better get an electrician to have a look at it this time.

Mother : There's that new electrician at the end of the street. He might be able to repair it.

Son : He's good. I had my tape-recorder mended there. He didn't charge very much and he was very quick.

C.

Mother : If you're going to that New Years' party, you'd better have your long dress cleaned.

Daughter : There's no point in having it cleaned. It's too dirty, and anyway it's out of fashion now.

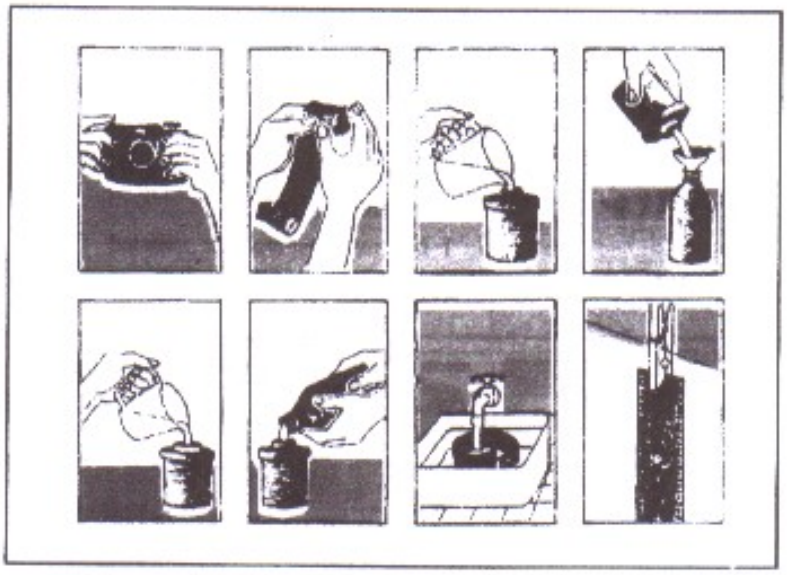
Mother : What are you going to wear then?

Daughter : I was thinking of having something made.

Mother : I don't think there's enough time to have anything made.

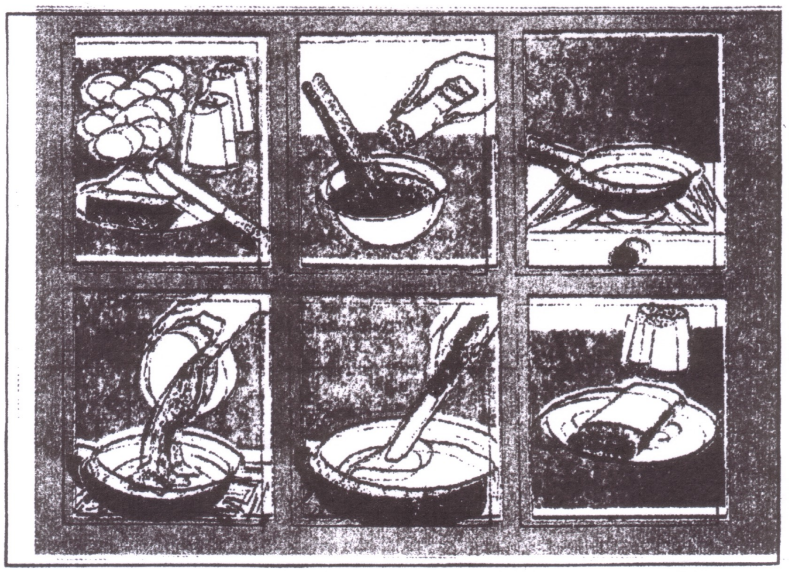
Daughter : Well, I could probably get that dressmaker in Bright Street to make something pretty quickly.

Mother : You always leave things to the last minute. You should've thought about this weeks ago.



If you have a camera and are interested in photography, it's a good idea to try to develop your own films, either at home or at school. First of all, you need a dark-room, or a room, which can easily be made into a dark-room by keeping out the light. This is necessary since films are very sensitive to light and must be developed without being exposed to any form of it. After taking the film out of the camera, put it into a tray containing developing fluid. Remove the film, and before putting it into a third tray filled with fixing fluid, wash it in a tray of water. After taking it out of the fixing fluid, rinse the film in running water.

When it is dry, your film is ready for printing.



To make an "omelette for four people, you need:

8 eggs

1 teaspoonful of salt

a pinch of pepper

2 tablespoonsful of butter

After beating the eggs, mix them thoroughly and add the seasoning. Before heating the butter in the pan, make sure the pan is perfectly clean and dry. (It's better, if you can, to keep a special pan only for making omelettes, and, instead of washing it after use, clean it by wiping it). When the butter has melted, pour in the eggs, keeping a good heat under the pan so that the omelette will cook quickly. As soon as the underside begins to set, start lifting the edge of the omelette, first in one place and then in another. Tilt the pan while doing so, so that the liquid egg runs underneath. When no more liquid will run underneath, your omelette is done, but it should still be quite moist on top. Using a knife, roll the omelette over now, and tip it onto a hot plate. It should be golden-brown on the outside and still moist inside.

STUDY THESE

1. Make meaningful sentence.

Tomorrow I'll	have get	my hair cut. the radio repaired. the fuse mended. my brown suit cleaned. the car polished. the walls painted.

2. Make meaningful sentences.

I	had	the electrician the secretary him	cut the grass. mend the fuse. type the letters. repair the electric iron.

3. Make meaningful sentences.

John	got	his brother his English teacher his sister his father	to	drive him to school. translate his letter. help him with his homework. iron his shirts.

4. Study the following.

The policeman		John	stop the car.
The director	made	him	do all the work.
The physics teacher		Mary	study harder.

PART II
READING

In England people begin to think about their summer holiday as early as January. As soon as the Christmas and New Year celebrations are over, newspapers and magazines are full of advertisements for exciting and exotic holidays both in Britain and abroad. Colourful brochures are sent out by travel agents, and in airline offices, as well as in bus and railway stations, attractive posters begin to appear. There are posters of places and buildings of historical interest and importance, of all kinds of scenery, mountains, valleys, lakes, rivers, woods and forests, but more than anything else, there are posters of beautiful, sandy beaches and calm blue seas, with luxurious hotels in the background.

Much of this advertising is successful, and as a result, many people decide to take a 'package holiday' of one type or another. Everything is organised for them-their travel arrangements, hotel reservations and sightseeing trips. Package holidays are usually cheap and that is probably why they are so popular. However, there are some people who refuse to travel in groups and who dislike having everything arranged for them; they prefer to travel alone or to spend their holiday with their family or with people they know.

Turkey has Turkish hospitality and friendship to attract tourists. These mean a lot to the visitor who is fed up with package tours and standard hotel rooms. Turkey offers many hotels, motels, resort areas, villages, and other types of accommodation. In Turkey one can find every type of climate, scenery and historical site. In other words, there are a lot of interesting places for a traveller to see. Popular centres include İstanbul, which is situated on seven hills, and is more than 2600 years old. Attractively decorated with minarets and domes İstanbul was once the capital of the East Roman Empire and the Ottoman Turks and is now a modern city of ten million people.

In Ankara in Central Anatolia, one can visit the tomb of the founding father of modern Turkey, Kemal Atatürk and the Museum of Anatolian Civilizations. In the south along the Mediterranean Coast. there are not only sunny beaches but also sites steeped in fascinating history like those at Side, Antalya, Perge which attract the tourists all the year round. If one is interested in more ancient cities like Ephesus, Pergamon, one must travel to the west.

People who plan to go on holiday in summer have to make arrangements a long time advance if they want to find a place that is reasonably priced. Many people, however, for one reason or another, have to wait until the last moment before deciding what they want to do or where they want to go. There are still others who never think of spending their holiday away from home. For them the ideal holiday is a holiday at home doing those things they have not had time to do all year long.

PART III
EXERCISES

A. *Look at the example and then answer in the same way.*
Example:

The windows were very dirty,
but I _____ now. (clean)
The windows were very dirty, but I've had them cleaned now.

1. His hair was very long, but he_____ now. (cut)
2. My watch wasn't working, but I _____ now. (mend)
3. Her skirt was too short, but she _____ now. (lengthen)
4. His car was very dirty, but he _____ now. (wash)
5. Her skirt was very creased, but she _____ now. (iron)
6. The chairs were very uncomfortable, but they _____ now. (repair)
7. His car never used to start in the mornings, but he _____ now.(lix)
8. His report was very difficult ro read, but he _____ now. (type)

B. *Join the following sentences as in the example.*

Example:

I had my hair cut. Zeki cut it for me.
I had Zeki cut my hair.

1. They had their television repaired. Ahmet repaired it for them.
_____
2. We had the windows cleaned. Fatma Hanım cleaned them for us.
_____

3. He had the grass cut. Nuri Efendi cut it for us.

_____

4. I had my report typed. The secretary typed it for me.

_____

5. He had his battery checked. The mechanic checked it for him.

_____

6. She had her temperature taken. The nurse took it for her.

_____

7. He had his trousers lengthened. The tailor lengthened them for him.

_____

C. *Repeat the above exercise using 'get _____ to' as in the following example:*

Example:

I had my hair cut. Zeki cut it for me.
I got Zeki to cut my hair.

D. *Change the following sentences as in the example.*

Example:
John had to give up smoking.
John's doctor made him give up smoking.

1. Ahmet painted the whole house.
Ahmet's wife _____

2. Ali did six exercises.
Ali's teacher_____

Appendix: 2 Results of the word analysis for unit one

```
                                    unitone_range
Unit one range analysis

Processing file: D:\Documents and Settings\Melike Agan\Desktop\tez\UNIT ON1.doc


  Number of lines: 493
  Number of words: 12459

  Reading: D:\Documents and Settings\Melike
Agan\Desktop\tez\22.02.2006\RANGE32[1][1].vocab.in text.Nation\BASEWRD1.txt
  Reading: D:\Documents and Settings\Melike
Agan\Desktop\tez\22.02.2006\RANGE32[1][1].vocab.in text.Nation\BASEWRD2.txt
  Reading: D:\Documents and Settings\Melike
Agan\Desktop\tez\22.02.2006\RANGE32[1][1].vocab.in text.Nation\BASEWRD3.txt


WORD LIST              TOKENS/%              TYPES/%              FAMILIES

one                     5842/46.89            315/40.54             257
two                       87/ 0.70             60/ 7.72              58
three                     18/ 0.14             12/ 1.54              11
not in the lists        6512/52.27            390/50.19           ?????

Total                  12459                  777                   326


 Number of BASEWRD1.txt types: 4119  Number of BASEWRD1.txt families: 998
 Number of BASEWRD2.txt types: 3708  Number of BASEWRD2.txt families: 988
 Number of BASEWRD3.txt types: 3107  Number of BASEWRD3.txt families: 570

Table of Ranges: Types

   777  Words appear in  1 input files

Table of Ranges: Families

   326  Words appear in  1 input files


Types Found In Base List One

TYPE                                RANGE     FREQ       F1
A                                       1      198      198
ABLE                                    1        4        4
ABOUT                                   1        3        3
AGO                                     1        1        1
AGREED                                  1        1        1
AIR                                     1        4        4
ALL                                     1        2        2
ALONG                                   1        1        1
ALREADY                                 1        3        3
AN                                      1        1        1
AND                                     1       25       25
```

	unitone_range		
BEING	1	3	3
BELIEVE	1	3	3
BELOW	1	2	2
BETTER	1	1	1
BIGGER	1	2	2
BOOK	1	1	1
BOOKS	1	1	1
BUILD	1	1	1
BUILDERS	1	1	1
BUILDING	1	2	2
BUILDINGS	1	1	1
BUILT	1	5	5
BURNING	1	1	1
BUT	1	1	1
BY	1	3	3
CAN	1	10	10
CAREFUL	1	1	1
CAREFULLY	1	1	1
CAUSED	1	1	1
CITIES	1	8	8
CITY	1	6	6
CLASS	1	1	1
COAL	1	2	2
COME	1	5	5
COMMITTEE	1	1	1
COUNTRIES	1	1	1
CROWDING	1	1	1
DAY	1	2	2
DEVELOP	1	1	1
DID	1	3	3
DIDN	1	1	1
DIRECTLY	1	1	1
DO	1	5	5
DOES	1	1	1
DOESN	1	1	1
DOING	1	1	1
DON	1	2	2
DONE	1	1	1
DR	1	3	3
DRAW	1	1	1
DRIVING	1	1	1
EACH	1	1	1
EITHER	1	2	2
END	1	1	1
ENJOYED	1	1	1
ENJOYING	1	1	1
EVEN	1	2	2
EXAMPLE	1	8	8
EXPERIMENT	1	1	1
FACTORIES	1	4	4
FARMING	1	1	1
FINISH	1	1	1
FINISHED	1	2	2
FIRST	1	2	2
FOR	1	6	6
FORCING	1	1	1
FRESHER	1	1	1
FRIENDS	1	1	1
FROM	1	2	2
FUTURE	1	2	2
GAME	1	1	1
GET	1	2	2
GETTING	1	3	3

unitone_range

WEEK	1	3	3
WELL	1	1	1
WHEN	1	1	1
WHICH	1	1	1
WHO	1	1	1
WHY	1	11	11
WILL	1	12	12
WINDOWS	1	1	1
WITH	1	3	3
WON	1	7	7
WORK	1	2	2
WORLD	1	1	1
WOULD	1	1	1
WRITE	1	1	1
YEAR	1	1	1
YES	1	2	2
YET	1	2	2
YOU	1	16	16

Types Found In Base List Two

TYPE	RANGE	FREQ	F1
AFRAID	1	4	4
AFTERNOON	1	3	3
AGRICULTURE	1	1	1
ATTEND	1	1	1
BATH	1	1	1
BREATHING	1	1	1
BROWN	1	1	1
BUS	1	1	1
CENTURY	1	2	2
CHIMNEYS	1	1	1
CLEAN	1	2	2
CLOCK	1	1	1
COMPOSITION	1	1	1
COPPER	1	1	1
CRIME	1	1	1
DINNER	1	1	1
DOUBLED	1	1	1
EXPLODED	1	1	1
FILM	1	2	2
FLATS	1	1	1
FOOTBALL	1	2	2
FRUIT	1	1	1
HARBOUR	1	1	1
HOTEL	1	1	1
IMPROVE	1	1	1
LESSON	1	2	2
LOT	1	2	2
LUNCH	1	4	4
MANAGER	1	3	3
MATCH	1	1	1
MEAT	1	1	1
MEND	1	1	1
METRES	1	1	1
MINERALS	1	1	1
MODEL	1	2	2
PAIR	1	1	1
PARKS	1	1	1
PERFORM	1	1	1
PLANE	1	1	1
POURED	1	1	1
PRACTISE	1	1	1
PROBABLY	1	2	2
RAPIDLY	1	1	1
RAW	1	1	1
REPLACE	1	1	1

	unitone_range		
RESPONSIBLE	1	1	1
SENTENCE	1	2	2
SHOPS	1	1	1
SOLUTIONS	1	1	1
SOLVE	1	3	3
SOLVED	1	2	2
SORRY	1	1	1
SOUP	1	1	1
SWIMMING	1	1	1
TELEPHONE	1	2	2
TOMORROW	1	5	5
TOWER	1	1	1
TRIP	1	1	1
WASTES	1	1	1
WIRE	1	1	1

Types Found In Base List Three

TYPE	RANGE	FREQ	F1
AID	1	1	1
CHART	1	1	1
CLASSICAL	1	3	3
CONTRIBUTED	1	1	1
CONVERTED	1	1	1
COUPLE	1	1	1
DECADES	1	1	1
DESIGNERS	1	1	1
ELEMENT	1	1	1
ELEMENTS	1	2	2
EXPERTS	1	3	3
TRANSPORT	1	2	2

LIST OF FAMILY GROUPS

BASE ONE FAMILIES	RANGE	TYFREQ	FAFREQ	F1
A	1	198	199	199
ABLE	1	4	4	4
ABOUT	1	3	3	3
AGO	1	1	1	1
AGREE	1	0	1	1
AIR	1	4	4	4
ALL	1	2	2	2
ALONG	1	1	1	1
ALREADY	1	3	3	3
AND	1	25	25	25
ANSWER	1	1	1	1
ANY	1	0	2	2
AS	1	2	2	2
AT	1	14	14	14
BACK	1	1	1	1
BE	1	23	205	205
BECAUSE	1	1	1	1
BEFORE	1	2	2	2
BEGIN	1	0	1	1
BELIEVE	1	3	3	3
BELOW	1	2	2	2
BEST	1	0	1	1
BIG	1	0	2	2
BOOK	1	1	2	2
BUILD	1	1	10	10
BURN	1	0	1	1
BUT	1	1	1	1
BY	1	3	3	3
CAN	1	10	10	10
CARE	1	0	2	2

	unitone_range			
WAY	1	2	3	3
WE	1	16	16	16
WEEK	1	3	3	3
WELL	1	1	1	1
WHEN	1	1	1	1
WHICH	1	1	1	1
WHO	1	1	1	1
WHY	1	11	11	11
WILL	1	12	20	20
WIN	1	0	7	7
WINDOW	1	0	1	1
WITH	1	3	3	3
WORK	1	2	2	2
WORLD	1	1	1	1
WOULD	1	1	1	1
WRITE	1	1	1	1
YEAR	1	1	1	1
YES	1	2	2	2
YET	1	2	2	2
YOU	1	16	16	16

BASE TWO FAMILIES	RANGE	TYFREQ	FAFREQ	F1
AFRAID	1	4	4	4
AFTERNOON	1	3	3	3
AGRICULTURE	1	1	1	1
ATTEND	1	1	1	1
BATH	1	1	1	1
BREATHE	1	0	1	1
BROWN	1	1	1	1
BUS	1	1	1	1
CENTURY	1	2	2	2
CHIMNEY	1	0	1	1
CLEAN	1	2	2	2
CLOCK	1	1	1	1
COMPOSE	1	0	1	1
COPPER	1	1	1	1
CRIME	1	1	1	1
DINNER	1	1	1	1
DOUBLE	1	0	1	1
EXPLODE	1	0	1	1
FILM	1	2	2	2
FLAT	1	0	1	1
FOOT	1	0	2	2
FRUIT	1	1	1	1
HARBOR	1	0	1	1
HOTEL	1	1	1	1
IMPROVE	1	1	1	1
LESSON	1	2	2	2
LOT	1	2	2	2
LUNCH	1	4	4	4
MANAGE	1	0	3	3
MATCH	1	1	1	1
MEAT	1	1	1	1
MEND	1	1	1	1
METRE	1	0	1	1
MINERAL	1	0	1	1
MODEL	1	2	2	2
PAIR	1	1	1	1
PARK	1	0	1	1
PERFORM	1	1	1	1
PLANE	1	1	1	1
POUR	1	0	1	1
PRACTISE	1	1	1	1
PROBABLE	1	0	2	2
RAPID	1	0	1	1
RAW	1	1	1	1
REPLACE	1	1	1	1

	unitone_range			
RESPONSIBLE	1	1	1	1
SENTENCE	1	2	2	2
SHOP	1	0	1	1
SOLVE	1	3	6	6
SORRY	1	1	1	1
SOUP	1	1	1	1
SWIM	1	0	1	1
TELEPHONE	1	2	2	2
TOMORROW	1	5	5	5
TOWER	1	1	1	1
TRIP	1	1	1	1
WASTE	1	0	1	1
WIRE	1	1	1	1

BASE THREE FAMILIES	RANGE	TYFREQ	FAFREQ	F1
AID	1	1	1	1
CHART	1	1	1	1
CLASSIC	1	0	3	3
CONTRIBUTE	1	0	1	1
CONVERT	1	0	1	1
COUPLE	1	1	1	1
DECADE	1	0	1	1
DESIGN	1	0	1	1
ELEMENT	1	1	3	3
EXPERT	1	0	3	3
TRANSPORT	1	2	2	2

Types Not Found In Any List

FICTION	1	1	1
FUSES	1	1	1
KANGAROO	1	1	1
PHYSICS	1	2	2
REHEARSE	1	1	1
SUITCASES	1	2	2
SURGEON	1	1	1
TV	1	1	1
VOLCANO	1	1	1

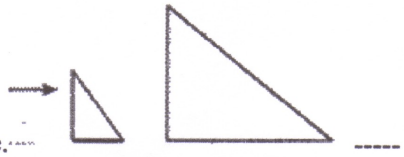
```
 time taken was : 1 Seconds

Number of cache nodes read:                132122
Number of cache nodes written:              61474
Number of disk nodes read:                      0
Number of disk nodes written:                   0
Number of nodes per second,                     0
Total Number of words,                      12459
Number of words per second,                 12459
Number of unique words in tree,               777
Number of unique words per second,            777
Memory used,                                 1.86 MB
Size of node,                                 316 bytes
Memory used by nodes,                        0.23 MB


 ...Finished
```

Appendix:3 1000 Word Level Test

VOCABULARY TEST: 1000 LEVEL TEST A

Instructions: There are 39 questions. Write"T" if a sentence is true. Write "N" if a sentence is not true. Write "X" if you do not understand the sentence. The first one has been answered for you.

We cut time into minutes, hours, and days. T

1. This one is little.  -----

2. You can find these everywhere.

3. Some children call their mother Mama

4. *Show me the way to do it* means 'show me how to do it.'

5. This country is part of the world.

6. This can keep people away from your house.

7. When something falls, it goes up.

8. Most children go to school at night.

9. It is easy for children to remain still.

10. One person can carry this.

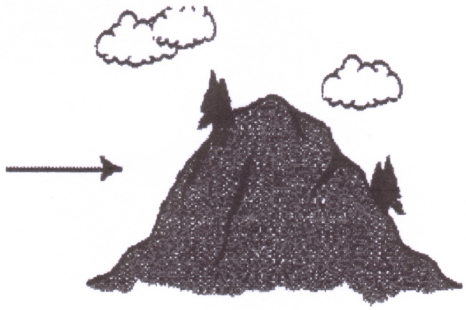
11. A scene is part of a play.

12. People often think of their home, when they are away from it.

13. There is a mountain in every city.

14. Every month has the same number of days.
15. A chief is the youngest person in a group.
16. Black is a colour.
17. You can use a pen to make marks on paper.
18. A family always has at least two people.

19. You can go by road from London to New York.
20. Silver costs a lot of money.



21. This is a hill.

22. This young person is a girl.

23. We can be sure that one day we will die.
24. A society is made up of people living together.
25. An example can help you understand.
26. Some books have pictures in them.
27. When some people attack other people, they try to hurt them.
28. When something is ancient, it is very big.
29. Big ships can sail up a stream.
30. It is good to keep a promise.
31. People often dream when they are sleeping.
32. This is a date - 10 o'clock.
33. When something is impossible, it is easy to do it.
34. Milk is blue.
35. A square has five sides.
36. Boats are made to travel on land.
37. Cars cannot pass each other on a wide road.
38. When you look at something closely, you can see the details.

39. This part is a handle.

A Vocabulary Levels Test

This is a vocabulary test. You must choose the right word to go with each meaning. Write the number of that word next to its meaning. Here is an example.

1 business	
2 clock	_____ part of a house
3 horse	_____ animal with four legs
4 pencil	_____ something used for writing
5 shoe	
6 wall	

You answer it in the following way.

1 business	
2 clock	__6__ part of a house
3 horse	__3__ animal with four legs
4 pencil	__4__ something used for writing
5 shoe	
6 wall	

Some words are in the test to make it more difficult. You do not have to find a meaning for these words. In the example above, these words are *business, clock, shoe*

Try to do every part of the test.

THE 2,000 WORD LEVEL

1 copy	
2 event	_____ end or highest point
3 motor	_____ this moves a car
4 pity	_____ thing made to be like another
5 profit	
6 tip	

1 accident
2 debt
3 fortune
4 pride
5 roar
6 thread

_____ loud deep sound
_____ something you must pay
_____ having a high opinion of yourself

1 birth
2 dust
3 operation
4 row
5 sport
6 victory

_____ game
_____ winning
_____ being born

1 clerk
2 frame
3 noise
4 respect
5 theatre
6 wine

_____ a drink
_____ office worker
_____ unwanted sound

1 dozen
2 empire
3 gift
4 opportunity
5 relief
6 tax

_____ chance
_____ twelve
_____ money paid to the government

1 admire
2 complain
3 fix
4 hire
5 introduce
6 stretch

_____ make wider or longer
_____ bring in for the first time
_____ have a high opinion of someone

1 arrange
2 develop
3 lean
4 owe
5 prefer
6 seize

_____ grow
_____ put in order
_____ like more than something else

1 blame
2 elect
3 jump
4 manufacture
5 melt
6 threaten

_____ make
_____ choose by voting
_____ become like water

ÖZGEÇMİŞ

Melike Agan 1982 yılında İstanbul'da doğdu. Uludağ Üniversitesi İngilizce Öğretmenliği Bölümü'nden 2003 yılında mezun oldu. İleri derecede İngilizce bilen Melike Agan, 2003 yılından beri devlet okullarında İngilizce öğretmeni olarak görev yapmaktadır. Evlidir ve Bursa'da yaşamaktadır.

VITAE

Melike Agan was born in İstanbul, in 1982. She graduated from the. Depatment of English Language Teaching, Uludağ University in 2003. Her English skill is advanced and she has been working as an English teacher at state high schools since 2003 . She is married and and she lives in Bursa.