GAZİ ÜNİVERSİTESİ
Gazi
1926

A CORPUS STUDY OF ACADEMIC TURKISH VOCABULARY: IMPLICATIONS FOR TEACHING TURKISH AS A FOREIGN LANGUAGE

Mustafa DOLMACI

Ph.D. DISSERTATION

DEPARTMENT OF ENGLISH LANGUAGE TEACHING

GAZI UNIVERSITY

INSTITUTE OF EDUCATIONAL SCIENCES

NOVEMBER, 2015

COPYRIGHT AND CONSENT TO COPY THE DISSERTATION

AUTHOR:

Name: Mustafa

Last name: Dolmacı

Signature:

Date of delivery: November, 2015

DISSERTATION:

Title of dissertation in Turkish: "Akademik Türkçe Kelime Bilgisi Üzerine Bir Derlem Çalışması: Yabancı Dil Olarak Türkçe Öğretimine Dair Çıkarımlar"

Title of dissertation in English: "A Corpus Study of Academic Turkish Vocabulary: Implications For Teaching Turkish As A Foreign Language"

DECLARATION OF CONFORMITY TO ETHICS

I declare that I have complied with the scientific ethical principles within the process of typing the dissertation that all the citations are made in accordance with the principles of citing and that all the other sections of the study belong to me.

Name and last name of the author: Mustafa Dolmacı

Signature of the author:

Dissertation Approval

We certify that the dissertation entitled "A Corpus Study of Academic Turkish Vocabulary: Implications for Teaching Turkish as a Foreign Language" prepared by Mustafa Dolmacı has been unanimously found satisfactory by the jury for the degree of doctorate of philosophy in the subject matter of English language teaching at Gazi University, Department of English Language Teaching.

Supervisor   Assist. Prof. Dr. Abdullah ERTAŞ
                School of Foreign Languages, Atılım University

Chairman   Prof. Dr. Abdülvahit ÇAKIR
                ELT Department, Gazi University

Member   Assoc. Prof. Dr. Nurdan ÖZBEK GÜRBÜZ
                ELT Department, METU

Member   Assoc. Prof. Dr. Kadriye Dilek AKPINAR
                ELT Department, Gazi University

Member   Assist. Prof. Dr. Cemal ÇAKIR
                ELT Department, Gazi University

Date of dissertation defence: 09/11/2015

I certify that this dissertation has complied with the requirements of degree of Doctorate of Philosophy in the subject matter of English Language Teaching.

Prof. Dr. Servet Karabağ   ....

Director of Institute of Educational Sciences

iii

ACKNOWLEDGEMENTS

AKADEMİK TÜRKÇE KELİME BİLGİSİ ÜZERİNE BİR DERLEM ÇALIŞMASI: YABANCI DİL OLARAK TÜRKÇE ÖĞRETİMİNE DAİR ÇIKARIMLAR

(Doktora Tezi)

Mustafa DOLMACI

GAZİ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Kasım 2015

ÖZ

Kelime bilgisi, diğer dil becerilerinin üzerine inşa edildiği sağlam bir temel oluşturur. Wilkins (1972) dilbilgisi olmadan çok az şey aktarılabileceğini, fakat kelime bilgisi olmadan hiçbir şey aktarılamayacağını ileri sürmektedir. Bu yüzden, bu çalışma akademik Türkçe kelime bilgisine ve aynı zamanda Türkçe öğrenen herhangi bir kişinin akademik kelime öğrenmeye başlamadan önce bilmesi gereken temel kelime bilgisine odaklanmaktadır. İlk olarak, 2000 kelime içeren *Genel Türkçe Kelime Listesi* oluşturulmuştur. İkinci olarak, lisans düzeyindeki ders kitaplarından derlenen bir *Akademik Türkçe Derlem* ve bu derlemden 1010 kelime içeren *Akademik Türkçe Kelime Listesi* elde edilmiştir. Üçüncü olarak, akademik derlemden ve akademik kelime listesinden faydalanılarak *Akademik Kelime Testi* geliştirilmiştir. *Akademik*

Kelime Testi, öğretim dili Türkçe olan 6 farklı üniversitenin Türkçe Eğitim Merkezlerinde Türkçe öğrenen 196 öğrenci tarafından cevaplanmıştır. Son olarak, akademik Türkçe kelime bilgilerine yönelik fikirlerini ortaya çıkarmak için öğretim dili Türkçe olan üniversitelerde öğrenim görmekte olan 41 anadili Türkçe olmayan lisans öğrencisiyle görüşme yapılmış ve katılımcıların fikirlerini ayrıntılarıyla açıklamak için nitel veri analiz edilmiştir. Araştırma sonuçları şunları göstermektedir: 1) *Genel Türkçe Kelime Listesi*nin ilk 1000 kelimesi oluşturulduğu derlemin yaklaşık yarısını kapsamaktadır, 2) *Akademik Türkçe Kelime Listesi*, *Akademik Derlem*in yaklaşık üçte birini oluşturmaktadır bu yüzden akademik Türkçe derslerine dahil edilmelidir, 3) Katılımcıların *Akademik Derlem*e dair genel kelime bilgisi düzeyleri *yeterli anlama düzeyi* (Hu and Nation, 2000) için düşüktür, 4) Öğrenciler Türkçe kullanarak sınıf içinde ve dışında iletişim kurmada kısmen başarılı olmalarına rağmen, Türkçe öğretim yapılan bir üniversitede öğrenci olmalarına dair problemleri Türkçeye hakim olmalarındaki zayıflıklarından, psikolojik ve bireysel farklılıklardan kaynaklanmaktadır. Bu çalışma akademik amaçlı Türkçe alanında *Akademik Türkçe Kelime Listesi* sunmasıyla alanında benzersizdir.

Anahtar kelimeler: Genel Türkçe kelime listesi, akademik Türkçe kelime listesi, akademik derlem, akademik kelime testi.

Sayfa sayısı: 195

Danışman: Yrd. Doç. Dr. Abdullah ERTAŞ

A CORPUS STUDY OF ACADEMIC TURKISH VOCABULARY: IMPLICATIONS FOR TEACHING TURKISH AS A FOREIGN LANGUAGE

(PhD Dissertation)

DOLMACI, Mustafa

GAZI UNIVERSITY
GRADUATE SCHOOL OF EDUCATIONAL SCIENCES

November 2015

ABSTRACT

Vocabulary knowledge is solid bedrock, upon which other language skills can be built. Wilkins (1972) asserted that "without grammar very little can be conveyed, without vocabulary nothing can be conveyed" (p. 111). Therefore, this study focuses on academic Turkish vocabulary and also aims to investigate the basic vocabulary that any Turkish language learner should have before starting to learn academic vocabulary. First, *the General Turkish Wordlist* containing 2000 words was developed. Second, *The Academic Corpus* derived from undergraduate textbooks and *the Academic Turkish Wordlist* containing 1010

words developed from *the Academic Corpus* were acquired. Third, *the Academic Vocabulary Test* was developed utilizing *the Academic Corpus* and *the Academic Wordlist*. *The Academic Vocabulary Test* was answered by 196 students learning Turkish at the Turkish Teaching Centers offering Turkish as the medium of instruction. Lastly, 41 undergraduate non-native Turkish students studying at universities offering Turkish medium instruction were interviewed in order to discover opinions of foreign students on their knowledge of academic Turkish vocabulary and the qualitative data from interviews was analyzed to elaborate the opinions of the participants. The results showed that 1) the first 1000 words of *the General Turkish Wordlist* covered nearly half of the corpus it was generated from, 2) *the Academic Turkish Wordlist* covered nearly one third of *the Academic Corpus,* so it should be included in academic Turkish courses, 3) the overall vocabulary knowledge level of the participants regarding *the Academic Corpus* was low for *the sufficient comprehension level* (Hu & Nation, 2000), 4) Although, students are more or less successful in communicating through Turkish language in and out of the classroom atmosphere, their problems regarding being a student studying at a university with Turkish as the medium of instruction mainly result from having a low mastery of Turkish and from psychological or individual factors. This study is unique since it will be the first research study suggesting *the Academic Turkish Wordlist* in the field of Turkish for academic purposes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Introduction

Since the pioneering research of Michael West (1953), in which he presented a list of roughly 2000 words called The General Service List (GSL), there have been many attempts to make up different corpora like the British National Corpus, Corpus of Contemporary American English, Bank of English or Australian Corpus of English. Similarly, many academic or general wordlists have been published in English or in other languages (Academic Vocabulary List [Gardner & Davies, 2013], Academic Wordlist [Coxhead, 2000], New General Service List [Browne, 2014]). On the other hand, the lack of a Turkish General and Academic Wordlists which meet the academic criteria such as being collected from a valid corpus has made this research necessary in the field of teaching Turkish as a foreign language.

In consideration of the things mentioned above, the purpose of the present study is to form a General Turkish Wordlist (GTW) and an Academic Turkish Wordlist (ATW) from an academic corpus which has been made up from university textbooks and to discover whether foreign students learning Turkish in order to study at undergraduate and graduate programs in Turkey need academic vocabulary support in order to fully pursue their courses in Turkish.

Compared to Teaching English as a Foreign Language (TEFL), Teaching Turkish as a Foreign Language (TTFL) is on the bottom rung of the ladder and Çakır's (1989) dissertation was the first phd study on teaching of Turkish as a foreign language and Karababa (2009) points out that the most important problem in TTFL is the inadequate curricula and studies and states that in order to solve this problem, there is a need for a suitable curriculum and books and

materials prepared in accord with this curriculum and the level of the students. This situation is more crucial when teaching Academic Turkish is addressed due to the lack of sufficient course materials in terms of teaching Academic Turkish.

Brown (2001) identified language for academic purposes as a term "that is very broadly applied to any course, module, or workshop in which students are taught to deal with academically related language and subject matter" (p. 123) and it aims to raise the learners to a point where they are capable of getting the most out of their coming course of study (Scrivener, 2011, p. 315).

Vocabulary knowledge allows students to use the language, and language use fosters the increase of vocabulary knowledge. While it is a known fact that vocabulary knowledge helps learners in every language skill act effectively, it should not be forgotten that having sufficient vocabulary does not make everything a piece of cake (Nation & Waring, 1997).

Foreign language learners already have the concepts in their minds and are able to produce these concepts using their mother tongue. Learners have great difficulty in expressing these concepts or ideas in another language mainly because of the linguistic structure of the target language. Knowing the equivalents of utterances in the target language is a prerequisite followed by presenting them using the appropriate grammar form and vocabulary learning is seen as the most important and time consuming activity by language learners although it is only one component of language skills.

Currently, descriptive analyses of corpora largely inform not only vocabulary teaching, which is a prominent part of language teaching, but also teaching materials such as coursebooks, which are the main source of foreign language teaching courses. Moreover, language learners, teachers and also coursebook writers need a comprehensive and authentic language database to present information in addition to what they could gather from reference books. At this point, a well prepared corpus can help them with its frequency list or concordance outputs.

Statement of The Problem

According to Hu and Nation (2000), the minimum amount of vocabulary coverage for a comprehensible reading of a text is clearly above 80%. Furthermore, if the percentage of known words in a text is 98%, comprehension and incidental vocabulary learning through reading are likely to increase (Nation, 2001).

The problem of this research is whether the vocabulary instruction applied in Turkish Teaching Centers (TTC) of universities in Turkey is efficient enough for students who have learned Turkish as a foreign language to study at universities offering Turkish courses. In line with this situation, another question arises: What is the general Turkish vocabulary wordlist that should be aimed for a general Turkish course and what is the academic Turkish wordlist that students need to acquire for their undergraduate programs?

There were three coursebooks used at the Turkish Teaching Centers of Selçuk, Gazi, Konya Necmettin Erbakan, Karatay, Düzce, Niğde and Abant İzzet Baysal Universities, where the Academic Vocabulary Test made up from the Academic Turkish Wordlist created in the current study was administrated. These coursebooks were: 1) İstanbul Üniversitesi Dil Merkezi İstanbul Seti (İstanbul Series), 2) Gazi Üniversitesi TÖMER Yabancılar İçin Türkçe Seti (Gazi Series), 3) Ankara Üniversitesi TÖMER Yeni Hitit Yabancılar İçin Türkçe Öğretim Seti (Ankara New Hitit Series). It is automatically assumed that these coursebooks would provide enough input for the students in order to follow their courses in Turkish at the universities in Turkey and to use Turkish for general purposes. Although these coursebooks are not intended for academic purposes, it is important to find out their usefulness on university education since most of the students learning Turkish as a foreign language at Turkish Teaching Centers aim to study at Turkish Universities which offer Turkish medium instruction. These coursebooks are supposed to prepare students for academic studies because students complete the C1 level (Effective Operational Proficiency) before they finish their courses. According to Common European Framework (Council of Europe, 2001), a student who is at the C1 level "can use language flexibly and effectively for social, academic and professional purposes" (p. 24) Therefore, another question is how much ready these students are for their undergraduate programs in terms of vocabulary knowledge.

Significance of The Study

The current study is significant for several reasons. This pioneering study focuses on academic Turkish vocabulary and it also aims to investigate the basic vocabulary that any Turkish language learner should have before starting to learn academic vocabulary. None of the earlier studies has tried to develop an academic Turkish corpus or an academic Turkish wordlist and evaluated its efficacy for the students studying Turkish as a foreign language. Therefore, this study is unique since it will be the first research study in the field of Turkish for academic purposes which suggests: a) a general Turkish wordlist, b) an academic Turkish wordlist, and c) a research on the knowledge level of foreign students studying Turkish regarding academic Turkish vocabulary.

As the creator of the AWL (Academic Wordlist), Coxhead (2000) points out:

> The use of this research method, however, does not imply that language teaching and learning should rely on decontextualised methods. Instead, the AWL might be used to set vocabulary goals for EAP courses, construct relevant teaching materials, and help students focus on useful vocabulary items (p. 227).

Additionally, Gardner & Davies (2013) state that "any well-conceived list of high-coverage words brings some order to what otherwise would be vocabulary chaos (Where do we start? What can our learners focus on now, next, etc.?)" (p. 6). Therefore, it is apparent that there is a need for a general Turkish wordlist and an academic Turkish wordlist in order to select the words to teach and develop the contents of courses and evaluate students' products in Turkish, and this study will fill this gap in the field of teaching Turkish as a foreign language and it should also be noted here that this study will be the first attempt to create an academic wordlist in Turkish for foreign undergraduate students. The wordlists developed in this study will also be made available for further research in the field.

Purpose of The Study

The lack of a general Turkish wordlist that includes the basic vocabulary any learner should have before starting to learn academic vocabulary makes it necessary to make up a general

Turkish wordlist that consists of more than half of the general texts in Turkish in the first phase of the study (the GSL covers up to 75% of nonfiction texts [Hwang, 1989], and up to 76% of the Academic Corpus [Coxhead, 2000]). Therefore, it can be said that this General Turkish Wordlist (or some parts of it) is the one that students learning Turkish as a foreign language should learn before they start their academic studies. We needed a general Turkish wordlist in order to achieve an academic wordlist using Coxhead's research design in which she excluded the first 2,000 most frequently occurring words of English, as represented by West's (1953) GSL (in our study first most frequent 300 words from the General Turkish Wordlist are excluded). This wordlist could also give language practitioners the idea of where to start or which words should be taught first regarding general Turkish vocabulary acquisition studies.

In the second phase of the study, an academic Turkish wordlist was formed. In order to achieve this, an academic corpus was derived from the E-books on the E-Learning Portal of Anadolu University. It was preliminarily assumed that the Academic Turkish Wordlist would account for 10% of the total tokens in the Academic Corpus. 10% of the total tokens was targeted since Coxhead (2000) points out that "the AWL (the Academic Wordlist in English) contains 570 word families that account for approximately 10.0% of the total words (tokens) in academic texts." (p. 213)

In the third phase of the study, the Academic Vocabulary Test consisting of 30 items chosen from the Academic Turkish Wordlist was answered by 208 participants studying at the Turkish Teaching Centers of Selçuk, Gazi, Konya Necmettin Erbakan, Karatay, Düzce, Niğde and Abant İzzet Baysal Universities.

Last but not the least, 41 undergraduate students who had previously studied Turkish as a foreign language at the Turkish Teaching Centers were interviewed in order to discover their opinions on their Turkish language competency.

The aim of implementing the Academic Vocabulary Test and holding an interview with undergraduate foreign students was to find out whether students studying Turkish as a foreign language with a goal of attending an undergraduate or graduate program need additional

academic Turkish vocabulary instruction before they start studying at a university with Turkish as the medium of instruction.

Research Questions

The current study addresses the following research questions:

1. What is the basic Turkish vocabulary that could be aimed at before starting to learn academic vocabulary in terms of Turkish as a foreign language?

2. What are the most frequent academic Turkish words that need to be known in order to be able to follow courses at universities offering Turkish medium instruction in Turkey?

3. To what extent do students studying Turkish at the Turkish Teaching Centres know the words for academic Turkish?

4. Do foreign undergraduate students who studied Turkish at the Turkish Teaching Centers need academic Turkish vocabulary instruction to improve their Turkish vocabulary knowledge so that they can follow academic courses?

Definitions of Some Key Concepts

Corpus: A corpus is a body of written text or transcribed speech which can serve as a basis for linguistic analysis and description (Kennedy, 1998, p. 1).

Academic Corpus: A corpus derived from journal articles, book chapters, textbooks, laboratory manuals, and course notes that are representative of the academic genre and written for an academic audience.

The General Turkish Wordlist: The list of most frequent 2000 Turkish words (lemmas) that includes the basic vocabulary that any learner should have. In this study, the General Turkish Wordlist was developed from the most frequent 11000 words of the Turkish National Corpus (Aksan et al. 2012).

The Academic Turkish Wordlist: The first academic Turkish wordlist developed from the Academic Turkish Corpus compiled from Turkish undergraduate textbooks in this study.

Token: Token is a "word" within a corpus. It is used most often to talk about word count and the size of a corpus (Tang, 2015).

Type: A unique word form in a corpus. Types are placed in a word list arranged most often in order of frequency or alphabetical order, and usually shown with frequency count (Tang, 2015).

Lemma: "A set of lexical forms having the same stem and belonging to the same major word class, differing only in inflection and/or spelling" (Francis & Kucera, 1982:1)

Word Family: "A word family is defined as the base form of a word plus its inflected forms (third person –s, -ed, -ing, plural –s, possessive –s, comparative –er, and superlative –est) plus derived forms made from certain uses of the following affixes (-able, -er, -ish, -less, -ly, -ness, -th, -y, non-, un- , -al, -ation, -ess, -ful, -ism, -ist, -ity, -ize, -ment, in-)" (Hirsh & Nation, 1992, p.692). Inflected and regularly derived forms of a known base word are thought to be known words if the learners have a good knowledge of the affixes.

Textbook: A book used as a standard work for the study of a subject. They are generally produced according to the demands of educational institutions. In the current study, "textbook" refers to the books that have been written to be used at the undergraduate programs.

Coursebook: A Coursebook is a book that is printed for the use of students and teachers as they study a particular subject and it is the primary classroom material in many foreign language classrooms.

CHAPTER 2

REVIEW OF LITERATURE

A Concise History of Corpus Linguistics In Relation With SLA

Since the late 1980's, corpora, which is defined as the bodies of written or spoken texts, have been considered in the field of linguistics. Since the advancement of technology, their use in research has only grown and more than ever. With the assistance of technology and software programs its data is effortlessly and easily transcribed, compiled, and analyzed. The corpora can also be specifically designed for its particular usage. Corpus has been defined in a number of ways by different scholars: "bodies of texts assembled in a principled way" (Johansson, 1995, p. 19), "a collection of texts, written or spoken, usually stored in a computer database" (McCarthy, 2004, p. 1) and "a collection of texts based on a set of design criteria, one of which is that the corpus aims to be representative" (Cheng, 2012, p.3). Over its relatively short life corpus linguistics has experienced many changes. In comparison with other sub-fields of linguistics, Bonelli (2012) cites a previous work (Bonelli & Sinclair, 2006) which classifies the history of corpus linguistics into three generations:

1. from 1960 to 1980, there were no electronic materials and a corpora which contained up to a million words;
2. from 1980 to 2000, the invention of a new method and device called "scanner" which allowed the investigators to build corpora containing up to twenty million words;
3. and from 2000 to the present, in which unlimited amounts of corpus data has been made available; (p. 208)

Advances in technology have also allowed scholars to have virtual access to the information needed for corpora studies. Moreover, corpus linguistics began as a methodology for

describing natural language data since the impetus behind corpus linguistics has changed with the times as well, largely due to the rise of technology and the quantities of information available. The methodology has also taken a different shape. It has changed from a top-down to a bottom-up approach. In doing so, the data (e.g. corpora) are already given and the researcher describes the patterns of usage that already exist. This is different from other research methodologies in linguistics, in which researchers set out to gain certain data from their participants or target discourses.

Bonelli (2012) describes the transition from a methodological approach (mainly qualitative) in corpus linguistics to a theoretical one that has happened thanks to a quantitative explosion that has "shaken the underlying assumptions behind many well-established theoretical positions in the field" (p. 17). To put it in a nutshell, the expansive improvements in technology have allowed corpus linguistics to literally explode as a field, and in exchange, researchers are now capable of observing patterns of change in language use through accessing huge amounts of data that have not been observed before.

Corpus research has generally been growing since the 1960's. Granger (2002) discusses that the kind of corpus research which included the Second Language Acquisition (SLA) has only been active since the 1980's. It is designated as a way corpus linguistics can be used to provide evidence of how speakers use language. Before the subject started to rise in popularity of usage, corpus linguistics and second language research were not connected at all and had never been incorporated in regards with research. However, later on, the SLA research in the field of corpus linguistics proved that techniques used within corpus linguistics could be trustworthy and valid devices for analysis.

Methodological approaches to SLA elicit data from generally a small number of texts or resources which can be participants in some research studies. Corpus linguistics employs the kind of language as it is used naturally in speech and writing from many participants or textbooks from any resources at any level of discourse. 'Learner corpora' are found either in qualitative and quantitative analysis, both of which are of primary significance in the world of empirical research. To put it in detail, a learner corpus is built from data gathered from different learners learning different languages. In these corpora, scholars tend to analyze the

10

language that learners produce in natural ways and situations, and thus they can attempt to arrive at conclusions from a large cloud of data about such subject as patterns of usage and frequency of occurrence of a particular linguistic item. This can take place in both written and oral basis.

Most learner corpora are cross-sectional, providing data at one point in time. Longitudinal and quasi-longitudinal data are much less common, as it is difficult and time-consuming to collect these across time. Despite the short history of corpus linguistics, if the data are collected locally, the type of data gathered from learner corpora can facilitate to inform classroom teaching and help to improve instructional methods.

One of the main usages of learner corpora is in the field of second language pedagogy (O'Keeffe, McCarthy, and Carter, 2007). With the assistance of corpus data, applied linguists with particular interest in second language education have started to explain how learners acquire and use language. These kinds of analyses of learner corpora can result in important implications in how second language development is taking place and how it is identified, including the language used in classrooms and the academic discourse which emerges in writing activities as well as spoken exercises. O'Keeffe et al. (2007) maintain that the study of corpora could change long-held notions of education and pedagogy. In the same fashion, it can also bridge the gap between the science of cognition in linguistics and many other areas including teaching, structural studies, translation techniques and sociolinguistics.

Research findings from the previous studies have been successful enough to bridge this gap. Amy Tsui, in *How to Use Corpora in Language Teaching* (Sinclair, 2004), explains some of the prominent studies done in corpus linguistics, which have investigated the following four fields: lexical collocation, genre analysis, syntactic patterning, and discourse structure. The mentioned research studies include analyses conducted with corpora such as the Brown Corpus (Francis & Kucera, 1982) and Carter & McCarthy's Spoken English Corpus (Carter & McCarthy, 1997).

It is worth noting that any learner corpus consists of a framing and classifying of learner language or interlanguage (Selinker, 1972). Interlanguage (IL) can be defined as a system of language used by second language learners that activates them to gain a sense from the second

language, and which resonates the learners' momentary guesses and hypotheses about the target language. An interlanguage, thus, shows a framed and regular rule governed use of language. Although the language is non-native-like, it includes errors which reflect natural processes of language acquisition. Transfer from the first language or overgeneralization of certain second language rules (Selinker, 1972) can be named from among those mistakes and errors.

The more the understanding of the linguists increases and the more they make sense of a range of concepts of interlanguage, the less unlighted second language instruction will be. This understanding and process would lead to producing more proficient and 'native-like' language learners in classrooms.

To give an example of an effective way of producing native-like language learners, the use of prefabricated patterns can be addressed. They are also called multiword items (MWIs), in general also referred to as 'multiword expressions' (Baldwin, 2005). Cobb (2003) states that MWIs permit the speaker to plan ahead and cognitively get ready for a discourse. In addition, MWIs in general is that they are idiosyncratic in their lexicon, syntax, semantics, as well as pragmatics. They give the meaning that they are due when assumed as a whole and they are not meaningful when the words are translated individually. The meanings are always the same and they carry predictable meanings and are predicted to be always put together. The euphemism "to kick the bucket", which conveys the semantic meaning of "to die" can be considered as an example.

The Concept of "Word" and Its Designation In Corpus Studies

Words are the core components of a language, and a large amount of word knowledge is fundamental for the mastery of a language. Therefore, 'word', as a key concept, needs to be defined accurately. In this context, word form, lemma, and word family are three relevant perceptions which need clarification.

Word Form, Lemma and Word Family

Word form, lemma and word family are three important concepts that shape corpus studies. While some studies such as Gardner and Davies (2013) use lemmas, others such as Coxhead (2000) use word family structure to produce pedagogical word lists:

Word form: it is any orthographical form of a word.

Word families: a word family consists of a root word and all its derivatives and inflections which can be understood by a learner without having to learn each form separately.

Lemma: it can be defined as the head word in a dictionary without any morphological changes.

Hence, in the following example, there are four word forms (*happy*, *happiness*, *happily*, and *unhappy*), four lemmas (the adjectives *happy* and *unhappy*, the noun *happiness*, and the adverb *happily*) and one word family (*happy*):

*When I saw her I thought that she was **happy** and the **happiness** in her face was clear and then we talked **happily** about the thing which had made me **unhappy** the day before that day.*

Thus, it becomes clear that 'word' can vary from treating all inflections having a single root, to treating whole orthographic word forms clearly separate and different. Thereof, in analyzing any text the limit of analysis must be elucidated. For this purpose, Bauer & Nation (1993) set up different levels "with a focus on recognition of written words":

Level 1: *Each form is a different word*. At this level the assumption is that learners consider 'book' and 'books' as two distinctive word forms. Gardner (2007) believes that this is a pessimistic view since the primary assumption is that the learners do not have knowledge about morphological relationships.

Level 2: *Inflectional Suffixes*. At this level, learners comprehend that words with the same base and inflections are members of the same word family.

Level 3: *The most frequent and regular derivational affixes*. Words are considered members of a single lexical unit provided that they have the same base and contain inflections and/or

derivational affixes. Suffixes such as *–able, -er, -ish, -less, -ly,-ness, -th,-y, non-,* and *un-* fall in this category.

Level 4: *Frequent, orthographically regular affixes.* At this stage, whether an affix is frequent is more important than whether it is productive or not. The same is true for orthographic criteria which are considered more crucial than phonological criteria.

Level 5: *Regular but infrequent affixes.* A number of less frequent affixes whose behavior is regular are welcome at this level. (e.g. *-age, -al, -an, -hood, -let, anti-, arch-, and bi*).

Level 6: *Frequent but irregular affixes.* Those affixes which cause problems of segmentation are included in this stage. Orthographic modifications to the base are necessary in order to let the affixation to occur. Affixes at this level are *-able, -ee, -ic,-ify, -ion, -ist, -ition, -ive, -th, -y, pre-,* and *re-*.

Level 7: *Classical roots and affixes.* As the title suggests, this stage belongs to the classical roots which are numerous in English. Also this level includes prefixes which are frequent, for example *ab-, ad-, com-, de- dis-, ex-,* and *sub-*. (p. 258-262)

So far, it has been believed that understanding inflectional and derivational morphology enables language learners to extract the underlying meaning from words without too much effort. A great support for this comes from the findings of a study by Nagy, Anderson, Schommer, Scott, and Stallman (1989). They examined the role of morphological relatedness in word recognition using a lexical decision task, and revealed the considerable impact of inflectional and derivational family frequency on the speed and accuracy of recognition of stems.

Mastery of a great deal of derivatives, though, is considered challenging. As a result, many scholars maintain that two factors play a pivotal role in deciding which unit of analysis must be used: 1) the aim of the word list (e.g., receptive vs. productive), and 2) the target user group's proficiency level and learning background.

Taking the goal of learners from manipulating word lists into account, it has been proposed that in order to maximize the amount of receptive vocabulary, a comprehensive definition of the word family is needed as the basis for the word lists (Coxhead, 2000; Nation, 2001; Nation & Webb, 2010).

On the other hand, for developing productive abilities, lemma, as a less comprehensive unit of analysis, is relevant. As claimed by Nation & Webb (2010), learners do not necessarily acquire various forms of a word by solely using one form of it.

Language learners' proficiency level is the second influential factor in the selection of unit of analysis. Knowing this, Gardner (2007) puts the emphasis on issues related to language learners and offers a brief amalgamation of research findings relevant to morphological acquisition:

1. *Base Forms + Regular Inflections*: younger children; low general English proficiency; low English literacy skills; no specific morphological training.
2. *Base Forms + Regular Inflections + Irregular Inflections + Derivational Prefixes*: older children and adolescents, intermediate general English proficiency, intermediate English literacy skills; some morphological training.
3. *Base Forms + Regular Inflections + Irregular Inflections + Derivational Prefixes + Derivational Suffixes (regular) + Derivational Suffixes (irregular)*: adults, high general English proficiency; high English literacy skills; extensive morphological training or experience (p. 258-262)

It can be inferred from Gardner's implications that his major idea was that the lower the language proficiency is and the less the morphological intake is, the lower expectations we should have regarding a learner's ability to distinguish and realize the relationships among word forms and families.

The most efficient word lists have been established on word families. The 5000-word list, which was detailed in Davies and Gardner's frequency dictionary of American English (2010), is one example of a contemporary list that is not based on word families. These 5000 words comprise lemmas, using Francis and Kucera's (1982) definition of lemma: "a set of lexical forms having the same stem and belonging to the same major word class, differing only in inflection and/or spelling" (p. 1). For example, the lemma *BASE* (n.) has the members *base*(n.) and *bases* (n.); the separate lemma *BASE* (v.) has the members *base* (v.), *bases* (v.), *basing* (v.), *based* (v.). In their introduction to this list, Gardner and Davies (2013) introduce lemma as the unit of analysis, based primarily on the collocational information we would lose from a more inclusive unit.

15

Schmitt (2010) offers a number of additional benefits of using the lemma as the unit of analysis, two of which are particularly relevant to this dissertation:

1. "The unit is relatively straightforward, which means that consumers of research studies will know what it means."
2. "It takes a lot of vocabulary to function in a language, and estimates based on word families may give the impression that less is necessary than is the case, especially as many consumers may simply interpret word family figures as 'words'." (p. 193)

As Gardner (2007) and Schmitt (2010) discuss, there are certain disagreements regarding the use of lemmas and word families. Schmitt reminds that, by deploying word families as the unit of analysis, we are considering that the students have a great deal of morphological competency. To give an example we are considering that a learner who comprehends *the word form* is able to realize the common root and properly use the meaning of suffixes/infixes/ prefixes to the following word family members: *former*, *formation*, and *formative*. This idea most probably overestimates knowledge, thus alludes the fact that the learners know more than you might be able to consider and faints the challenge that they should possess to learn more.

In addition to operationally defining the construct "word" and other lexical units of meaning, a great deal of work has gone into understanding the distributions of vocabulary in written texts and into classifying these words according to their use and distributions. There has been some variability in terms of classifications made for the vocabulary found in academic texts, as well as of the terms used for these distinctions, but, in general, a three-way distinction can be made. Nation (2001) uses the following terms:

1. High-frequency vocabulary
2. Specialized vocabulary
3. Low-frequency vocabulary

For the most part, the classification given above shares both qualitative and, in a general sense, as well as quantitative attributes across studies. Each classification can be seen as a level defined in large part by increasing distributional restrictions, specifically frequency (i.e., the number of occurrences in a corpus) and range (the number of texts or sections in which a word occurs), and in some cases, dispersion (the evenness of occurrence distribution across texts or

sections in a corpus) (e.g., Coxhead, 2000; Nation, 2004). At the least restricted level is high-frequency vocabulary. Less frequent vocabulary with a narrower range is specialized vocabulary, and even more restricted vocabulary is low frequency vocabulary. It is not the case, however, that as distributions become more restricted, we simply end up with a subset of a less restricted set. For example, specialized vocabulary is not simply a subset of general vocabulary. Rather, there is often a line drawn between classifications making them mutually exclusive. For example, specialized vocabulary is the set of items whose distribution is too narrow to meet the criteria of high-frequency vocabulary. Conversely, high-frequency vocabulary is found across too many registers to be considered specialized.

The Number of Words Needed For Sufficient Comprehension

In order to know the disposition and nature of vocabulary learning, and to set certain goals to achieve in designing a curriculum and a teaching methodology, research has emphasized that some questions (research strands) has to be posed. The mentioned trend in vocabulary research has sought to figure out how many words are necessary to fulfill some tasks that learners hope to attain. There are two primary questions in the mentioned fields of study. Once we answer these two questions, we can begin to estimate the number of words learners require to achieve the required level of comprehension.

1. What kind of task are the learners trying to accomplish?

2. What proportion of lexicon coverage is needed for sufficient comprehension?

Several vocabulary researchers such as Waring & Nation (1997) have classified tasks in order to determine whether they are *comprehension* or *production* based. However, there are various studies which focus on the vocabulary inquiries of *comprehension* tasks (e.g., Hirsh & Nation, 1992; Nation, 2006; Ward, 1999; Webb & Rogers, 2009). The division is primarily made based on the fact that the mode of input can differ between 'aural' (i.e., listening comprehension tasks) vs. 'written' (i.e., reading comprehension tasks). Furthermore, comprehension tasks are typically classified by genre (e.g., listening to television programs vs. watching movies; reading newspapers vs. novels).

In connection with reading tasks, it was believed that a receptive vocabulary which allows for 95% text coverage will be adequate for a suitable and acceptable level of comprehension. This figure was related to a study by Laufer carried out in 1989. It had a strong impact on practitioners' belief in a way that for many years they believed that it is the right estimation for the amount of words needed for coverage of a passage.

Laufer (1989) carried out a research using Israeli EAP students as participants. She had them underline unknown words in a passage and then answer comprehension questions which were related to the text. She then probed the relationship between the number of unknown vocabulary underlined and reading comprehension marks. She used 55% comprehension as the measure for sufficient comprehension level. She concluded that significantly those participants who scored 55% could better understand words accounting for 95% or more of the passage. Therefore, she resulted that knowing 95% vocabulary of a certain text stands for the threshold level of adequate reading comprehension.

This vocabulary threshold for a sufficient understanding of a text was later changed to 98% by Hu & Nation (2000). In that study, they considered "adequate comprehension" in a different way: reading comprehension test scores equal to those earned by test takers who had vocabulary knowledge providing 100% text coverage (a score of approximately 71% when scores on two comprehension tasks were averaged). They used 71% comprehension as the measure for sufficient comprehension level. They concluded that significantly those participants who scored 71% could better understand words accounting for 98% of the passage. Their findings suggested that 98% word coverage allowed most participants to achieve adequate comprehension, as measured by the test (i.e., a comprehension score of 71%).

Schmitt, Jiang, and Grabe (2011) ran a much better-designed investigation regarding the relationship between vocabulary levels and reading comprehension. They asked the participants to identify their knowledge of vocabulary items doing a yes/no checklist test format. The information gained was compared with their performance on a reading comprehension test. Their findings implied that the vocabulary coverage percentage needed to gain an adequate comprehension of a text is closer to the figure proposed by Hu & Nation

(2000): to achieve over 68% comprehension, participants are required to have vocabulary knowledge that covers at least 98% of the text.

A number of other studies have also attempted to distinguish how many words are needed to attain adequate comprehension. In the literature, adequate comprehension has been referred to as either a merest score on a comprehension test (e.g., Laufer, 1992), or, more often and more contemporarily, simply the threshold vocabulary text coverage recently proposed by researchers for adequate comprehension (e.g., Adolphs & Schmitt, 2003; Nation, 2006; Webb and Rogers, 2009a).

In another study, Laufer (1992) compared the vocabulary size of Israeli EAP students (as determined by a vocabulary levels test) with their reading comprehension scores to determine what level of vocabulary knowledge would allow for adequate comprehension. Participants needed a receptive vocabulary of 3000 word families to achieve this score as adequate comprehension was operationalized as a score 56% or better on a reading comprehension test. She also found that for every additional 1000-word level that participants tested at, their comprehension scores would increase by approximately 7%.

Waring and Nation (1997) proposed that 95% coverage of written texts could be attained with 3000-5000 word families, based on text coverage of the Brown Corpus (Francis & Kucera, 1982) and of a corpus of adolescent fiction (Hirsh & Nation, 1992). Interestingly, in this study, Waring and Nation additionally propose that for productive tasks (i.e., speaking or writing), "it is possible to make use of a smaller number [of words]"—2000-3000 (p. 10), though it is unclear how they arrived at this figure.

Adolphs and Schmitt (2003), six years later, researched the vocabulary level required to take part in everyday spoken conversation or any kind of discourse. They suggested a tentative figure of approximately 3000 word families (or 5000 individual word forms) as a reasonable target (i.e., approximately 96% word coverage) for participation in everyday spoken discourse (p. 433). They investigated text coverage in two spoken corpora: the 10-million word spoken component of the BNC and the five million-word Cambridge and Nottingham Corpus of Discourse in English (CANCODE). They certified that "there may be no one lexical coverage figure which would supply an adequate amount of vocabulary in every situation" (p. 433).

A contemporary trend in estimating vocabulary size needs for language learners in order to accomplish different tasks is to base estimations on the BNC sub-lists (Nation, 2006). In short, in the British National Corpus, the BNC lists consist of a series of 1000-word lists arranged in order of frequency of occurrence. For instance, the BNC 1K is the 1000 most frequent word families in the BNC, the BNC 2K is the second 1000 most frequent word families, etc. The researchers then, in order to meet the desired coverage (e.g., 98%) evaluated the coverage of a text or a corpus provided by the BNC 1K, BNC 2K, etc. The vocabulary size required is then based on the BNC frequency band at which the desired coverage is met. For example, researchers conclude that a vocabulary of 5000 words is required to comprehend a kind of text if 98% coverage of a that kind of text is met at the BNC 5K level.

Nation (2006) used this methodology and proposed that to read a novel or a newspaper a vocabulary of 8-9000 words is required, and that vocabulary of 6000-7000 words is needed to understand a children's movie (i.e., *Shrek*) or unscripted spoken interaction (i.e., interview and conversation from the BNC). Webb and Rogers were interested in the comprehension of television programs (2009a) and movies (2009b) and the number of words needed to do so. They employed Nation's methodology for assessing the vocabulary size required to achieve adequate understanding. Their estimates offer a much wider range of themes and topics and the related required vocabulary size. They suggested that vocabulary needs differ depending on genre (e.g., crime, drama, comedy) and, to some degree, variety of English (i.e. British or American English). In order to understand 95% coverage of TV programs a list of 2000-4000 words while in order to understand 98% a vocabulary of 5000-9000 words is needed. Comprehending comparable coverage in movies requires a bit higher vocabulary level: 3000-5000 words for 95% coverage, and 6000-10000 words for 98% coverage.

As the researches discussed above attempted the vocabulary challenge presented by a variety of domains, two notable studies have evaluated this demand with specific relation to academic reading comprehension. A more often-cited study by Hazenburg & Hulstijn (1996) probed the vocabulary load of readings encountered by Dutch university students. They concluded that a minimum of 10000 headwords (akin to word families) + proper nouns allowed for 95% word coverage of the Dutch texts, thus the contention that 10000 words be the minimum target for students hoping for adequate comprehension of academic texts.

20

More recently, Laufer & Ravenhorst-Kalovski (2010) revised the topic, discussing Laufer's (1989) operationalization of "adequate comprehension" from a 55-56% reading comprehension score to a norm-referenced measure. When adequate comprehension was met as the power to read independently (which was predicted by an entrance text in an Israeli national university) the text coverage needed was understood to be about 98%, achieved by BNC bands up through 6K-8K.

Before going forward, it is important here to distinguish between Nation's (2006) method, also used by Webb & Rogers (2009a; 2009b) and Laufer & Ravenhorst-Kalovski (2010), and previous calculations of assessing vocabulary size requirements. Previous studies (e.g., Adolphs & Schmitt, 2003; Hazenburg & Hulstijn, 1996) figured out the number of actual words which were needed to reach the desired coverage in a certain text, corpus or discourse.

A more different question is posed in the recent methodology used by Nation (2006) and others. It challenges the BNC 1000-word lists asking the number of the words which are required to account for the desired coverage of a text are needed to achieve adequate coverage of a certain text or corpus, and how many of the BNC 1000-word lists are required to account for the desired coverage of a text. For example, Nation (2006) found that the novel *Lady Chatterley's Lover* was written with just over 5000 word families. However, he did not conclude that it takes a bit over 5000 word families to read this novel. Instead, because words up through BNC 9K were required to provide 98% coverage of this novel, he concluded that reading a novel requires a vocabulary of 8000-9000 word families. Table 1, from Nation (2006, p. 70) illustrates how this calculation was made.

Table 1. Cumulative Percentage Coverage Figures for Lady Chatterley's Lover

Word list (1,000)	Coverage without proper nouns (%)	Coverage including proper nouns (%)
1	80.88	82.93
2	88.09	90.14
3	91.23	93.28
4	93.01	95.06
5	94.08	96.13
6	94.77	96.88
7	95.38	97.43
8	95.85	97.90
9	96.17	98.22
10	96.41	98.46
11	96.62	98.67
12	96.82	98.87
13	96.93	98.98
14	96.96	99.01
Not in the lists	97.92	1000.00

Corpus-Based Language Learning

A corpus is a large collection of language as it occurs naturally which is collected systematically and is usually stored and processed digitally. Corpora existed before computers and were created primarily for linguistic research. Since then, the volumes of corpora have become bigger and bigger (about several hundred million words) and the possibilities of using corpora have changed and gone beyond the simple linguistic research to language teaching and research.

The concept of a principled corpus may also cover, among other parameters, text length, variety of authorship, time constraints, and register. One important component of well-designed corpora is that they can show accurate tendencies within the targeted area of language and will provide, as McEnery and Wilson (2001) noted, "a broad range of authors and genres which, when taken together, may be considered to 'average out' and provide a reasonably accurate picture of the entire language population in which we are interested" (p.

30). A more notable additional issue which is taken from "relative frequency" is saliency. It refers to assessing the degree of the importance of text related items found in a particular text or groups of texts, which are often interpreted by researcher or through the use of "keyness" (Baker, 2006; Scott & Tribble, 2006). Keyness is a descriptor for words (usually referred to as keywords) within a text "whose frequency is unusually high in comparison with some norm" (O'Keefe et al., 2007, p. 12).

Corpora are identified often either large amount of texts generalized or small amount of texts which are specialized corpora. In order to make broad claims generalized corpora are often used and are related to language use. They generally include multiple sub-corpora composed of large parameters like register, spoken vs. written form, or location. The 100-million-word British National Corpus can be categorized as corpora which is normally used and is generalized. It includes an even broad proprietary corpora deployed by publishers as means of investigation tools for corpus-based dictionaries (e.g., Rundell, 2007; Sinclair, 2001); grammars (e.g., Biber et al., 1999; Carter & McCarthy, 2006; Collins Cobuild English Grammar, 1990), and textbooks (e.g., Bunting, Diniz, and Reppen, 2013; Conrad & Biber, 2009). However, specialized corpora, such as MICASE (Michigan Corpus of Academic Spoken English), pertain to a narrower focus and are usually smaller. These technical and specialized corpora are often produced for particular aims. Some of the targets then can be introduced here are such as conceptualizing EFL or ESL learners' use of idioms or phrases or analyzing the word diversity in a small scale corpus (i.e. a corpus of newspaper or magazine summaries). The names of useful lists of corpora, including either generalized or specialized corpora which are accessible, are given in Bennett (2010) and O'Keefe, McCarthy, and Carter (2007).

The technology developed over time and corpus-based language learning received an increasing amount of attention. Teachers and researchers believe that it is an influential second and foreign language teaching approach facilitating course design (Flowerdew, 1993; Tribble and Jones, 1997), teacher training (Allan, 2002; Tsui, 2004), materials development (McCarthy & O'Dell, 2005; O'Dell & McCarthy, 2008; Willis, 1998), classroom implementations (O'Keeffe, McCarthy, & Carter, 2007), learners' writing skills (Conzett, 2000; Gilmore, 2009; O'Sullivan & Chambers, 2006; Sun, 2007), writing feedback (Gaskell &

Cobb, 2004), vocabulary (Sun, 2000, 2003), grammar (Conrad, 2000; Meunier, 2002), reading (Brodine, 2001) and the transfer of learned vocabulary knowledge to writing (Kaur & Hegelheimer, 2005).

In language teaching, corpora have been utilized in writing dictionaries (such as the Collins COBUILD English Language Dictionary) and developing concordancers, or computer programs to be used with ESL and EFL teaching. Today, some corpora and concordancers are available online.

Related to corpora and concordancers, a collocation is finding two or more words within a short distance from each other in a passage. This tool is used in finding language patterns, such as grammatical templates and rules of language. The number can be adjusted for different research purposes (Baker, 2006). A good example of a collocation can be seen in English verbs and their relationships with to-infinitive and gerunds, i.e., certain verbs are preferred to certain forms of complements because the verb phrase structure might have a big or a small difference in meaning. For example, the difference can be understood when two different complements are used with the verb "talk" (e.g. stop to talk and stop talking).

Corpus-based language learning practitioners have suggested the application of using the concordancer in designing language teaching syllabuses and materials to be used in foreign language education (Cain, 2002; Ciesielska-Ciupek, 2003; Davis & Russell-Pinson, 2004; Flowerdew, 1993; Fox, 1998; Lewis, 1993, 1997, 2000; Osborne, 2003; Tribble & Johns, 1997; Wichmann, Fligelstone, McEnery, & Knowles, 1997; Willis, 1990; Willis, 1998). They have offered that the use of corpora in language education can provide not only a way to discover about a language and its culture, but also chances for using it communicatively and interactively, while focusing on meaning and not form. Thus, they suggest that using corpora in second and foreign language education can be a beneficial tool to develop reading and writing skills and understand and produce particular types of texts.

From the view of the language learner, corpus-based second and foreign language instruction can provide an important tool by which language learners can increase their motivation and make themselves more autonomous (Aston, 2001; Woolard, 2000).

Accessing the authentic language can be mentioned as one of the benefits of using a corpus for second and foreign language education, which is not possible through using a traditional dictionary. In most dictionaries limited and artificial sentences are used as examples to clarify meanings. The use of concordancing is suggested referring to the aforementioned fact. Some of the recent studies can be named here as follows: collocation learning (Howarth, 1998; Sun and Wang, 2003), lexical acquisition (Cobb, 1999), writing (Sun, 2007; Tribble, 2001), critical literary (Louw, 1997), stylistics (Kettemann, 1995), and grammar (Sun, 2000, 2003).

Although the usefulness of context in vocabulary learning is emphasized a lot, a more important issue which is discussed in language learning strategies is involved in corpus-based language learning. Inductive and deductive learning are two vital and long debated issues in vocabulary learning studies. Usually, corpus-based language learning mechanism is based on inductive learning.

Learners look up words in the concordance lines and infer the meanings and usages of the words. Moreover, corpus-based language learning offers deductive as well as inductive learning. The concordance lines can be lined up in both the descending and ascending order in which a keyword happens so that the learners can see the usage of the vocabulary related to the frequency of usage. This frequency list allows the learners to deductively learn the usage of the vocabulary. Corpus-based language learning, therefore, uses advantages of both learning strategies.

To put it in a nutshell, the learners not only benefit from the corpus-based instruction for their linguistic knowledge correction, but the instruction also promotes their lexico-grammatical awareness. Further, the more the students get imposed and familiar with corpora the more independent and confident language producers they become. Thus, the issue of practicality of corpus comes into question and becomes of major concern. Hereafter, the issue of corpus-based language learning will be discussed more in detail addressing its most important elements along with the relevant research. It will focus on more on the problems and how it does not meet the teachers' needs as the core element putting the idea of corpus-based learning into practice.

Corpus Linguistics and Teachers

Since teachers show different reactions to different innovations in a variety of ways, it is important to go beyond an ideal model of teacher research, and to move into areas that attest teacher intentions, knowledge, effect, and reasoning in addition to behaviors before, during, and/or after teaching in order to fully understand how the language instructors do what they do when they teach, and to gain information regarding cognitive and affective processes. Shavelson and Stern (1981) argued that

> a solely behavioral approach… cannot account for predictable variations in teachers' behavior arising from differences in their goals, judgments, and decisions,[but] research linking teachers' intentions to their behavior will provide a sound basis for educating teachers and implementing educational innovations. (p. 455-456)

Shi and Cumming (1995) examined through a series of weekly interviews each teacher's post-lesson thinking and beliefs about writing instruction. The results indicated that teachers' decision-making was based primarily on each teacher's unique set of existing beliefs and practices, "considered largely in personal, practical terms" (p. 104). The findings also implied that some teachers might accept innovation when it suits their beliefs, or they might choose to ignore parts of innovation or modify them to fit those existing beliefs and practices. They could decline the innovation entirely, however, if it cannot fit their existing practices and beliefs. The teachers, whether accepting, modifying, or rejecting the innovation, generally integrated it within several months into their "usual belief systems and teaching routines" (p. 104).

Corpus-Based Teaching Methodologies

The methodologies used in corpus-based teaching include the various ways in which corpus tools can be employed to provide learning outcomes, which in turn can be imagined in several ways. In an earlier approach by Biber, Conrad and Reppen (1998), three basic pedagogical uses of corpus findings were identified:

Many of the findings of corpus-based studies, as well as the techniques in corpus-based research, have valuable applications in educational settings… [in]… three ways: the dissemination of findings from corpus-based studies, the development of educational materials, and the design of classroom activities. (p. 236)

Another way to consider the corpus-based teaching methodologies is to take into account Gavioli's (2005) explanation of how corpus analysis can be an aid for foreign language instructors: its use in syllabus design, as support for teacher explanations, and direct use by learners either with teacher guidance or autonomously.

This first issue stated above primarily means that findings from corpus-based studies are used to identify grammatical, lexical, or discourse points to include, exclude, or organize according to importance in the curriculum. The second issue, which is the use in supporting teacher explanations, can be represented by the use of corpora to provide authentic materials to highlight specific grammatical, lexical or discourse points and also to provide critical examination of existing published or self-created materials. Researchers, teachers, and learners find "that there exist considerable differences between what textbooks are teaching and how native speakers actually use language as evidenced in the corpora" (McEnery & Wilson, 2001, p. 120). In a third view of corpus-based teaching methodologies, Römer (2011) made a distinction between direct and indirect applications of corpora and findings from corpus-based studies through the use of 1) the corpora themselves and any software used to analyze the data, and 2) techniques used to analyze and interpret the results of corpus investigations. One example of the indirect application is to compare corpus-based findings with choices made in textbooks, identifying "any mismatches between naturally occurring and the type of [language] that is put forward as a model in the examined teaching materials" (pp. 208-209).

Related Studies

Farr (2003) used a corpus of spoken English derived from an Irish university setting to examine how engaged listenership was signaled in meetings between tutors and graduate students. The various linguistic devices employed by both parties for this purpose were quantified and functionally analyzed. The three strategies examined in the study included

minimal response tokens (e.g. mm hm, mm, yeah), non-minimal response tokens (e.g. really, right, fine, good), and simultaneous speech and interruptions. These items were found to differ quantitatively and functionally and were shown to be highly relevant for the effective functioning of students in such a spoken EAP context.

Harwood (2005) explored the various anti-textbook arguments in the literature to determine their relevance to the field of EAP. He distinguished between a strong and a weak anti-textbook line, then reviewed the corpus-based studies which compare the language EAP textbooks teach with corpora of the language academic writers use. After sorting out the methodology of some of the studies, he claimed that the weak anti-textbook line most accurately described the state of the EAP textbook market. Although this view was not against textbooks per se, it held that the current state of commercial materials was highly unsatisfactory, and that publishers and materials writers must make greater efforts to ensure that research findings were put into use in textbooks.

Cookson, Hunter, Jackson, and Sick (2005) presented an overview of the criterion essay management system in the context of an English for Academic Purposes Writing course. They discussed student and teacher productivity, providing feedback on student writing, using corpus linguistic techniques in conjunction with Criterion, and learning outcomes. The first semester of English for Academic Purposes Writing course, which was a compulsory course for English literature and language students at Obirin University, focused on expository writing, typical of the TOEFL[R] writing exam. The second semester focused on writing about literature. To facilitate their writing all students were provided with a user account for Criterion, a web-based essay management system. Essays submitted to Criterion were evaluated by computer, and returned to the writer in approximately one minute. Submitted essays were kept on the Criterion server, where they could be accessed and evaluated by teachers, and further revised by students.

Flowerdew (2015) investigated current corpus-based research on written academic English and illustrated the tight links that exist between corpus research and pedagogic applications. She first explicated Sinclair's concept of the 'lexical approach', which sustains much corpus research and pedagogy. She then discussed studies which focus on individual lexical items,

phraseology (including lexical bundles) and rhetorical functions. Finally, she examined studies that combine Swalesian move structure analysis with corpus analysis. Throughout the paper, she underlined the importance attached to lexis in both corpus research and pedagogic applications of corpus findings in the realm of written discourse in English for Academic Purposes (EAP), covering both general and more specific purposes.

Since previous studies were mainly focused on corpus use in classroom settings, more research was needed to reveal how students react to using corpora on their own and what should be provided to help them become autonomous corpus users, considering that their ultimate goal is to be independent scholars and writers. In a study conducted by Chang (2014) in an engineering lab at a Korean university over 22 weeks, data on students' experiences and evaluations of consulting general and specialized corpora for academic writing were collected and analyzed. The findings showed that, while both corpora served the participants well as reference sources, the specialized corpus was particularly valued for its direct help in academic writing because, as non-native English-speaking graduate engineering students, the participants wanted to follow the writing conventions of their discourse community. The participants also showed disparate attitudes toward the time taken for corpus consultation due to differences in factors such as academic experience, search purposes, and writing tasks. The article concluded with several suggestions for better corpus use with EAP students regarding the compilation of a corpus, corpus training, corpus competence, and academic writing.

Tribble and Wingate (2013) reported on the first phase of a discipline-specific academic writing project at King's College London. They believe that the instructional approach takes as its starting point the assessed academic genres which students have to write. Working within the tradition of genre-based approaches to academic writing instruction, they used the KCL Apprentice Writing Corpus to provide a databank of exemplars. These texts, along with grading and feedback information from academic departments, and input from disciplinary specialists, have provided the basis for discipline-specific writing courses which focus on strategically important assessed disciplinary genres (Critical Assignments in Applied Linguistics, and Laboratory Reports in Pharmacy). They discussed issues in corpus compilation, the selection of exemplars, and the rationale which has informed the development of the teaching resources. They also provided examples from the workshops offered as part of

the courses. They concluded with a brief account of the evaluation of the courses with a discussion of how corpus informed materials were being developed to extend the project in its second phase.

Vincent (2013) demonstrated an approach based on findings from phraseology which can be used to identify potentially useful phrases in a text by starting with continuous or discontinuous sequences of very frequent words. These combinations were then searched in a corpus of academic texts to find their common collocates and ascertain whether the particular wording chosen in the text was an exemplar of a recurrent phrase. The phrases identified were also evaluated in terms of whether they were worthy of pedagogic attention bearing in mind their distinctiveness to academic prose and their functions. The approach thus offers practitioners of English for Academic Purposes (EAP) a way of identifying and evaluating common phrases in their own texts. Pedagogic materials were also presented which were based on the findings of such searches and which aim to raise learner awareness of phraseology in academic English.

Charles (2012) reported on the feasibility and value of an approach to teaching EAP writing in which students constructed and examined their own individual, discipline-specific corpora. The approach was trialed in multidisciplinary classes of advanced-level students (mostly graduates). The course consisted of six weekly 2-hour sessions. Data were collected from initial and final questionnaires, which provided background information and asked students to evaluate the corpus work. Data from 50 participants were presented and showed generally positive results. Over 90% of students found it easy to build their own corpora and most succeeded in constructing a corpus of 10-15 research articles. Most students were enthusiastic about working with their own corpora: about 90% agreed that their corpus helped them improve their writing and intended to use it in the future. This suggests that even corpora of this size and type can provide a useful resource for writing discipline-specific texts. They discussed the data on participants' attitudes and experiences and considered the issues and problems that arose in connection with do-it-yourself corpus-building. She also argued that this approach need not be restricted to small groups of well-resourced students, but can be implemented in mainstream EAP classes.

The study by Krishnamurthy and Kosem (2007) summarizes the processes and factors to be considered in the design and compilation of an EAP corpus (e.g., the selection and acquisition of texts, metadata, data annotation, software tools and outputs, web interface, and screen displays), especially one intended to be used for teaching. Such a corpus would also facilitate EAP research in terms of longitudinal studies, student progression and development, and course and materials design. Their study was informed by the preparatory work on the EAP sub-corpus of the ACORN corpus project at Aston University.

Hyland and Tse (2007) considered the notion of academic vocabulary: the assumption that students of EAP should study a core of high frequency words because they are common in an English academic register. They examined the value of the term by using Coxhead's (2000) Academic Word List (AWL) to explore the distribution of its 570 word families in a corpus of 3.3 million words from a range of academic disciplines and genres. The findings suggested that although the AWL covers 10.6% of the corpus, individual lexical items on the list often occur and behave in different ways across disciplines in terms of range, frequency, collocation, and meaning. The results suggested that the AWL might not be as general as it was intended to be and, more importantly, questions the widely held assumption that students need a single core vocabulary for academic study. They also argued that the different practices and discourses of disciplinary communities undermine the usefulness of such lists and recommend that teachers help students develop a more restricted, discipline-based lexical repertoire.

Richard (2003) discussed that most previous work in EAP has focused more on the content of teaching than on the methodology. By examining reports of EAP teaching practice, he identifies six key approaches to the teaching of EAP: inductive learning, process syllabuses, learner autonomy, authenticity, technology, and team teaching. Reasons for the emphasis on these approaches were given. The paper concluded by arguing that a greater emphasis needs to be placed on methodology in EAP.

Taking a corpus-based perspective to the analysis, Eniko (2006) compared the language of university classroom talk to academic prose and face-to-face conversation, positioning university classroom talk on the language continuum of speech and writing. More specifically,

looking at a large number of linguistic features working together, she described the language of 196 university class sessions (1.4 million words) collected at six universities across the United States. The analysis was based on Biber's multi-dimensional analytical framework (Biber, 1988). Overall, the results indicated that in these classrooms language features associated with both informational focus (as in academic prose) and involved discourse (as in face-to-face conversation) were equally present. Hence, this evidence-based research supported the argument that North American university classrooms exhibit language that can be treated as an interface on an oral–literate continuum.

Read (2002) focused on the form of the input for EAP listening tests. He assumed that while there was a great deal of interest currently in the use of visual media for listening assessment, it was likely that tests with purely auditory input will continue to have a significant role. The article reported on the development of a test in two audiotaped versions: a scripted monologue and an unscripted discussion of the same topic by three speakers. The test was applied to two matched groups of learners taking an intensive pre-sessional EAP course and it was found out that the monologue version was significantly less difficult than the discussion. Various possible reasons for the difference in findings were demonstrated and it was concluded with a consideration of what can be learned from the research for the design of listening test tasks with interactive input.

Grant (2011) investigated the frequency and functions of 'just' in British academic spoken English. It adopted the meanings of 'just' established by Lindemann & Mauranen, 2001, taken from the occurrences of 'just' across five speech events in the Michigan Corpus of Academic Spoken English (MICASE) to see if they also apply to occurrences of 'just' in the British Academic Spoken English (BASE) corpus. One fifth of the mostly monologic lecture data and one fifth of the more interactive seminar data was examined across the four disciplines of the BASE corpus. Results from corpus examples showed that as in MICASE, 'just' occurs most in academic spoken English as a minimizer or mitigator, across all four disciplinary groupings, often in metadiscursive or 'teacher talk' frames. Although the highest frequency of 'just' was in Physical Sciences, there was little difference in the function of 'just' among the different disciplinary groupings. There was, however, some difference between student and lecturer use

of 'just', showing that there was a place for the teaching of 'just' in English for academic purposes classes for students preparing for university study.

Focusing on the use of two functional categories (articles and prepositions) and three content categories (nouns, adjectives and verbs), Genç (2013) in his study compared a small, spoken non-native speaker corpus (10973 words) with a smaller, spoken native speaker corpus (2331 words). Participants in both groups were assigned a speaking task, the result of which produced the non-native speaker and native speaker corpora and those corpora were analyzed quantitatively with the help of computer software titled Concordance. The results indicated that due to the limited number of choices, there was a large overlap in the use of articles and prepositions; close similarity between the use of nouns and verbs; and a remarkable difference in the use of adjectives.

Wood & Appel (2014) assert that some studies have examined which units might be of greatest utility for teaching English for academic purposes (EAP), dealing with corpora of academic language and it is a known fact that most EAP programs focuses on preparing students for first year university courses, in which textbooks play the key role. They used a corpus analysis method focusing on units called multiword constructions (MWC), and addressed three hitherto unexplored key issues for EAP teachers and materials developers: What MWC existed in first year textbooks used in business and engineering, the most popular EAP student majors at a large Canadian university; whether the first year textbook MWCs were present in the reading texts in popular EAP textbooks; whether the EAP textbooks contained pedagogical treatment of MWC. The results showed that the MWC were weakly present in the EAP materials, and that they were not dealt with pedagogically.

Reviewing the literature, it is important to note that the advancement of technology has enabled the researchers to easily compile large-scale corpora (both generalized and specified). It is done as an authentic way of collections of electronically-stored language texts and information which are analyzed quantitatively. A concordancer or a computer technology can be used for linguistic analysis, in order to distinguish the frequency of particular words and expressions. As Svartvik (1990) mentions, computer/corpus-based applied linguistics yields "more academic cross-fertilization and fresh approaches to old problems" and leads to "a

33

better understanding of the complexities of natural language and the marvel of human language processing." (p. 85)

Briefly reviewing the related studies carried out in the field suggest the fact that more researchers in the fields of TEFL and SLA have recognized the potential pedagogical and practical importance of learner corpora. However, less studies have been carried out to fill the gap of learning English in Turkey as an EFL context.

Corpus Studies In Turkish Context

Middle East Technical University (METU) has proposed a Turkish Corpus which is a collection of two million words of post-1990 written Turkish samples. It is said on the website that "the words of METU Turkish Corpus were taken from 10 different genres. At most 2 samples from one source is used; each sample is 2000 words or the sample ends when the next sentence ends." (METU, n.d.)

TS Corpus (2012) is "a general-purpose corpus containing 491 million tagged tokens". The data of TS Corpus is web-based and it is derived from the internet such as newspapers, instructions on the university pages, meeting reports, The Grand National Assembly of Turkey reports. Unlike other corpora in Turkish, this corpus serves a Turkish Corpus which

- presents a tagged corpus
- provides open access on internet and user-friendly sign up via internet
- presents a platform free (Mac OS, Linux, MS Windows, etc) corpus
- is compatible with all modern browsers
- is accessible by desktop PCs, net-books, tablet computers or smart phones (TS Corpus, 2012)

In another research carried out at Mersin University, a corpus called Turkish National Corpus (TNC) has been made up. It is "designed to be a balanced, large scale (50 million words) and general-purpose corpus for contemporary Turkish" and Aksan et al (2012) claim that TNC generally followed the framework of British National Corpus. For that reason, this corpus is used in the current study as the basis for creating the General Turkish Wordlist and a list of words (e.g. pronouns) that will be excluded from the Academic Turkish Wordlist. It is a 50-

million-word corpus consisting of written texts (98%) across a wide variety of genres covering a period of 20 years (1990-2009). 2% of TNC consists of transcribed spoken data. It is suggested that TNC is a balanced and a representative corpus of contemporary Turkish in which the distribution of number of words is determined proportionally for each text domain, time and medium of text (Aksan et al, 2012).

Using suggestions from "Dictionary of Word Frequency of Written Turkish" (Göz, 2003) for words to be taught to the students in basic level, Açık (2013) developed a vocabulary list for basic level Turkish vocabulary instruction. Although her list could have been used as a General Turkish List that is aimed in this study, it was not chosen because the source of her list, Dictionary of Word Frequency of Written Turkish, was thought to be inappropriate for the current study.

There are also several dissertations on corpus. In her study, Kurtoğlu (2013) analyzed values to be taught in Turkish textbooks using corpus-based discourse analysis method. Her main aim was to show how corpus-based discourse analysis can be used to decide whether the Turkish textbooks serve to achieve the values in the objectives and principles of Ministry of Education.

Yıldız (2012) investigated the predicates in the introduction and conclusion sections of Turkish academic texts that are written in humanities, basic and applied sciences; the frequencies and the clusters of them as verbs, nouns, adjectives and adverbs; the similarities and differences among their usages in the disciplines; and the effects of them in constituting academic genre.

Şanal's research (2007) addressed "multidimensional lexical aspects of a machine-readable corpus of the writing of Turkish students of English as a foreign language." His doctoral dissertation aimed at discovering: "1) learners' lexical complexity and richness, 2) how far the learner corpus is deviant from the reference corpus in terms of the features and percentages of the top most 200 frequent tokens, 3) what are the most salient and stereotype features of the learner corpus? And how far the learner corpus influenced by the learners' L1." Findings in his study showed that: "1) the learner corpus is much less complex in terms of lexical diversity and density than the reference corpus. 2-3) Learners' top 200 tokens are markedly

characterized by vague lexica, underuse and overuse of some lexica, resulting from the influence of the linguistic and rhetorical features of learners' L1."

Ağçam's study (2014) "intended to cover the epistemic stance devices used in conveying author stance in Academic English". It was investigated that whether there is a statistically significant difference across the doctoral dissertations authored in English by native, Spanish-speaking and Turkish-speaking candidates with respect to the use of these devices. A total number of 136 doctoral dissertations written by native and non-native speakers of English were analyzed. Epistemic stance devices found in the dissertations were identified and examined in four groups (epistemic adjectives, epistemic adverbs, epistemic nouns and epistemic verbs).

Thus, this research intends to answer how the TFL learners receive input that has sufficient quality and quantity. It will also find an answer suggestive of some of the weak points and strong points of teaching Turkish as a foreign language situation in Turkey. The research will be following the way to discover answers that will help the teachers fill this gap by providing the right corpora and general/academic wordlists. Another question which remains unanswered is whether the TTCs provide the foreign students with adequate opportunity to employ whatever they learned in independent reading. It should also be analyzed to see if the texts which are used in the textbooks are representative of the sample texts available in the field. Additionally, it is necessary to know whether the number of the words needed for a successful academic life brought in the textbooks correspond well with what their common uses are.

CHAPTER 3

METHODOLOGY

In second language learning, vocabulary is the essential part of all comprehension both in authentic tasks such as checking out in a hotel and in academic reading, which is in turn fundamental for the academic success of language learners. Instructors in content area classes frequently teach content specific vocabulary, but they need to be aware of the obstacles that can occur when students do not know the basic words. Therefore, this study focuses on both basic words and academic words.

Preliminary Considerations

Wordlists such as General Service List (West, 1953), The Academic Word List (Coxhead, 2000), and Academic Vocabulary List (Gardner & Davies, 2013) were created to assist students and teachers regarding English language teaching. Similar to these wordlists, using suggestions from "Dictionary of Word Frequency of Written Turkish" (Göz, 2003) for words to be taught to the students in basic level, Açık (2013) developed a vocabulary list for basic level Turkish vocabulary instruction. Although her list could have been used as a General Turkish List that is aimed in this study, it was not chosen because the source of her list, Dictionary of Word Frequency of Written Turkish, was thought to be inappropriate for the current study and the main reason for that was that Açık's list was lack of the spoken data since Dictionary of Word Frequency of Written Turkish did not contain any transcribed spoken data. Therefore, the first 11000 most frequent words from Turkish National Corpus (Aksan et al, 2013) were used as the basis for generating the General Turkish Wordlist since

this corpus is "a balanced and a representative corpus of contemporary Turkish" with a size of nearly 50 million words and "consists of samples of textual data across a wide variety of genres covering a period of 20 years (1990-2009)".

Research Design

This research was designed as a mixed-method research study employing quantitative and qualitative approaches. However, it was administered with a more focus on quantitative aspect. An Academic Vocabulary Test was created and implemented in order to describe TFL learners' knowledge of Turkish academic vocabulary. So, a quantitative method was used and the results were statistically analyzed.
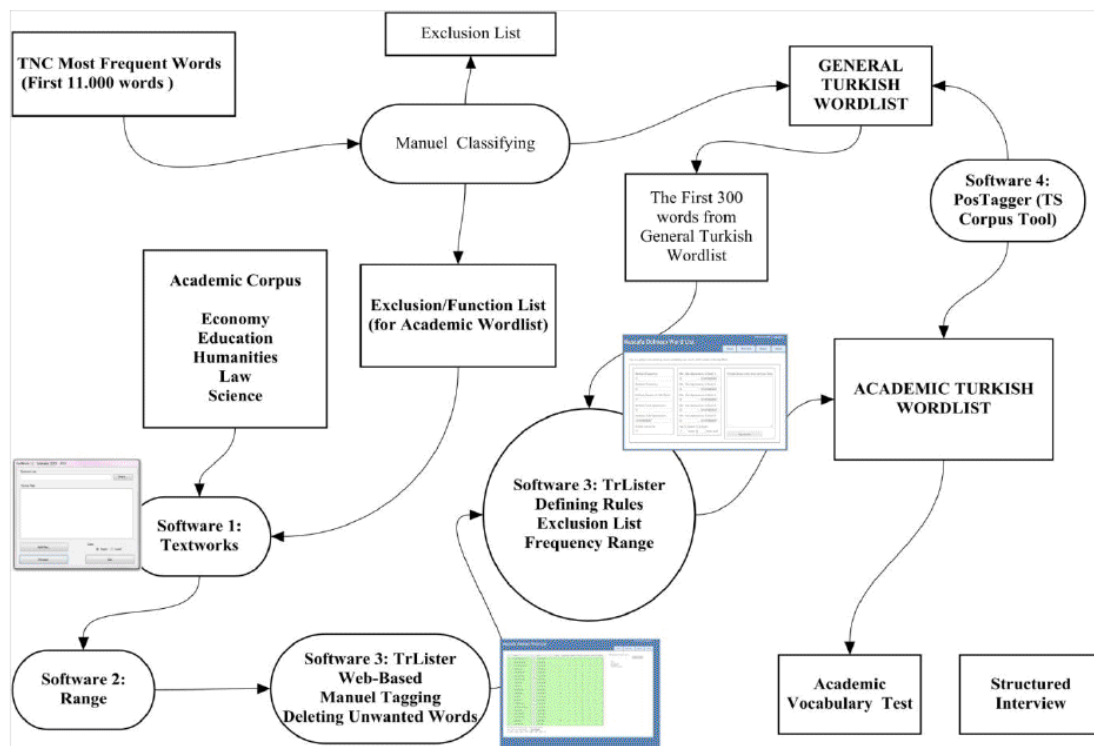
Figure 1. Research design of the current study

The qualitative aspect of the study, on the other hand, involves qualitative examination of the interview forms to expand and elaborate the quantitative results. The design of the study is illustrated above in Figure 1. The steps of creating general and academic wordlists will be explained in the subsequent sections.

Developing The General Turkish Wordlist

Using a frequency list derived from a comprehensive corpus is the most logical way to create a general language wordlist. Therefore, as mentioned in the previous sections, prior corpus studies in Turkish were reviewed beforehand and the first 11000 most frequent words from Turkish National Corpus were used to create the General Turkish Wordlist. In English, grouping the inflections of the words under lemmas is a relatively easy process because inflectional suffixes are not confusing and limited in number. On the other hand, Turkish is an agglutinative language so a frequency list does not give us the most common lemmas easily. The most frequent 11000 words from Turkish National Corpus were selected to make a list of lemmas as the unit of analysis. In other words, the most frequent 11000 words of Turkish National Corpus were manually grouped together under lemmas differing only in inflectional forms and belonging the same major class.

Lemmas, or the groups of lexical forms with the same stem that belong to the same word class, (Francis & Kucera, 1982) are used for the creation of the General Turkish Wordlist and Academic Turkish Wordlist. They are understood by learners as follows: words with the same base and inflections are members of the same word family as defined by Level 2 of Bauer and Nation's (1993) scale. An example of the verb k*ullanmak* (to use) can be seen in Table 2.

As it can be seen in the table below, the verb *kullanmak* (to use) has 27 different inflections whereas the same verb form in English has only 4 different inflectional forms (use, used, using, uses).

Table 2. Inflectional Forms of the Verb "kullanmak"

	Word	Frequency	Lemma	Total Frequency
1	kullanan	2210	kullanmak	45524
2	kullanarak	3499	kullanmak	45524
3	kullandı	1127	kullanmak	45524
4	kullandığı	2607	kullanmak	45524
5	kullandığımız	563	kullanmak	45524
6	kullandıkları	891	kullanmak	45524
7	kullanılabilir	1449	kullanmak	45524
8	kullanılacak	1074	kullanmak	45524
9	kullanılan	8466	kullanmak	45524
10	kullanılarak	1586	kullanmak	45524
11	kullanıldığı	1390	kullanmak	45524
12	kullanılır	2838	kullanmak	45524
13	kullanılıyor	833	kullanmak	45524
14	kullanılmaktadır	1831	kullanmak	45524
15	kullanılması	2819	kullanmak	45524
16	kullanılmasına	525	kullanmak	45524
17	kullanılmasını	540	kullanmak	45524
18	kullanılmaya	639	kullanmak	45524
19	kullanılmış	811	kullanmak	45524
20	kullanılmıştır	1498	kullanmak	45524
21	kullanır	765	kullanmak	45524
22	kullanıyor	1019	kullanmak	45524
23	kullanma	1973	kullanmak	45524
24	kullanmak	2288	kullanmak	45524
25	kullanması	527	kullanmak	45524
26	kullanmaya	1077	kullanmak	45524
27	kullanmayı	679	kullanmak	45524

While finding out lemmas, some words were omitted from the list. The words excluded from the General Turkish Wordlist were:

1) All proper names (*Rıza, Afrika, Akdeniz* etc.)

2) All numbers including ordinal numbers and years

3) Abbreviations (*ABD, ytl, YÖK* etc.)

4) Letters and misspelled words

5) Question Particles (*mı, mıdır, mısın* etc.) and exclamations

A full list of these words can be seen in Appendix 1.

Building The Academic Turkish Corpus

Researchers who have compiled academic corpora included a variety of academic texts. As it can be seen in Table 3 below, Coxhead's corpus includes "414 academic texts by more than 400 authors, containing 3513330 tokens (running words) and 70377 types (individual words)" in four subcorpora: arts, commerce, law, and science. Her academic corpus contains articles from academic journals, university textbooks or coursebooks and texts from some other corpora.

Table 3. Composition of Coxhead's (2000) Academic Corpus

	Discipline				
	Arts	**Commerce**	**Law**	**Science**	**Total**
Running Words	883214	879547	874723	875846	3513330
Texts	122	107	72	113	414
Subject Areas	Education	Accounting	Constitutional	Biology	
	History	Economics	Criminal	Chemistry	
	Linguistics	Finance	Family and Medicolegal	Computer Science	
	Philosophy	Industrial Relations	International	Geography	
	Politics	Management	Pure Commercial	Geology	
	Psychology	Marketing	Quasi-commercial	Mathematics	
	Sociology	Public policy	Rights and remedies	Physics	

Another study entitled "A new academic vocabulary list", carried out by Gardner & Davies (2013), used the academic portion of the Corpus of Contemporary American English (COCA) (http://corpus.byu.edu/coca). The academic section of COCA contained the following nine disciplines with 120 million words: 1) Education 2) Humanities 3) History 4) Social science 5) Philosophy, religion, psychology 6) Law and political science 7) Science and technology 8) Medicine and health 9) Business and finance.

In the light of the studies mentioned above, the textbooks currently used at the undergraduate programs of Open Education Faculty of Anadolu University were used in order to build a sample *Academic Turkish Corpus* since undergraduate foreign students learning Turkish as a foreign language are the target population in this study. They need to learn Turkish in order to follow their courses at Turkish Universities offering Turkish-medium instruction. Therefore, listening especially understanding the lecturers, reading comprehension, and writing are the

most important skills in their academic lives and their textbooks are their primary instructional material.

Corpus Design and Compilation

Similar to the studies mentioned previously (Coxhead / Gardner & Davies), textbooks which had open-access in the 2013-2014 academic year and could be downloaded from the website of Open Education Faculty of Anadolu University were chosen across five different categories. These categories were: 1) Humanities 2) Education 3) Economy (Commerce) 4) Law 5) Science. Textbooks were selected in different numbers to ensure a fair distribution across the number of words in each category. While 30 textbooks were added to the categories of Education, Economy and Law (10 for each), 11 textbooks were added to Humanities and 16 to Science. The details of the Academic Turkish Corpus can be seen in Appendix 2 and Appendix 3.

Developing The Academic Turkish Wordlist

Developing the Academic Turkish Wordlist was a really painful process. To make it better understood, the whole process is elucidated as below:

1) 57 textbooks containing 4032442 words (tokens) were put under five different categories. The parts not required such as the front and back cover, table of content and references were extracted one by one using Foxit PDF Editor V. 2.0 computer software package.

2) The textbooks in the same category were merged and one pdf file for each category was obtained. Table 4 shows the number of pages of these pdf files in each category. Each pdf file was processed using AbbyyFineReader 10.0.software. This software was used to convert PDF files in each category into txt file format.

Table 4. The Number of Pages in Each Category

Category	The number of textbooks	The number of pages
Economy	10	2121
Education	10	2124
Humanities	11	2446
Law	10	2028
Science	16	3020
Total	**57**	**11739**

3) At this point, txt files had many unwanted letters, numbers or misspelled words. Therefore, another software was needed to extract these unwanted words and to organize remaining words as word per line format. Upon request, a computer programmer wrote a small scale program (software) called Textworks, which has three main functions:

a) Deleting some of the unwanted words automatically

b) Excluding the list of words prepared beforehand.

c) Organizing words as word per line format

To delete some of the words automatically, Textworks uses some principles. The first principle is to delete all the numbers in the input file. The second is to delete the words which have letters that do not exist in Turkish alphabet (The letters of *q*, *w*, and *x*). Textworks can leave out the list of words from the input file. While generating the General Turkish Wordlist, two exclusion lists were formed. The first exclusion list was the list of the words which were excluded from the first 11000 words of TNC which was previously mentioned in the section of "Developing the General Turkish Wordlist" (Appendix 1). The second exclusion list was formed at same time to exclude function words from the Academic Turkish Corpus. This second exclusion list therefore was called the Function list. Function list or the second exclusion list contained pronouns, determiners, exclamations, some certain conjunctions,

adjectives and the nouns which were considered inappropriate to be included in the Academic Corpus. Pronouns, determiners, exclamations and some certain conjunctions in the Function list were thought to be directly associated with grammar and the Academic Wordlist was intended to include content words. The function list is given in Appendix 4. The last and the most important function of Textworks is to organize words as word per line format. We needed input files (txt files in five categories of the Academic Corpus) as word per line format for two reasons. The first reason is that the software, RANGE (Heatley, Nation, and Coxhead, 2002), which is used in the next step of the study, needs input files as word per line format. The second reason is that when the words are in word per line format, it is easier to leave out the misspelled or unwanted words. For that reason, Textworks was written to sort the words firstly in alphabetical order and then according to the number of the letters in each word. In this way, the longest words, probably the misspelled ones, appeared at the end of the output file and were deleted manually afterwards. Presented in Figure 2 is the interface of Textworks.
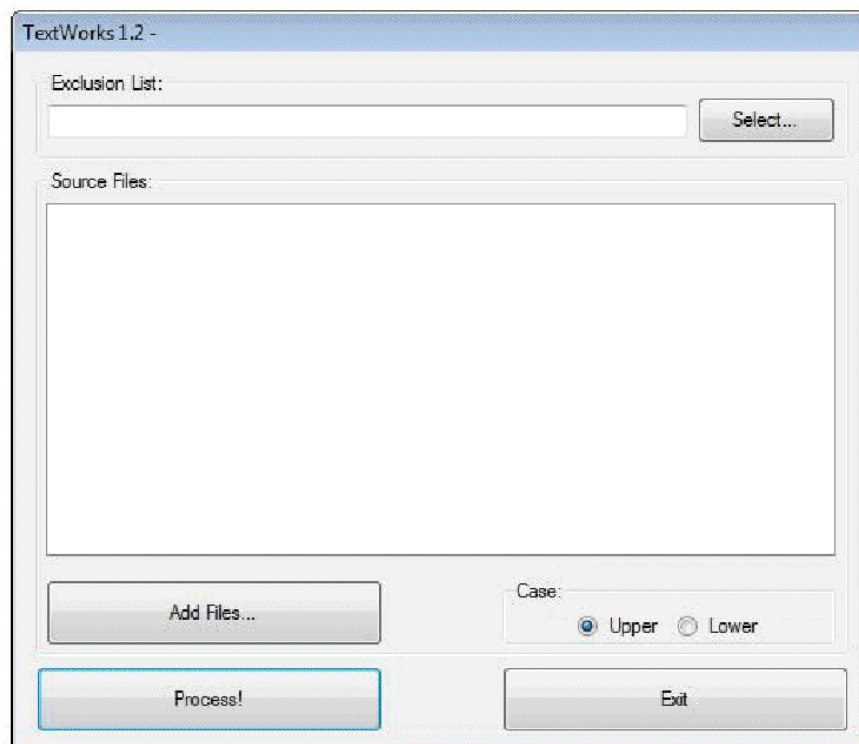
Figure 2. The interface of Textworks

45

4) After having mostly organized five categorical txt files that are free from numbers, most misspelled words, most abbreviations, most function words and most proper nouns, another software RANGE (Heatley et al, 2002) was used to generate a frequency list of these txt files in one txt file. All the words in 5 categorical txt files were grouped and sorted out in an alphabetical order using RANGE. Sample output file from RANGE can be seen in Table 5. In the table, F1, F2, F3, F4, and F5 show the categories in the corpus, "Range" shows how many categories the word exists in and "Freq" shows how many times the word occurs in all categories.

Table 5. Sample Output File From RANGE

TYPE	RANGE	FREQ	F1	F2	F3	F4	F5
ABAJUR	1	2	2	0	0	0	0
ABAKÜS	2	2	0	0	1	1	0
ABAKÜSLERİN	1	1	1	0	0	0	0
ABAKÜSÜ	1	1	0	0	0	1	0
ABANOZ	1	3	0	0	3	0	0
ABANT	1	1	1	0	0	0	0
ABARTABİLİR	1	1	0	0	1	0	0

The RANGE output file has 172156 types of words. It means that 172156 different words having different frequency rates exist in the Academic Corpus.

5) Using RANGE output file to make a wordlist in English would be very easy to make a frequency-based one. As mentioned earlier, grouping the words under lemmas is relatively comfortable in English. Table 6 shows an example of the inflections of a word in English and Turkish:

46

Table 6. Inflections of The Word "Vehicle" (Araç) in The Academic Turkish Corpus

	Inflections	Total occurrence
English	vehicles	**1**
Turkish	araçla, araçlar, araçlara, araçlarda, araçlardaki, araçlardan, araçlardandır, araçlardır, araçları, araçlarıdır, araçlarıdırlar, araçlarımız, araçlarımızın, araçların, araçlarına, araçlarında, araçlarındaki, araçlarından, araçlarındandır, araçlarını, araçlarının, araçlarıyla, araçlarla, araçsa, araçta, araçtaki, araçtan, araçtı, araçtır, araçtırlar,	**30**

The English word *vehicles* is a noun that is inflected for number, specifically to express the plural by adding the suffix -s; these two morphemes (*vehicle* and *–s*) together form the inflected word *vehicles*. Since Turkish is an agglutinative language in which extensive use of suffixes is common, it is not surprising to encounter 30 inflections of the word "vehicle" (*araç*) in the Academic Turkish Corpus. Under these circumstances, another computer program was needed to help us connect the words to lemmas. Two computer programmers were asked to develop a web-based program to classify (connecting inflections of a word to a lemma) 172156 tokens (RANGE output file) and an internet-based software called TrLister, which runs on a server, was written for the current study.

Developing TrLister took weeks and the cooperation with the computer programmers was needed in every step while designing the program. Finally a program with the following features came out:

a) Content-specific, misspelled, and unwanted words are listed out by clicking on a button next to the words.

b) Lemmas are added manually and words are listed under the lemmas by checking the box next to the words and then clicking on the *Add to Root* button.

c) Deleted (listed out) words are colored red and words listed under lemmas are colored green. That makes everything easier to understand at a first glance.

d) TrLister keeps log and calculates the number of lemmas, the number of words connected to lemmas and the number of excluded (listed out) words.

e) After all the words are manually connected to lemmas, TrLister reports all details according to the preferences of the user and gives the results in xls format. The user adjusts minimum and maximum frequency rates of total occurrences of the words, and the number of categories each lemma must exists in. (see Figure 3 below)

Figure 3. Report page of TrLister

f) TrLister can exclude a list of words from the report page. (See Figure 3) (See more for the Technical Specifications of TrLister in Appendix 15)

6) As soon as TrLister was ready to work on a server, the RANGE output file which had 172156 types to be connected to lemmas manually was installed on the program and 3 undergraduate students studying at the Turkish Education Department of Konya Necmettin Erbakan University helped the researcher to associate the 172156 types of words with the lemmas which were typed in manually. There were 8608 pages to be processed on TrLister (see Figure 4 for a sample page), so it took weeks to connect all the words to lemmas.

Figure 4. A sample page from TrLister

7) After all lemmas were added to the program, it was time to define the rules to pick out the Academic Turkish Wordlist. Frequency (ratio) and range (the number of categories each lemma has to occur) were taken into consideration similar to the other studies in literature [(Coxhead, (2000), Gardner & Davies, (2013), Brezina & Gablasova, (2013)]. On the other hand, there is no concrete way to set up rules to make a wordlist from corpora in the literature. Gardner & Davies (2013) describe the procedure of finding a rule of frequency for their academic wordlist as follows:

> There is nothing particularly special about the 1.5 Ratio, as there is no commonly accepted value for this measure. We performed extensive experimentation with values as high as 2.0 and as low as 1.2, and simply observed which words entered into and left the academic core list as we adjusted the values. At too high of a figure (e.g. 2.0) we would lose words like *system, political, create, require,* and *rate*, all of which we would argue are representative words for an academic core. On the other hand, if we set the Ratio too low, then we would pull in too many general high-frequency words. (p. 11)

Therefore, the report page of TrLister was used to define the rules and the frequency rate was initially adjusted to 100. It meant that a lemma had to occur at least 100 times in each category. And then other frequencies of occurrence (50, 40, 30, and 20 times) were tested. At the same time, the range (the number of categories each lemma had to occur) was changed to 2, 3, and 4. Range was implemented to make the list free from academic technical words or "those that appear in a narrow range of academic disciplines" (Gardner & Davies, 2013).

The third exclusion list was prepared and the first 300 words from the General Turkish Wordlist were excluded from the Academic Turkish Wordlist. The first 300 words were chosen because they thought to be too general to be included in the Academic Turkish Wordlist.

The main purpose was to have an academic wordlist containing nearly 1000 words. In the end, the frequency rate was defined as 22 and the range was 4, in other words, the words (lemmas) in the Academic Turkish Wordlist occurred at least 22 times in each category and they occurred at least in 4 categories and the first 300 words from the General Turkish Wordlist excluded from the list. Then, some of the words which were thought to be inappropriate in the

academic list were cleared out and 1010 words (lemmas) were included in the Academic Turkish Wordlist.

8) In order to present the classes of the words (part of speech) in the General and Academic Turkish Wordlists, PosTagger, one of the TS Corpus Tools (TS Corpus, 2012), was used. Unfortunately, PosTagger can parse the words with a success of 80-82 % (Sezer & Sezer, 2013), so outputs of PosTagger were checked manually.

Generating The Academic Turkish Vocabulary Test

To answer the third research question which was "To what extent do students studying Turkish at Turkish Teaching Centres of universities know the words for academic Turkish?", a quantitative method was carried out and a vocabulary test was prepared to find out to what extent TFL students know the words for academic Turkish. Nation & Beglar (2007) proposed "*a vocabulary size test*" in which ten items from each 1000 word level existed. Their test consisted of 140 items for the most frequent 14000 word families in English. They chose multiple-choice format based on the following reasons:

1 to allow the test to be used with learners from a variety of language backgrounds.

2 to control the level of difficulty of the answers. That is each item attempts to tap roughly the same degree of knowledge of a word.

3 to make marking as efficient and reliable as possible.

4 to make learners demonstrate knowledge of each item. (p. 11)

Therefore, a multiple choice vocabulary test including 30 items was prepared. The *demographics* section provided data about the age, gender, mother tongue, foreign languages and Turkish background of the participants. 1st, 2nd and 3rd items were from the first 100 word level and 4th, 5th and 6th items were from the second 100 word level and so on. All the sentences and distractors were chosen from the Academic Turkish Corpus derived in this study by the researcher. The distractors were in the same format, so any grammatical knowledge did not help the test takers get the right answer. Before the test was implemented, the draft version of the test was distributed to 20 foreign language instructors to check the items in order to find out whether there were any errors and inappropriate items. Corrections

51

were made and the final version was ready to carry out. Cronbach alpha value was found to be 0.853 in the reliability test performed for the Academic Vocabulary Test. This result indicated that the test was highly reliable. The Academic Turkish Vocabulary Test can be seen in Appendix 5.

Interview With Students

The fourth research question was "Do foreign undergraduate students who previously studied Turkish at Turkish Teaching Centres of some universities need academic Turkish vocabulary instruction to improve their knowledge of Turkish vocabulary so that they can follow academic courses?" As mentioned earlier, the qualitative aspect of the study, which addressed the third research question above, involved qualitative scrutiny of the interview forms to expand and elaborate the quantitative results. A semi-structured interview form was prepared in accordance with the views of the experts in the field. At the top of the paper a few questions were placed asking the participants about their age, gender, nationality, Turkish learning experiences and the length of the time they have been studying or dealing with Turkish. Based on the purpose of the study, the interview form had seven open-ended questions. Open-ended questions aimed to discover opinions of foreign students studying at universities offering Turkish medium instruction on their knowledge of academic Turkish vocabulary. Most of the interviewees filled in the form in person, and interview forms were distributed to few others to fill in at home and they were collected later. Dörnyei (2007) asserts that semi-structured or unstructured interviews need to be recorded. Moreover, he also claims that many people do not like to be recorded. The interview form was prepared both in Turkish and in English and the participants were asked to feel free to choose either of the languages. (See more for English and Turkish versions of the interview form in Appendix 6 and Appendix 7)

The data gathered from the interviews were saved on the computer after being checked by the interviewees and their consent was received. Then, the data were analyzed using descriptive analysis and summarized and interpreted through the themes determined in advance. The aim of the descriptive analysis is to present the findings to the readers in an organized and interpreted way (Yıldırım & Şimşek, 2006). During the data analysis, the data was analyzed

together with a different researcher in order to ensure the reliability of the research and the formula (P (agreement percentage) = (Na (Agreement/Na (Agreement) + Nd (Disagreement)) x 100)) which was developed by Miles and Huberman (1994) was used. According to this formula, the reliability between the researchers was found to be 88% and the qualitative part of the research was accepted as reliable. In addition, the findings were supported by giving examples from the answers of the students by coding them such as P1 (participant 1), P2 (participant 2).

Participants

There are two kinds of participants in the study. The first group of participants consisted of the students learning Turkish at the TTCs of universities with the goal of studying at a Turkish-medium university. The Academic Turkish Vocabulary Test was implemented to this group. They were all at the C1 level and they were asked to answer the Academic Vocabulary Test at the end of their Turkish course. As mentioned in Chapter 1, three different coursebooks were in use at the Turkish teaching centers where the Academic Vocabulary Test was implemented. Table 7 shows the number of participants from Turkish Teaching Centers of various universities in Turkey and the coursebooks used:

Table 7. The Number of Participants and Coursebooks

Universities	Number of participants	Coursebook
Gazi University	70	Gazi Üniversitesi TÖMER
Bolu A.İ.B University	23	Yabancılar İçin Türkçe Seti
Selçuk University	39	İstanbul Üniversitesi Dil Merkezi
Konya N.E. University	27	İstanbul Seti
Karatay University	27	Ankara Üniversitesi TÖMER Yeni
Niğde University	10	Hitit Yabancılar İçin Türkçe Öğretim Seti
TOTAL	**196**	

There were also 12 participants from Düzce University but they were not included in the study since all of them were Syrian and had a Turkish background.

The age of the first group ranged from 15 to 39 years, with an average of 22 years. They were from 52 different countries (see more in Appendix 8) and nearly all of them knew one or two more foreign languages in addition to Turkish and had been living in Turkey for 9 months.

The second group of participants consisted of 41 second and third grade undergraduate students. Semi-structured interview forms were completed by them. They had been living in Turkey for more than 3 years and all of them studied Turkish for a year at the Turkish Teaching Centers in the first year. The age of the second group of participants ranged between 20 and 35, with an average of 28 years. They were from 14 different countries and studying Law, Finance, Engineering, History, Medicine and International Relations.

Data Analysis

In addition to an academic corpus development procedures and general and academic wordlist development procedures, both quantitative and qualitative analyses of the data were included in the research since it had a mixed-methods design. The Statistical Package for Social Sciences (SPSS) version 20.0 was used to conduct descriptive statistics (frequency, percentage, the maximum and minimum scores, means, and standard deviations) and reliability analysis of the Academic Vocabulary Test. The qualitative data from interviews was analyzed to elaborate the opinions of the participants. The interview forms were translated for data analysis. The results were organized according to the emerging themes of the interviews and presented in an interpretive narrative style.

CHAPTER 4

FINDINGS AND DISCUSSION

Introduction

Chapter 3 presented a description of the methods and the tools used in this study to explore the following research questions:

1) What is the basic Turkish vocabulary that could be aimed at before starting to learn academic vocabulary in terms of Turkish as a foreign language?

2) What are the most frequent academic Turkish words that need to be known in order to be able to follow courses at universities offering Turkish medium instruction in Turkey?

3) To what extent do students studying Turkish at Turkish Teaching Centers know the words for academic Turkish?

4) Do foreign undergraduate students who studied Turkish at the Turkish Teaching Centers need additional academic Turkish vocabulary instruction to improve their Turkish vocabulary knowledge so that they can follow academic courses?

This chapter will demonstrate the findings on: 1) The General Turkish Wordlist, 2) The Academic Turkish Wordlist, 3) Quantitative research findings (The Academic Turkish Vocabulary Test), and 4) Qualitative research findings (Semi-structured Interviews).

The General Turkish Wordlist

The first 11000 most frequent words from the Turkish National Corpus (Aksan et al, 2013) were chosen to generate the General Turkish Wordlist because the Turkish National Corpus is a balanced and representative Turkish corpus. It also includes "transcriptions from spoken data, which constitute 2% of TNC's database" and "the distribution of the number of words in the corpus is determined proportionally for each text domain, time and medium of text following the model of BNC." Appendix 9 shows the composition of the written component and domains of the Turkish National Corpus.

After all proper names, numbers, abbreviations, letters, misspelled words, and question particles were removed from the first 11000 words of the Turkish National Corpus, the remaining words were grouped by lemma. There were 3614 lemmas and the first 2000 of them were presented as the General Turkish Wordlist in this study. PosTagger, one of the TS Corpus Tools (TS Corpus, 2012), was used to define the classes of the words (part of speech) and the list was also checked manually in terms of the classes of the words since PosTagger can parse the words with a success of 80-82 % (See more for the General Turkish Word List in Appendix 10). Table 8 shows the distribution of the words according to the classes in the General Turkish Wordlist.

Table 8. The Distribution of the Words in the General Turkish Wordlist

	N	%
Verbs	207	10.35
Nouns	1198	59.9
Adjectives	409	20.5
Adverbs	92	4.55
Other	94	4.7
Total	**2000**	**100**

In the table above, "Other" indicates conjunctions, pronouns, determiners, interjections, postpositions and slangs. Actually, most of them are function words. The number of these

kinds of words is relatively low compared to the content words in the list. The list of the first 100 words of each class (part of speech) can be seen in Appendix 11.

Table 9 compares the first most frequent 10 words from Dictionary of Word Frequency of Written Turkish (Göz, 2003) with Ts Corpus and the General Turkish Wordlist. As it can be inferred from the Table 9, all the words except "*bir*" (one) exist in all lists in similar ranks. The General Turkish Wordlist does not include "*bir*" (one) because it was assumed as a number so it was omitted from the list.

Table 9. Comparison of the First 10 Words

Dictionary of Word Frequency of Written Turkish		Ts Corpus		General Turkish Wordlist	
1	bir	3	Bir	not in the list	bir
2	ve	1	Ve	1	ve
3	olmak	2	Olmak	2	olmak
4	bu	4	Bu	3	bu
5	için	10	İçin	6	için
6	ben	22	Ben	8	ben
7	demek	5	demek	10	demek
8	çok	12	Çok	9	çok
9	yapmak	7	yapmak	7	yapmak
10	ne	28	Ne	14	ne

In order to demonstrate the coverage of the General Turkish Wordlist, the frequencies of the first 1000 words (lemmas) and the second 1000 words were accumulated. Table 10 shows the cumulative number of frequencies of the first and the second 1000 words and the percentages of the words in the General Turkish Wordlist.

Table 10. The Coverage of the General Turkish Wordlist

	N	%
The First 1000 words	23779386	49.54
The Second 1000 words	2319050	4.83
Total (The First 2000 Words)	**26098436**	**54.37**

The first 1000 words cover nearly half of the words in the Turkish National Corpus which has 48000000 words. Therefore, it can be assumed that the first 1000 words cover the 50 % of a non-academic text such as a novel or a short story.

The Academic Turkish Wordlist

As explained earlier in Chapter 3, a lot of steps needed to be taken in order to generate the Academic Turkish Wordlist and all the procedure took a lot of time to complete. Notably, 57 textbooks were chosen from five different disciplines and an academic corpus containing 4032442 words was formed. After unwanted and misspelled words were excluded from the corpus for the first time, there remained 3415378 words in the Academic Corpus (see Appendix 2 and 3). RANGE software was used to provide a range or distribution figure (how many times the word occurs in five different categories or disciplines) and a frequency figure for each of the texts the word occurs in. The Range output file was used as a databank for TrLister which is a software written for this study to group words by lemmas. There were 172156 different types having different frequency rates in the databank and they were grouped by lemmas manually. In the end 140349 words were grouped under 15456 lemmas which were typed manually into TrLister and 31807 words were listed out because of either being misspelled words or being content-specific words such as "kardiovasküler", "karyoplazma" or being proper nouns. Other calculations were done taking into consideration 15456 lemmas.

The main aim was to have an academic wordlist containing nearly 1000 words. Therefore, the frequency rate was defined as 22 and the range was 4, in other words, the words (lemmas) in the Academic Turkish Wordlist occurred at least 22 times in each category and they occurred at least in 4 categories and the first 300 words from the General Turkish Wordlist excluded

from the list. Finally 1115 words (lemmas) were determined and some of the words which were thought to be inappropriate in the academic list were cleared out and 1010 words were included in the list. Appendix 12 shows the Academic Turkish Wordlist.

It is noteworthy that West (1953) used both a quantitative and a qualitative approach in designing the General Service List. Having said that, the quantitative paradigm is more of usage for the purpose of developing wordlists by practitioners. West's quantitative approach involved word frequency and his qualitative criteria contained some principles which included: 1) the ease of learning, 2) necessity, 3) cover, and 4) stylistic and emotional neutrality. This qualitative criteria add a great deal of subjectivity into his product. Conversely, Coxhead (2000) selected the words in "The New Academic Wordlist" based on three criteria:

> 1. Specialized occurrence: The word families included had to be outside the first 2,000 most frequently occurring words of English, as represented by West's (1953) GSL.
>
> 2. Range: A member of a word family had to occur at least 10 times in each of the four main sections of the corpus and in 15 or more of the 28 subject areas.
>
> 3. Frequency: Members of a word family had to occur at least 100 times in the Academic Corpus. (p. 221)

In a recent study by Gardner & Davies (2013), "A New Academic Vocabulary List" was developed using these principles: 1) Ratio 2) Range 3) Dispersion 4) Discipline Measure. In another study by Brezina & Gablasova, (2013), developing a core general vocabulary list was targeted by comparing wordlists based on the four corpora (LOB, BNC, BE06, and EnTenTen12). To sum up, a totally quantitative research design is accepted by most scholars in developing general or academic wordlist (Coxhead, 2000; Gardner & Davies, 2013; Brezina & Gablasova, 2013) in literature. Similarly, a quantitative measure of word frequency and range was used in this study. However, listing out 105 words added a slight subjectivity to our study. These words included some erroneous words like *edici, profosyonel, olunmak* which were misgrouped in TrLister software and some content-specific words like *elektrik, elektronik, enerji, anayasa, piyasa, hukuk* which had a higher frequency in one category, and

some basic level general words like *hayvan, kardeş, beyaz*. In fact, these words were excluded objectively.

Table 11 shows the distribution of the words according to their classes (part of speech) in the Academic Turkish Wordlist in comparison with the General Turkish Wordlist.

Table 11. The Distribution of the Words in the Academic Turkish Wordlist

	Academic Turkish Wordlist		General Turkish Wordlist	
	Number of Words	%	Number of Words	%
Verbs	287	28.41	207	10.35
Nouns	496	49.10	1198	59.9
Adjectives	200	19.80	409	20.5
Adverbs	19	1.88	92	4.55
Other	8	0.79	94	4.7
Total	**1010**	**100**	**2000**	**100**

According to the table above, it can easily be understood that nearly half of the words in the academic wordlist are nouns and more than a quarter of the words are verbs and adjectives are less than verbs in the academic wordlist. However, the percentage of nouns in the general wordlist is higher but the percentage of the verbs in the general wordlist is lower than those in the academic wordlist. The percentages of adjectives in both lists are nearly same (19.80% in academic, 20,5% in general wordlist) and there is a slight difference in terms of adverbs and other kinds of words. The significant difference is between the percentages of the verbs. The percentage of the verbs in the Academic Turkish Wordlist is almost three times higher than the percentage of the verbs in the General Turkish Wordlist.

In order to demonstrate the coverage of the Academic Turkish Wordlist, the frequencies of the lemmas were accumulated. Table 12 shows the frequency and the percentage of words in the Academic Turkish Wordlist. The first 100 words constitute approximately 11% of the Academic Corpus and second 100 words constitute 6% of the corpus. All the words in the list

cover nearly one third of the words in the Academic Corpus which has 3415378 words. Therefore, it can be assumed that all words in the list cover approximately one third of an academic text at undergraduate level such as a textbook.

Table 12. Coverage of the Words in the Academic Turkish Wordlist

Words	N	%
0-100	379420	11.11
100-200	203610	5.96
200-300	140395	4.11
300-400	104045	3.05
400-500	80596	2.36
500-600	64420	1.89
600-700	50920	1.49
700-800	40234	1.18
800-900	30835	0.90
900-1010	23158	0.68
Total (0-1010)	**1117633**	**32.71**

Table 13 below shows the distribution of the words in the Academic Turkish Wordlist in terms of part of speech. Unsurprisingly, nouns constitute the biggest part of the total tokens of the list and verbs comes next. They together make up the 83% of the total tokens of the list.

Table 13. The Distribution of the Words in the Academic Turkish Wordlist in terms of Part of Speech

	N	Total Tokens	%
Verbs	287	294368	26.34
Nouns	496	628557	56.24
Adjectives	200	161656	14.46
Adverbs	19	15489	1.39
Other	8	17563	1.57
Total	**1010**	**1117633**	**100.00**

Quantitative Research Findings

To answer to the third research question, the results of the quantitative data of the Academic Turkish Vocabulary Test are presented in this section.

The Vocabulary Size Test proposed by Nation & Beglar (2007) had ten items from each 1000 word level and they used only words from the first 1000 of West's (1953) General Service List for the first and second 1000 word levels. Their distractors were the same part of speech as the correct answer. Unlike the Vocabulary Size Test, all items of the Academic Turkish Vocabulary Test were chosen from the Academic Corpus since not only knowing the meaning of the words but also understanding the items (sentences from the Academic Turkish Corpus) was targeted in this study. Items consisted of only one or two sentences from the Academic Corpus. On the other hand, the distractors in the Academic Turkish Vocabulary Test belonged to the same part of speech with the correct answer in a similar manner with the Vocabulary Size Test.

Before participants started to answer the Academic Turkish Vocabulary Test, they were asked to give their opinions on their Turkish Language Competency using a three-point scale regarding four language skills. Table 14 shows participants' responses for their Turkish Language Competency.

Table 14. Participants' Responses for Their Turkish Language Competency

What do you think about your Turkish Language Competency?								
	Speaking		**Listening**		**Reading**		**Writing**	
Beginner	16	8%	13	8%	9	5%	12	7%
Pre-Intermediate	108	60%	98	58%	88	53%	107	65%
Advanced	54	30%	57	33%	68	41%	46	28%
Total	**178**	**100%**	**168**	**100%**	**165**	**100%**	**165**	**100%**

Although there were 196 participants who answered the Academic Turkish Vocabulary Test, some of them (nearly 10%) did not give their opinions about their Turkish Language Competency. All of them were studying TFL for 9 months and they were using a C1 level coursebook.

In terms of speaking skill, only 8% of students (n=16) expressed that their speaking skill was at the beginner level. Since they had been living in a Turkish speaking environment for nine months, the reason why 16 students thought that their speaking skill was at beginner level could be explained by psychological factors like being introvert or less willing to communicate. 60% of the students thought that their speaking skill was at moderate level and the rest of the students (30%) thought that their speaking skill was at the advanced level

In terms of listening skill, 8% of the students asserted that their listening skill was at the beginner level and the rest of the students thought that their listening skill was at moderate or advanced level. The findings for the writing skill was nearly the same as the speaking and listening skill (7% beginner, 65% moderate and 28% advanced). The significant difference was between reading skill and others. Merely 5% of the students (n=9) expressed having beginner level reading skill and 53% of the students (n=88) thought that they were at moderate level and 41% of them (n=68) claimed to be at advanced level in terms of reading skill. According to students' opinions for their reading skill it could be hoped that they would get higher scores from the Academic Vocabulary Test since they were studying a C1 level coursebook and that means they could use language effectively for academic and professional purposes.

The answers of 196 students studying Turkish at the Turkish Teaching Centers with the goal of studying at a Turkish-medium university included in the quantitative analyses of the Academic Vocabulary Test. Table 15 indicates the means and standard deviations of each individual item in a decreasing order of the means. As seen in the Table 15, three highest mean scores were .80 (Item 3), .73 (Item 30) and .71 (Item 2) and 3 lowest mean scores were .21 (Item 25), and .23 (Item 8), and .26 (Item 24).

Table 15. The Statistical Analysis of the Academic Vocabulary Test Items

	Mean	SD		Mean	SD
Item 3	.80	.40	Item 22	.48	.50
Item 30	.73	.44	Item 1	.47	.50
Item 2	.71	.45	Item 29	.46	.50
Item 13	.65	.47	Item 17	.45	.49
Item 21	.65	.47	Item 14	.42	.49
Item 6	.63	.48	Item 5	.42	.49
Item 10	.60	.49	Item 7	.40	.49
Item 11	.60	.49	Item 23	.37	.48
Item 27	.57	.49	Item 26	.35	.47
Item 28	.57	.49	Item 9	.34	.47
Item 4	.55	.49	Item 16	.33	.47
Item 12	.53	.50	Item 20	.32	.46
Item 19	.52	.50	Item 24	.26	.44
Item 18	.51	.50	Item 8	.23	.42
Item 15	.49	.50	Item 25	.21	.41

Item 3 which tested the word *ihtiyaç* (necessity) had the highest mean score because it is both a frequently used word in the academic corpus and it can also be encountered in everyday speech. Item 30 tested *davranışsal* (behavioral) which is the least frequent word in the vocabulary test. It is the 945th word in the Academic Wordlist. Surprisingly, it had the second

highest mean score. This could be explained by having the word *problem* in the sentence (item 30) which collocates the word *davranışsal* (behavioral). Item 3 which tested the word *araştırma* (research) had the third highest score. Similar to item 30, *araştırma* (research) had collocations in the item, which was prepared to test it.

Item 25 had the lowest mean score and this item had been criticized by language instructors who had checked the Academic Vocabulary Test. In the basis of their complaints for this item was that the item tested the figurative meaning of the target academic word *iklim* (climate-atmosphere). Therefore, the sentences in the academic corpus which included *iklim* were analyzed and it was found out that the figurative meaning of the word *iklim* was mostly used in the textbooks. To sum up, the reason as to why item 25 had the lowest mean score, it could be testing the figurative meaning of target word and most of the students were unable to guess the correct answer. Item 8 and Item 24 tested *yürütmek* (execute) and *kanıtlamak* (prove) respectively. They seem to be typical academic words and the word *yürütmek* (execute) is the 140th word in the Academic Turkish Wordlist and the 1050th word in the General Turkish Wordlist and *kanıtlamak* (prove) exists in the Academic Turkish Wordlist in the 745th place.

Table 16 indicates that students demonstrated a moderate level of achievement in the Academic Turkish Vocabulary Test.

Table 16. The Academic Vocabulary Test

	N of Items	Mean	%
AVT	30	14,7551	0.49

Participants answered nearly half of the test items correctly. This result reveals the fact that students could know approximately half of the words in The Academic Turkish Wordlist. However, for the sufficient comprehension level, which is 71% according to Hu and Nation (2000) and 68% according to Schmitt et. al. (2011), participants are required to have

vocabulary knowledge that covers at least 98% of the text and in our study 196 students' average vocabulary knowledge seemed to be below 50%.

The six universities where the research was carried out employed three different Turkish course books. Every two university used the same set of books to teach Turkish to the foreign students. The main coursebooks which were put into use are as follows accompanying with the number of the students who participated in the research:

1. Gazi Üniversitesi TÖMER Yabancılar İçin Türkçe Seti (n=93) (henceforth Gazi Series)

2. İstanbul Üniversitesi Dil Merkezi İstanbul Seti (n=66) (henceforth Istanbul Series)

3. Ankara Üniversitesi TÖMER Yeni Hitit Yabancılar İçin Türkçe Öğretim Seti (n=37) (henceforth Ankara New Hitit Series)

The success rate of the students who used Gazi Series was 42%, which seems to be moderate regarding the Academic Turkish Vocabulary Test. The achievement rate of the participants who were instructed through Ankara New Hitit Series was almost the same and were reported to be 39%. Nevertheless, the universities in which the students had used Istanbul Series books accomplished a higher rate of success and turned out to be 63%. This report indicates that 34% of the participants (n:66) taking Turkish lessons in their classes were higher achievers. Thus, the findings shown in Table 17 shows that only the competency of the students who employed the İstanbul Series were closer to the sufficient comprehension level. Hu and Nation (2000) identified students at the level of 68% and above as ones possessing sufficient comprehension level. However, assuming the yardstick as that of the Schmitt and et al's (2011) the target seems far from available. They most recently designated the definition for the above mentioned level as 71% with a vocabulary knowledge that covers at least 98% of the text.

Table 17. Coursebooks Used at The Turkish Teaching Centers

	N	Mean	SD
Gazi Series	93	.4258	.15410
Istanbul Series	66	.6379	.20042
Ankara New Hitit Series	37	.3973	.20748
Total	**196**	**.4918**	**.20865**

The results and findings reported above can be a consequence of different factors and cannot be justified as the quality of the course books. Some other elements can be found influencing the learning process namely, the class atmosphere, school educational policies and whether the students did the test themselves or in company of others. The ANOVA results of the current analysis engaging the three different course books in terms of the student performance is given below in Table 18.

Table 18. The ANOVA Analysis of the Coursebooks

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	2.144	2	1.072	32.604	.000
Within Groups	6.345	193	.033		
Total	**8.489**	**195**			

According to the table above a significant difference can be seen among the three sets of books in terms the students' performance. The details regarding the multiple comparisons made among the groups are given in Appendix 14.

Qualitative Research Findings

In this section, a qualitative description of the findings explored by qualitative method and tool (semi-structured interviews) will be presented to answer the following research question:

-Do foreign undergraduate students who previously studied Turkish at TTCs need academic Turkish vocabulary instruction to improve their knowledge of Turkish vocabulary in order to follow academic courses?

As mentioned in Chapter 3, the qualitative aspect of the study, which addressed the fourth research question above, deals with qualitative scrutiny of the interview forms to expand and elaborate the quantitative results of the research. In this chapter, the results are organized according to the emerging themes of the interviews, and presented in an interpretive narrative style.

There were four emerging themes from the interviews and the finds regarding these themes are explained below on an individual basis:

1) How do students evaluate themselves in terms of reading comprehension?

2) How do students evaluate themselves in terms of understanding the lectures and the assignments?

3) How do the students self evaluate themselves in communicating in turkish on and off the campus?

4) How did the students explain the difficulties they encountered in a university with turkish as the medium of instruction?

How Do Students Evaluate Themselves In Terms Of Reading Comprehension?

When the students were asked about whether they understood what they had read in such Turkish materials as short stories, novels, newspapers, they claimed that they were overall capable of understanding what they read although their reading comprehension level was supposed to be below the desired level (according to the quantitative results of the current study). They asserted that they were able to read simple texts and the ones which did not

contain unknown words more comfortably and added that they gradually improved their reading skills.

Furthermore, 15 students claimed that they completely understood what they read and one participant said that he did not understand what he read. The participants who were able to understand completely what they read chose moderate or advanced level for four skills when they were asked "How well do you use Turkish?" in terms of four language skills, and the participant who said that he did not understand what he read stated being beginner level in each language level. Some of the participants' responses are given below:

P1 If the content includes normal and clean sentences and simple words, I can understand.

P11 Yes, Sometimes I read story books in my leisure time. But, I had a lot of trouble in the first year. Thank God, I know better now. I can easily solve my problems and I am still solving my problems.

P31 I almost understand but I do not know special expressions so I have to look them up in the dictionary. Moreover, the foreign words from western languages ....

P2 I understand quite well

P33 No. Because it is too difficult. I cannot understand. I cannot read.

When students were asked to make a judgment on their understanding of textbooks, 17 of them stated that they were able to understand the textbooks (they used some expressions like *in an advanced level, mostly, and 80% or 90%*). Five of the participants stated that they had difficulty in understanding their textbooks. Their Turkish competency was either at a beginner or a moderate level. The general view of all the participants was that the students had mixed feelings regarding understanding the textbooks. From among the students who stated that they were able to understand their textbooks completely, there were ones who said that they had difficulty in understanding foreign words or some certain words. The students expressed that they did not know all the words in their textbooks and used dictionaries to understand.

Moreover, they admitted that they gradually understood more since they spent a lot of time on their textbooks. Some of the students responded as follows:

P13 Sometimes I do not understand my textbooks at all since they are too difficult. Everything is also new for me.

P27 Yes. I understand my textbooks better than other books. I hardly encounter unknown words.

P37 I moderately understand. I do not know 60-70% of the words. The terms in textbooks are too difficult for me.

P41 Yes. But not much. I do not know all the words.

It is possible to conclude that the students had a moderate level of reading comprehension skill. This finding also supports the quantitative results of the study. The students' expressing their ability in understanding the self study materials and the activities they did which engaged them in reading for pleasure implies the fact that the more they feel stimulated to learn to satisfy their inner feelings the better they can understand the passages. This feeling seems to be justifiable even for the case of native speakers of any language, where they could admit they are capable of understanding any text they are keen on without any difficulty. People would feel easy reading and understanding a text where they enjoy doing so which can also be called a bed-time reading per se.

However, the number of the students who reported to be struggling to comprehend the text was not rather noticeable (5 out of 41). This is conclusive to the fact that after three painstaking years of learning textbook contents it seems rather predictable that they will not have trouble in mastering the terminologies. However, from among them, there could be found some people who complained about a number of words which were challenging to them.

How Do Students Evaluate Themselves In Terms Of Understanding The Lectures And The Assignments?

During the interviews, the students were asked whether they understood the lectures given by their professors. Taking into account the responses for the question "Do you understand your lecturers?", it is possible to say that the participants generally understand their lecturers. They expressed that their lecturers' way of talking was understandable when they took the professors' speaking speed for granted. They said that they improved their listening skill through watching movies, communicating with friends, listening to music and reading books. Only 2 students said that they did not understand their lecturers at all and 6 students said they moderately understood and had little difficulty in understanding. Some of the responses are given below:

P1 Yes, at a level of 85%. It depends on the pronunciation and the speaking speed of the lecturer. At the same time, it also depends on my prior knowledge about the topic in some specific lectures.

P13 Sometimes I understand, sometimes I do not understand my lecturers at all. So I sometimes have problems. It is difficult to understand lectures in foreign language.

P38 No, I do not understand much.

Additionally, students were asked whether they had any difficulties in preparing projects or answering the exam paper (mid-term or final exam paper) in Turkish in terms of language. Only 6 students expressed that they did not have any difficulties in preparing projects or in completing exam papers. These students also added that they had had difficulty before but gradually they got used to it and then they could overcome the problem. Other students said that they had problems regarding exams and projects. They mentioned these problems as vocabulary insufficiency, sentence formation problems, time management problem due to difficulty in reading comprehension, failure in understanding the questions, being Latin alphabet ignorant, and sluggish writing. Moreover, the students had less difficulty in preparing

assignments because of the assistance they received from their circles and friends around. Some of their responses are brought here:

P10: I don't understand some questions in examinations. Besides, I leave them blank. Since there are words that I do not know. I understand some questions only partially.

P05 Foreign students experience problems in written exams. Turkish language is such a difficult language to write in.

P15 I remember the time when I first entered the classroom and I had big problems in writing. I was able to resolve them over time though.

They maintain that although they have such a slight difficulty in understanding the lecturers, they cannot express themselves as they are supposed to during the exams. This finding implies that the students have more difficulties in writing than in communication orally. This can lead the research agendas to investigations in written mode as opposed to the spoken mode which is more common and has become a household research in SLA field.

There are some points that can be hypothesized in regard to difficulties students have in listening comprehension. Students have particularly got this problem in their second or foreign language. People are able to comprehend a conversation on a noisy street in their mother tongue; however, they seem unable to transfer that skill comfortably to a second language. One way of tackling this problem is to show the foreign learners of Turkish how to spot the important words that they will encounter during listening. This is recommended to be done when they are taking lessons earlier when preparing them for the undergraduate programs. Nevertheless, this is feasible by incorporating "The Academic Turkish Wordlist" and "The Academic Corpus" into mainstream education.

How Do The Students Self Evaluate Themselves In Communicating In Turkish On and Off The Campus?

Firstly, the students were asked about their experiences regarding their in-class communication in Turkish with their peers and teachers. Some students (n=10) reported that they experienced a lot of difficulties in communicating in Turkish language in the classroom. To put it generally, these drawbacks were admitted in the following areas: understanding what is heard but unable in verbalizing what they needed to utter, having fluent communicative relationship with their peers but having an inhibitive relationship with their lecturers and trouble making accurate sentences. The other students reported having troubles in communication every now and then but not any serious break-downs in communication in total. Some of their responses are:

P11 I am bad at communicating with others. Moreover, I am not a sociable person.

P16 I understand the person I'm talking to pretty well but I cannot recall some of the words immediately when I want to speak so I sometimes have difficulty.

P18 Yes, especially with friends. However, I have problems with my lecturers regarding pertaining to the subject matter. We ask questions but the lecturer does not understand or he/she asks but we do not understand.

P34 Yes. I am able to communicate. Since the students in class are Turkish, I communicate using the same language and ask my questions.

Students were also asked about their experiences regarding their communication in Turkish in friendly atmospheres outside the classroom. 7 students expressed that they were unable to communicate because of various reasons namely, preferring English or another language, acting in accordance with the situation they are in, and having troubles in communicating using Turkish. Other students said that they were able to communicate in Turkish although they had various problems. Some of their responses are brought here:

P12 I say "Hi" and start the conversation. Yes, I am able to communicate. I chat with others.

P15 I can communicate in every context whether in a social environment or not.

P14 Yes, I can. However, I am sometimes unable to communicate with some people since I do not have a good mastery of Turkish.

P33 No, because I feel unqualified when it comes to talk to them.

It can be concluded that students are successful in communicating through Turkish language in and out of the classroom although they have some kind of difficulties. It goes without saying that every time in the past-related literature, foreign language learners have most problems in producing speech rather than comprehending. However, the fact that they are more comfortable talking in Turkish outside the classroom rather than inside the classroom is more related to their learning experience they had in the community. The learners spent most of their time in the settings which are informal and which are not educational like the conversations they had in the cafes or in the street.

How Did The Students Explain The Difficulties They Encountered In A University With Turkish As The Medium of Instruction?

Finally, students were asked about what kind of difficulties they encountered studying at a university offering Turkish-medium instruction. Some students (n=6) stated that they did not have any difficulties and added that they were able to overcome the difficulties in time.

By the time responses of other students are examined, it can be understood that they have problems due to having a low mastery of Turkish though there is a wide variety of statements:

Table 19. Problems Encountered by the Students

Students' Statements	N
Hearers' inability to understand me	3
Inability to ask lecturers questions	3
Low mastery of the language	3
Inability to understand the lecturer	2
Inability to understand the text when read	2
Unknown words	2
Having problems in the exams	2
Spending too much time for understanding the subject matter	2
Writing	2
Inability to read books/textbooks	2
Pronunciation	1
Inability to learn by heart	1
Inability to follow Turkish resources	1
Inability to do research	1
Inability to understand lectures	1
Spending too much time for reading comprehension	1
Polysemous Words	1
Inability to understand the speaker	1
Low mastery of grammar	1
Turkish is a difficult language	1
Inability to understand the concepts	1
Afraid of being made fun of	1
Feeling like an outsider	1
Cultural differences	1

3 students explained that they were afraid of being made fun of and would feel alienated partly due to the cultural differences. In fact, it can be thought that having low mastery of Turkish lies at the bottom of their problems. Some of the participants responded as follows:

P17 Before I began to study in my department, I spent a year in order to learn Turkish and another year to learn English as compulsory preparation courses. Therefore, I tried to learn Turkish for two years in Turkey and I do not have any problems regarding Turkish language. If I had learned Turkish for only one year, I could not have understood both my lecturers and my textbooks. I would certainly fail the class. My friends who have studied at the Turkish Teaching Centers have a lot of problems.

P16 Since I have difficulty in speaking, I prefer being silent. Sometimes I cannot even ask my lecturers any questions. Because my friends will laugh at me if I make an error in speaking.

P30 Only language. Since we are foreigners we have a lot of difficulty. For example, if a Turkish student study for 5 hours, I am not able to study like him/her in 5 hours. Unfortunately, we try a lot to understand what we are supposed to but sometimes we go crazy because we do not understand some parts of it.

CHAPTER 5

CONCLUSION

Introduction

The aim of this dissertation has been to make a research on the academic vocabulary in Turkish as a foreign language. Chapter 1 provides an overview of the purpose of the study, research questions, research limitations and definitions of some important terms. Chapter 2 explores the literature regarding corpus linguistics, corpus-based language learning, and studies based on different corpora. Chapter 3 presents an overview of the research design, data collection methods, data analysis procedures, instrumentation and study subjects. Chapter 4 presents the findings and analyses of the data gathered from the study. Chapter 5 summarizes the findings and recommendations for further research in the field.

In order to answer the first research question "What is the basic Turkish vocabulary that could be aimed at before starting to learn academic vocabulary in terms of Turkish as a foreign language?", prior corpus studies in Turkish were reviewed beforehand and the frequency list of the Turkish National Corpus (Aksan et al, 2013) was used to create the General Turkish Wordlist. Afterwards, proper names, numbers including ordinal numbers and years, abbreviations, letters and misspelled words, question particles, and exclamations were cleared out from the frequency list of Turkish National Corpus (Aksan et al, 2013), remaining words were grouped by lemmas and a general wordlist containing 2000 words was reached. The frequency list of the Turkish National Corpus (Aksan et al, 2013) also served as a basis to form exclusion lists containing proper names, numbers including ordinal numbers and years, abbreviations, letters and some misspelled words, question particles, exclamations and the

frequent words that were peculiar to general Turkish rather than academic language. These exclusion lists were used to develop an academic wordlist which would answer the second research question: "What are the most frequent academic Turkish words that need to be known in order to be able to follow courses at universities offering Turkish medium instruction in Turkey?". An academic corpus derived from fifty-seven textbooks was formed so as to discover the Academic Turkish Wordlist. There were 4032442 tokens in the academic corpus and various computer software programs some of which were written for the current study were used in order to acquire an academic wordlist and 1010 words were included in the final version of the Academic Turkish Wordlist. To answer the third research question: "To what extent do students studying Turkish at the Turkish Teaching Centers know the words for academic Turkish?", the Academic Turkish Vocabulary Test was created for the purpose of revealing the knowledge level of TFL students regarding the Academic Turkish Wordlist and the statistical findings of the vocabulary test were analyzed in Chapter 4. Second and third year undergraduate students were interviewed using a semi-structured interview form to answer the last research question: "Do foreign undergraduate students who studied Turkish at the Turkish Teaching Centers need academic Turkish vocabulary instruction to improve their Turkish vocabulary knowledge in order to follow academic courses?". Quantitative data from the interviews were also analyzed under four themes in the previous chapter.

Discussions

The General Turkish Wordlist

To put it in a nutshell, the General Turkish Wordlist consisted of 2000 words and the first 1000 words covered 49% of the Turkish National Corpus (Aksan et al, 2013) and the total words in the list covered more than the half of the corpus (54%). However, Brezina & Gablasova (2013) cited that West's list (General Service List) (West, 1953) covered more than 80% of *the Lancaster-Oslo-Bergen Corpus, The British National Corpus, The BE06 Corpus of British English* and *EnTenTen12* on an individual basis. This situation was the same for New-GSL (Brezina & Gablasova, 2013). It should be noted here that, there were 4114 lemmas (2000 word families) in West's list and 2494 lemmas in New-GSL. This factor could account

for at least some of the difference mentioned above. Another factor could be that the first most frequent 11000 words were used to develop the General Turkish Wordlist. Inflections of some of the words in the list might be outside of the first 11000 words so the frequencies of these words were not included in the cumulative coverage of the General Turkish Wordlist. Last but not least, linguistic structure of the two languages (English and Turkish) could account for the most part of the aforementioned difference. In English, every sentence needs a subject and most of them contain prepositions, auxiliary verbs, and determiners. On the other hand, in Turkish, only a word could convey the meaning of a sentence by adding suffixes to the word, which can be expressed by using several different words in English. This linguistic difference could explain the high coverage of general wordlists in English.

The Academic Turkish Wordlist

In order to create an academic wordlist, fifty-seven textbooks were chosen from five different disciplines and an academic corpus containing more than four million words was formed. Various computer software programs some of which were written for the current study were used in order to acquire an academic wordlist and 1010 words were included in the final version of the Academic Turkish Wordlist.

The main purpose was to have an academic wordlist containing nearly 1000 words. In the end, the frequency rate was defined as 22 and the range was 4, in other words, the words (lemmas) in the Academic Turkish Wordlist occurred at least 22 times in each category and they occurred at least in 4 categories and the first 300 words from the General Turkish Wordlist were excluded from the list. Then, some of the words which were thought to be inappropriate in the academic wordlist were cleared out and 1010 words (lemmas) were included in the Academic Turkish Wordlist.

The Academic Turkish Wordlist covered 32% of the academic corpus. However, Gardner & Davies (2013) stated that their academic wordlist (A New Academic Vocabulary List –AVL) covered 13.8% of the academic portion of COCA (an American corpus, from which the AVL was derived) and it also covered 13.7% of the academic portion of BNC (a British corpus, which was not used in the creation of the AVL). They also claimed that Coxhead's academic

Wordlist (2000) covered 7.2% of the academic portion of COCA and 6.9% of the academic portion of BNC. Comparing the Academic Turkish Wordlist with the other academic wordlists in English (AVL and AWL), it is evident that the coverage of the Academic Turkish Wordlist is a far-reaching one when compared to others.

One of the factors that account for some of the difference is the design of the wordlists. The Academic Turkish Wordlist had 1010 lemmas, which only had the inflections of the words, but a word family design was used to calculate the coverage of the other lists (AVL and AWL) and they consisted of 570 word families individually. Another factor could be that the academic corpus in the current study consisted of only textbooks, which were prepared for undergraduate programs. Therefore, the author's style and wording did not vary much, so the coverage of the Academic Turkish Wordlist is more than the coverage of other lists (AVL and AWL).

Knowledge Level of Students Regarding The Academic Turkish Wordlist

To answer the third research question, which was "To what extent do students studying Turkish at Turkish Teaching Centers know the words for academic Turkish?" a quantitative method was carried out and a vocabulary test was prepared to find out to what extent TFL students know the words for academic Turkish. It was a multiple choice vocabulary test including 30 items. All items of the Academic Turkish Vocabulary Test were chosen from the Academic Corpus since not only knowing the meaning of the words but also understanding the items was targeted in this study. Items consisted of only one or two sentences from the Academic Corpus. On the other hand, the distractors in the Academic Turkish Vocabulary Test belonged to the same part of speech as the correct answer in a similar manner with The Vocabulary Size Test.

The quantitative results indicated that the overall vocabulary knowledge of the participants regarding the Academic Corpus was 0.49%. However, for the sufficient comprehension level which is 71% according to Hu & Nation (2000) and 68% according to Schmitt et. al. (2011), participants are required to have vocabulary knowledge that covers at least 98% of the text and in our study 196 students' average academic vocabulary knowledge levels seemed to be below

50%. This percentage does not cover students' vocabulary knowledge regarding general words. However, it should be noted here that the Academic Turkish Wordlist covered 32% of the Academic Corpus, so knowing the most of the general Turkish words and half of the academic wordlist is still beyond the sufficient comprehension level that requires having the vocabulary knowledge that covers at least 98% of a text.

Needs Analysis For Additional Academic Vocabulary Instruction

The qualitative aspect of the study, which addressed the third research question, dealt with the qualitative scrutiny of the interview forms to expand and elaborate the quantitative results of the research. Consequently, the qualitative results of the study made the following cases known to the researcher:

a) How the students evaluated themselves in terms of reading comprehension

b) How the students evaluated themselves in terms of understanding the lectures and the assignments

c) How the students self-evaluated themselves in communicating in Turkish on and off the campus

d) How the students explained the difficulties they encountered in a university with Turkish as the medium of instruction.

In terms of reading comprehension (a) it is possible to conclude that the students had a moderate level of reading comprehension skill regarding both reading for pleasure (this also includes short stories, novels, newspapers) and reading academic texts, especially textbooks. This finding also supports the quantitative results of the study.

With regard to reading academic texts, which is the main concern of this study, less than half of the students (n:17) stated that they were able to understand textbooks. After three meticulous years of trying to comprehend the textbook contents, it seems rather predictable that they will not have trouble in mastering the terminologies. However, from among them, there were some who complained about a number of words which were challenging to them. Students' low level of academic vocabulary accounts for most of this problem and another

reason may have resulted from individual factors, namely, having a low level of reading comprehension, studying at an undesired undergraduate program, unwillingness to spend time reading or improving reading skills, etc.

The findings also revealed that the students have more trouble in writing than in oral communication (b). Surprisingly, quantitative results indicated that students' overall listening skills were sufficient to understand their professors regarding academic undergraduate lectures. Actually, they had been living in Turkey for at least three years. Learning the target language where it is spoken facilitates *passive learning*, *accent training* and *acquiring the skill intuitively* and it could also have helped them improve their listening skills. In this sense, some of the students said: "It was very difficult to understand everything at first, but there have remained no problems as the years passed by." In spite of having moderate listening comprehension skills, students admitted that they were unable to express themselves as they were supposed to during the exams and they mentioned their problems as vocabulary insufficiency, sentence formation problems, time management problems due to difficulty in reading comprehension, failure in understanding the questions, having no prior knowledge of the Latin alphabet, and sluggish writing. These drawbacks could be remedied through academic Turkish courses focusing on reading comprehension, academic writing and academic vocabulary instruction.

It seems that students are successful in communicating through Turkish in and out of the classroom (c) although they have some kind of difficulties. Living in the country where the target language is spoken and operating solely in the target language might result being comfortable in speaking Turkish.

From the responses regarding the difficulties students studying at a university with Turkish as the medium of instruction encountered (d), it can be understood that there is a wide variety of statements and their problems mainly result from having a low mastery of Turkish and from psychological or individual factors.

Conclusion

Being a corpus study, the current research study examined academic vocabulary in Turkish as a foreign language and it had four components: 1) generating the General Turkish Wordlist using the frequency list of a Turkish corpus 2) generating the Academic Corpus from textbooks and the Academic Turkish Wordlist 3) finding out the knowledge level of students regarding the Academic Turkish Wordlist 4) needs analysis for additional academic vocabulary instruction.

The General Turkish Wordlist consisted of 2000 words. The first 1000 words cover nearly half of the Turkish National Corpus, which consists of 48 million words. Therefore, this wordlist could help language practitioners, syllabus designers or material developers define where to start or which words should be taught first or later. Gairns & Redman (1986) pointed out that "The high frequency of an item is no guarantee of usefulness, but there is obviously a significant correlation between the two so it is worth examining some of the work on frequency word-counts that has been carried out over recent decades." (p. 58).

1010 words were included in the Academic Turkish Wordlist and they covered nearly one third of the Academic Turkish Corpus, which was compiled from 57 textbooks. It is evident that knowing the words in the academic list will help undergraduate foreign students with the goal of studying at a Turkish-medium university improve their academic language skills, especially reading comprehension.

The quantitative results indicated that the overall vocabulary knowledge of the participants regarding the academic corpus was below the desired level. In-sessional academic language courses, where students study while they are already attending an academic course or pre-sessional academic language courses, where students take the course before they go on to further academic studies could help them develop academic language skills efficiently. While subject specific academic language courses for the students learning Turkish as a foreign language seem to be cost-inefficient, common core academic language courses seem to be more feasible since foreign students study various subjects and grouping them under a common core academic language course is easier. If possible, pre-sessional common core

academic language courses and in-sessional subject specific academic language courses are strongly suggested.

Scrivener (2011) stated that an academic language course is likely to include work on some of these elements:

- Listening to and understanding lectures
- Note-taking
- Effective reading and researching
- Raising awareness of formal styles of language used in academic writing
- Essay writing
- Generic academic language points
- Academic stylistic conventions
- Making presentations
- Taking part in seminar discussions
- Learning to learn
- Contextual and cultural awareness (p. 315)

These elements are consistent with the problems which emerged from the qualitative results of the study. It was discovered that students' problems mainly resulted from having a low mastery of Turkish and academic language courses offering academic course materials would fill in this gap. Stubbs (2001) concluded that:

> Corpus study shows that language use is much more highly organized than previously suspected. It is governed, not by the kinds of categorical rules that linguistics has often dealt with, but by tendencies and probabilities. Only with large corpora and appropriate software is it possible to observe repeated patterns across the language use of many speakers and writers (p. 242)

Therefore, practitioners could utilize the General and Academic Turkish Wordlists developed in this study in various ways: 1) to choose the words for beginner level students as a starting point 2) to check out language learning materials in order to find out how many low/high frequency words (general or academic) they contain to find out the appropriateness of the material for the vocabulary level of their students 3) to compare students' writings in terms of how many academic words they are able to use 4) to prepare proficiency exams 5) to revise their language program in order to include more academic words to prepare their students for undergraduate programs.

Pedagogical Implications of The Study

Whereas this study did not explore the instruction of academic words, the results are suggestive in terms of C1 level Turkish learners' knowledge level of academic words. They lead to four pedagogical implications for the need of additional academic vocabulary instruction. First, C1 level Turkish learners need more exposure to academic words before they begin to study at their undergraduate program since they have a moderate knowledge level of academic words in Turkish. Second, Turkish learners may need explicit instruction on how to effectively use academic words in writing. Academic texts should be provided to students to be used as models for their own writing since undergraduate students have great difficulty in written examinations. Third, they need more exposure to and practice with academic words in terms of improving their reading comprehension skills. Lastly, practitioners should be aware of the individual factors that might cause communication breakdowns.

Theoretical Implications of The Study

According to "*The Languages for the Future report*", Turkish is one of the ten languages (Spanish, Arabic, French, Mandarin Chinese, German, Portuguese, Italian, Russian, Turkish and Japanese) which emerge as the most important for the UK over the next 20 years (British Council, 2014). These languages, which were thought to be the most vital to the UK, were chosen based on "economic, geopolitical, cultural and educational factors including the needs of UK businesses, the UK's overseas trade targets, diplomatic and security priorities, and prevalence on the internet". Therefore, it is important to find out the effectiveness of teaching Turkish as a foreign language at the Turkish Teaching Centers since the main goal of these institutions is to prepare learners for their studies at Turkish medium graduate or undergraduate programs. From a theoretical perspective, this study addresses the need for additional academic vocabulary instruction for learners of Turkish in terms of reaching the sufficient comprehension level and that the Academic Turkish Wordlist developed from the Academic Corpus compiled from textbooks could serve as a unique source for academic Turkish courses and course materials and it can also be used for further research in the field.

Limitations of The Study

While several billion words of English are scanned in The Cambridge English Corpus, The British National Corpus (BNC) is a 100 million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of current British English, both spoken and written (http://www.natcorp.ox.ac.uk/). On the other hand, nearly 50 million words are scanned in the Turkish National Corpus, which is considered to be efficient enough for the current study in order to make up a general Turkish wordlist. Therefore, the first most frequent 11000 words of Turkish National Corpus were used to develop the General Turkish Wordlist.

Fifty-seven textbooks from the E-Learning Portal of Anadolu University were used to develop the Academic Corpus. Because the study participants consisted of undergraduate students and students studying at Turkish Teaching Centers, only textbooks were used to make up an academic corpus. The words in the academic wordlist may differ when they are collected from another corpus derived from another source.

Lemmas, or the groups of lexical forms with the same stem that belong to the same word class (Francis & Kucera, 1982) were used for the creation of the General Turkish Wordlist and the Academic Turkish Wordlist. They are understood by learners as follows: words with the same base and inflections are members of the same word family as defined by Level 2 of Bauer and Nation's (1993) scale. Further research can focus on the next level (Level 3) in which words are considered as members of a single lexical unit provided that they have the same base and contain inflections and/or derivational affixes. This level contains suffixes such as *–able, -er, -ish, -less, -ly, -ness, -th,-y, non-,* and *un-*. However, it should be found out how much learners of Turkish are aware of these suffixes at first, and then further research can focus on the following level.

Forty-one undergraduate students and two hundred and eight students studying at Turkish Teaching Centers of seven different universities (the total number of participants was 249) participated in the study. Thus, it may not be appropriate to generalize the results to all students studying Turkish as a foreign language.

All the items (n=30) asked and distractors to these items in the Academic Turkish Vocabulary Test were from the Academic Corpus, which was developed for this study and this test was used to evaluate the knowledge level of students regarding the Academic Turkish Wordlist.

Suggestions For Further Research

It is logically expected that the research on this field is still immature and there are still vast areas that have not been yet taken into consideration since this is one of the few studies of its kind conducted on the academic vocabulary of students of Turkish as a foreign language via a corpus-based approach.

In view of the previous remarks, further research is definitely needed to:

1) develop a comprehensive academic Turkish corpus which includes various academic texts from various domains,

2) develop academic and general corpora in which all the words are tagged regarding stems, affixes, domains and frequencies of the words,

3) investigate lexical and grammatical collocations in general and academic corpora as well as learners' free writing,

4) create a dictionary of the problematic words that foreign students of Turkish are likely to encounter at different phases of their second language mastery,

5) build a syllabus that meets learners' lexical needs regarding academic Turkish vocabulary in the light of corpus studies,

6) examine the TFL (Turkish as foreign language) coursebooks in terms of their effectiveness on preparing students for their undergraduate and graduate studies,

7) develop a corpus-based academic Turkish learner's dictionary including the frequency of the words and the morphological and collocational information of its items

8) focus on the reasons why foreign students studying at a university offering Turkish medium instruction have great difficulty in examinations.

Corpus-based research can be conducted when a corpus is formed based on input from various domains and sources. Applying corpus analysis to vocabulary instruction would also be beneficial for improving the vocabulary knowledge of the students and a good example for this is using concordances. Using a concordance, specific functions could be utilized, such as word frequency, sorting to find patterns, collocational patterns and tracking distribution in texts. Furthermore, future studies could focus on corpus-based analysis in Turkish after a large, balanced Turkish Corpus like COCA or BNC which gives detailed information about its tokens is constructed.

As for curriculum and syllabus designing, it is sufficiently evident from the current study that Turkish learners need additional in-sessional and/or pre-sessional academic vocabulary instruction. Therefore, practitioners, syllabus designers and coursebook writers should benefit from corpus studies and the general/academic wordlists generated in this study.

REFERENCES

Açık, F. (2013). *Temel Türkçe (a1/ a2) için söz dağarcığı tespit denemesi. Abdurrahman Güzel için Armağan Kitabı*. Ankara: Akçağ.

Adolphs, S., & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics*, *24*(4), 425-438.

Ağçam, R. (2014). *A corpus-based study on author stance in academic English*. Phd Dissertation. Çukurova University Institute of Social Sciences, Adana.

Aksan, Y. et al. (2012, May). *Construction of the Turkish National Corpus (TNC)*. In Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012). Istanbul Lüfti Kirdar Convention & Exhibition Centre, İstanbul. Retrieved from http://www.lrec-conf.org/proceedings/lrec2012/papers.html on 12 Januay, 2014.

Allan, Q. G. (2002). The TELEC secondary learner corpus: A resource for teacher development. In S. Granger, J. Hung & S. Petch-Tyson (Eds.)*, Computer learner corpora, second language acquisition and foreign language teaching*. (pp. 195-211). Amsterdam: John Benjamins.

Aston, G. (2001). Learning with corpora: An overview. In G. Aston (Ed.), *Learning with corpora* (pp. 1-45). Houston, TX: Athelstan.

Baker, P. (2006). *Using corpora in discourse analysis*. London: Continuum.

Baldwin, T. (2005, April). *Looking for prepositional verbs in corpus data*. In Proceedings of the 2005 Meeting of the Association for Computational Linguistics: Workshop on the Linguistic Dimensions of Prepositions and their Use in Computational Linguistics Formalisms and Applications (pp. 115-126). Colchester, UK.

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

Bennett, G. (2010). *Using corpora in the language learning classroom*. Ann Arbor: University of Michigan.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: investigating language structure and use*. Cambridge: CUP.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. New York: Longman.

Bonelli, E. T. (2012). Theoretical overview of the evolution of corpus linguistics. In A. M. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (p. 14-27). Routledge.

Brezina, V., & Gablasova, D. (2013). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics,* 1–23.

British Council. Languages for the future: Which languages the UK needs most and why (2014). Retrieved from http://www.britishcouncil.org/organisation/publications/ languages-future on 12 March 2015.

Brodine, R. (2001). Integrating corpus work into an academic reading course. In G. Aston, (Ed.), *Learning with corpora* (pp. 138-176). Houston, TX: Athelstan.

Brown, H. D. (2001). *Teaching by principles: an interactive approach to language pedagogy*. 2nd ed. White Plains, NY: Longman.

Browne, C. (2014). The new general service list version 1.01: Getting better all the time. *Korea TESOL Journal*, *11*(1), 35-50.

Bunting, J., Diniz, L., & Reppen, R. (2013). *Grammar and beyond 4*. Cambridge: Cambridge University.

Cain, K. (2002). Syntactic awareness and reading ability: Is there any evidence for a special relationship? *Applied Psycholinguistics, 29*, 679-94.

Carter, R., & McCarthy, M. (1997). *Exploring spoken English*. Cambridge: Cambridge University.

Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English: A comprehensive guide. spoken and written English grammar and usage.* Cambridge: Cambridge University.

Chang. J. (2014). The use of general and specialized corpora as reference sources for academic English writing: A case study. *ReCALL*, *26*(2), 243-259.

Charles, M. (2012) "Proper vocabulary and juicy collocations": EAP students evaluate do-it-yourself corpus-building. *English for Specific Purposes*, *31*(2), 93-102.

Cheng, W. (2012). *Exploring corpus linguistics: language in action*. New York: Routhledge.

Ciesielska-Ciupek, M. (2003). Teaching the İnternet and corpus material: Preparation of the ELT materials using the internet and corpus resources. In Lewandowska-Tomaszczyk, B. (Ed.), *PALC 2001: Practical applications in language corpora* (pp. 521-531). New York: Peter Lang.

Cobb, T. (1999). Breadth and depth of vocabulary acquisition with hands-on concordancing. *Computer Assisted Language Learning 12*, 345-360.

Cobb, T. (2003). Analyzing late interlanguage with learner corpora: Quebec replications of three European studies. *Canadian Modern Language Review/La Revue canadienne des languesvivantes, 59*(3), 393-424.

COCA (n.d.) *Corpus of contemporary American English*. Retrieved from http://corpus.byu.edu/coca on 10 January 2014.

Collins Cobuild English Grammar. (1990). Glasgow: HarperCollins.

Conrad, S. M. (2000). Will corpus linguistics revolutionize the grammar teaching in the 21st century? *TESOL Quarterly 34*(3), 548-560.

Conrad, S., & Biber, D. (2009). *Real grammar: A corpus-based approach to English*. White Plains, NY: Pearson Longman.

Conzett, J. (2000). Integrating collocation into a reading and writing course. In M. Lewis (Ed), *Teaching collocation: Further developments in the lexical approach* (pp. 70-87). Hove, England: Language Teaching.

Cookson, S., Hunter, S., Jackson, D., and Sick, J. (2005, June). *Using criterion in an English for academic purposes course*. The Proceedings of JALTCALL 2005, 147–151.

Council of Europe (2001). Common European framework of reference for languages: learning, teaching, assessment. Council of Europe / Cambridge University. Available online: www.coe.int/.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213-238.

Çakır, A. (1989). The Teaching of Turkish as a foreign language. Phd Dissertation. Gazi University Institute of Social Sciences, Ankara.

Gardner, D., & Davies, M. (2013) A new academic vocabulary list. *Applied Linguistics*, 35: 1-24.

Davis, B. & Russell-Pinson, L. (2004). Concordancing and corpora for k-12 teachers: project MORE. In Connor, Ula/Upton, Thomas (Eds) *Applied Corpus Linguistics: A Multidimensional Perspective*. pp. 147-169 .Amsterdam: Rodopi,

Dörnyei, Z. (2007). Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies. Oxford: Oxford University.

Eniko C. (2006). Academic talk in American university classrooms: crossing the boundaries of oral-literate discourse? *Journal of English for Academic Purposes 5*, 117–135.

Farr, F. (2003) 'Engaged listenership in spoken academic discourse: the case of student-tutor meetings', *Journal of English for Academic Purposes 2*(1): 67–85.

Flowerdew, J. (1993). Concordancing as a tool in course design. *System, 21*(2), 231-244.

Flowerdew. L. (2015). Corpus-based research and pedagogy in EAP: From lexis to genre. *Language Teaching*, *48*(1), 99-116.

Fox, G. (1998). Using corpus data in the classroom. In B. Tomlinson (Ed.), *Materials development in language teaching* (pp. 25-43). Cambridge, England: Cambridge University.

Francis*, W.,* & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.

Gairns, R., & Redman, S. (1986). *Working with words: a guide to teaching and learning vocabulary*. New York: Cambridge University.

Gardner, D. (2007). Validating the construct of word in applied corpus-based vocabulary research: A critical survey. *Applied Linguistics*, *28*(2), 241-265.

Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? System, *32*(3), 301-319.

Gavioli, L. (2005). *Exploring corpora for ESP learning*. Amsterdam: John Benjamins.

Genç, B. (2013). Differences in English vocabulary use: Insights from spoken learner corpus and native speaker corpus. *Eğitim ve Bilim*, *38*(167), 41-49.

Gilmore, A. (2009). Using online corpora to develop students' writing skills. *ELT Journal*, *63*(4), 363-372.

Göz, İ. (2003). *Yazılı Türkçenin Kelime Sıklığı Sözlüğü*. Ankara: Türk Dil Kurumu.

Granger, S. (2002). A bird's-eye view of learner corpus research. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.) *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 517-533). John Benjamins.

Grant, L. E. (2011). The frequency and functions of just in British academic spoken English *Journal of English for Academic Purposes*, *10*(3), 183-p197.

Harwood. N. (2005). What do we want EAP teaching materials for? *Journal of English for Academic Purposes, 4*(2), 149-161.

Hazenberg, S., & Hulstijn, H. (1996). Defining a minimal receptive second-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics*, *17*, 145-163.

Heatley, A., Nation, I.S.P., & Coxhead, A. (2002). *RANGE and FREQUENCY programs*. Retrieved from http://www.vuw.ac.nz/lals/staff/Paul_Nation on 11 March, 2014

Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language, 8*, 689–696.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, *19*(1), 24-44.

Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, *13*, 403–430.

Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary"? *TESOL Quarterly*, *41*(2), 235-253.

Johansson, S. (1995). 'Mens sana in corpora sano: On the role of corpora in linguistic research'. *The European English Messenger, 4*(2), 19-25

Karababa Z. C. (2009). Teaching Turkish as a foreign language and problems encountered. *Ankara University, Journal of Faculty of Educational Sciences. 42*(2), 265-277

Kaur, J., & Hegelheimer, V. (2005). ESL students' use of concordance in the transfer of academic word knowledge: An exploratory study. *Computer Assisted Language Learning*, *18*(4), 287-310.

Kennedy, G. (1998). *An introduction to corpus linguistics*. London/New York: Longman.

Kettemann, B. (1995). Concordancing in stylistics teaching. In W. Grosser, J. Hogg, & K. Hubmayer (Eds.), *Style: Literary and non-literary – contemporary trends in cultural stylistics* (pp. 307-318). Lewiston, NY: The Edwin Mellen.

Krishnamurthy, R. & Kosem, I. (2007). Issues in creating a corpus for EAP pedagogy and research. *Journal of English for Academic Purposes, 6*(4), 356-373.

Kurtoğlu, Ö. (2013). *A corpus based discourse analysis: the properties of texts in Turkish textbooks for primary and secondary schools*. Phd Dissertation. Mersin University Institute of Social Sciences, Mersin.

Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren and M. Nordman (Eds). *Special language: from humans thinking to thinking machines* (pp. 316-323). Bristol, UK: Multilingual Matters.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In H. Bejoint and P. Arnaud (Eds.) *Vocabulary and Applied Linguistics* (pp. 126-132). NY: Macmillan.

Laufer, B., & Ravenhorst-Kalovski, G. (2010). Lexical threshold revisited: lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22*, 15-30.

Lewis, M. (1993). *The lexical approach: The state of ELT and a way forward.* Hove, England: Language Teaching.

Lewis, M. (1997). *Implementing the lexical approach*. Hove, England: Language Teaching.

Lewis, M. (Ed.). (2000). *Teaching collocation: Further development in the lexical approach*. Hove, England: Language Teaching.

Lindeman, S., & Mauranen, A. (2001). "It's just real messy": The occurrence and function of just in a corpus of academic speech". *English for Specific Purposes*, *20(1),* 459-476

Louw, B. (1997). The role of corpora in critical literary. In A. Wichmann, S. Fligelstone, T. McEnrey & G. Knowles (Eds.), *Teaching and language corpora*. (pp. 240-252). NY: Addison Wesley Longman.

McCarthy, M., & O'Dell, F. (2005). *English collocations in use. Cambridge, England: Cambridge University.*

McCarthy, M. (2004). *From corpus to classroom*. Cambridge: Cambridge University.

McEnery, T., & Wilson, A. (2001). *Corpus linguistics* (2nd Ed.). Edinburgh: Edinburgh University

METU. (n.d.). *METU Turkish Corpus*. Retrieved from http://ii.metu.edu.tr/content/metu-turkish-corpus on 10 January 2014.

Meunier, F. (2002). The pedagogical value of native and learner corpora in EFL grammar teaching. In S, Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching*. (pp. 119-141). Amsterdam: John Benjamins.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis (2nd ed.).* Thousand Oaks, CA: Sage.

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

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University.

Nation, I. S. P. (2004). A study of the most frequent word families in the British National Corpus. In P. Bogaards and B. Laufer (Eds.) *Vocabulary in a second language* (pp. 3-14). Philadelphia, PA: John Benjamins.

Nation, I. S. P. (2006). *Second language vocabulary*. In K. Brown (Ed.), Encyclopedia of language and linguistics (2nd ed.), (pp. 448–454). Oxford: Elsevier.

Nation, I.S.P. & Beglar, D. (2007) A vocabulary size test. *The Language Teacher*, *31(7),* 9-13.

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

Nation, I.S.P., & Webb, S. (2010). *Researching and analyzing vocabulary*. Boston, MA: Heinle.

O'Dell, F., & McCarthy, M. (2008). *English collocations in use: Advanced*. Cambridge, England: Cambridge University.

O'Keeffe, A. M., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge, England: Cambridge University.

Osborne, J. (2003). Integrating corpora into a language-learning syllabus (plenary). In B. Lewandowska-Tomaszczyk (Ed.), *PALC 2001: Practical applications in language corpora* (pp. 479-492). New York: Peter Lang.

O'Sullivan, Í., & Chambers, A. (2006). Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of Second Language Writing*, *15*(1) 49-68.

Read, J. (2002). The use of interactive input in EAP listening assessment. *Journal of English for Academic Purposes*, 1, 105–119

Richard W. T. (2003). EAP or TEAP? *Journal of English for Academic Purposes*, *2*(2), 147–156.

Römer, U. (2011). Corpus research applications in second language teaching. *Annual review of applied linguistics, 31*, 205-225.

Rundell, M. (2007, October). 'The dictionary of the future" in Granger, S. (Ed.) *Optimising the role of language in technology-enhanced learning.* Proceedings of Expert Workshop of the Integrated Digital Language Learning seed grant project, Louvain-la-neuve, Belgium: 49-51.

Schmitt, N. (2010). *Researching vocabulary: a vocabulary research manual.* NY: Palgrave Macmillan.

Schmitt, N. Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Modern Language Journal, 95*(1), 26-43.

Scott, M., & Tribble, C. (2006). *Textual patterns: Keywords and corpus analysis in language education.* Amsterdam: John Benjamins.

Scrivener, J. (2011). *Learning teaching.* Oxford: MacMillan Education.

Selinker, L. (1972). Interlanguage. *IRAL-International Review of Applied Linguistics in Language Teaching*, *10*(1-4), 209-232.

Sezer, B., Sezer, T. (2013, May). TS Corpus: Herkes için Türkçe derlem. *Proceedings of the 27th National Linguistics Conference.* Antalya. Hacettepe University, Linguistics Department, 217-225. Available online at http://tscorpus.com.

Shavelson, R. J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. *Review of Educational Research, 51*(4), 455-498.

98

Shi, L., & Cumming, A. (1995). Teachers' conceptions of second language writing instruction: five case studies. *Journal of Second Language Writing, 4*(2), 87-111.

Sinclair, J. (ed.-in-chief-) (1987). *Collins cobuild English language dictionary*. London: HarperCollins.

Sinclair, J. M. (ed.) (2004), *How to use corpora in language teaching*. Amsterdam: John Benjamins.

Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics*. Oxford: Blackwell.

Sun, Y. C. (2000). *Using on-line corpus to facilitate language learning*. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages, British Columbia, Canada.

Sun, Y. C. (2003). Extensive reading online: An overview and evaluation. *Journal of Computer Assisted Learning*, *19*(4), 438-446.

Sun, Y. C. (2007). Learner perceptions of a concordancing tool for academic writing. *Computer Assisted Language Learning*, *20*(4), 323-343.

Sun, Y. C., & Wang, L.-Y. (2003). Concordancers in the EFL classroom: Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning*, *16*(1), 83-94.

Svartvik, J. (Ed.). (1990). *The London-Lund corpus of spoken English: Description and research*. Lund, Sweden: Lund University.

Şanal, F. (2007). *A learner corpus based study on second language lexicology of Turkish students of English*. Phd Dissertation, Çukurova University Institute of Social Sciences, Adana.

Tang, W. M. (2015) *A corpus linguistic glossary*. Retrieved from http://wmtang.org/corpus-linguistic-glossary/ on 15March 2015

The British National Corpus, version 3 (BNC XML Edition). (2007). From Oxford University Computing Services on behalf of the BNC Consortium. Retrieved from http://www.natcorp.ox.ac.uk/ on 22 February 2014

Tribble, C. (2001). Small corpora and teaching writing. In M. Ghadessy, A. Henry, & R. L. Roseberry (Eds.). *Small corpus studies and ELT: Theory and practice*. (pp. 381-408) Amsterdam: John Benjamins.

Tribble, C., & Wingate, U. (2013). From text to corpus - A genre-based approach to academic literacy instruction. *SYSTEM*, *41*(2), 307-321.

Tribble, C., & Johns, G. (1997). *Concordances in the classroom: A resource book for teachers*. Houston, TX: Athelstan.

TS Corpus. (2012). *The Turkish Corpus*. Retrieved from http://tscorpus.com/frontend /files/userfiles/files/TS_Corpus_Mini_Paper_English.pdf on 15 December 2014.

Tsui, A. B. (2004). What teachers have always wanted to know–and how corpora can help. In J. H. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 39-66). John Benjamins.

Vincent, B. (2013). Investigating academic phraseology through combinations of very frequent words: A methodological exploration. *Journal of English for Academic Purposes*, *12*(1), 44-56.

Ward, J. W. (1999). How large a vocabulary do EAP engineering students need? *Reading in a Foreign Language, 12*(2), 309–324.

Webb, S., & Rodgers, M.P.H. (2009a). The lexical coverage of movies. *Applied Linguistics*, *30*, 407-427.

Webb, S., & Rodgers, M.P.H. (2009b). The vocabulary demands of television programs. *Language Learning*, *59*, 335-366.

West, M. (1953). *A General Service List of English Words*. London: Longman, Green and Co.

Wichmann, A., Fligelstone, S., McEnery, T., & Knowles, G. (Eds.) (1997). *Teaching and language corpora*. New York: Longman.

Wilkins, D. A. (1972*). Linguistics in language teaching*. London: Edward Arnold.

Willis, D. (1990). *The lexical syllabus: A new approach to language teaching*. UK: Collins Cobuild.

Willis, J. (1998). Concordances in the classroom without a computer: Assembling and exploiting concordances of common words. In Tomlinson, B. (Ed.). *Materials development in language teaching* (pp. 44-66). Cambridge England: Cambridge University.

Wood, D., & Appel, R. (2014). Multiword constructions in first year business and engineering university textbooks and EAP textbooks. *Journal of English for Academic Purposes*, 15, 1-13.

Woolard, G. (2000). Collocation-encouraging learner independence. In M. Lewis (Ed.), *Teaching collocation: Further development in the lexical approach* (pp. 28-46). Hove, England: Language Teaching.

Yıldırım, A. & Şimşek, H. (2006). *Sosyal bilimlerde nitel araştırma yöntemleri*. Ankara: Seçkin.

Yıldız, İ. (2012). *A corpus-based analysis of predicates used in the introduction and conclusion sections of turkish academic texts*. MA Thesis. Mersin University Institute of Social Sciences, Mersin.

APPENDICES

APPENDICES

Appendix 1 The list of words which was excluded from the General Turkish Wordlist

#	word	#	word	#	word	#	word
1	bir	41	ta	81	33	121	70
2	bire	42	tir	82	34	122	700
3	birer	43	veveya	83	3-4	123	72
4	birinci	44	ye	84	35	124	75
5	birincil	45	1993	85	36	125	8
6	birincisi	46	1994	86	37	126	80
7	çoluk	47	1995	87	38	127	800
8	da	48	1996	88	39	128	85
9	dan	49	1997	89	4	129	9
10	de	50	1998	90	40	130	90
11	den	51	1999	91	400	131	95
12	di	52	2	92	41	132	aaa
13	ikimiz	53	20	93	42	133	ab
14	ikincil	54	200	94	43	134	abbas
15	imiş	55	2000	95	44	135	abd
16	la	56	2001	96	45	136	abd'de
17	le	57	2002	97	4-5	137	abd'deki
18	mı	58	2002'de	98	46	138	abd'li
19	mıdır	59	2003	99	47	139	abd'nin
20	mısın	60	2004	100	48	140	abdullah
21	mısınız	61	2005	101	49	141	abd'ye
22	mıydı	62	2006	102	5	142	ab'nin
23	mıyım	63	2007	103	50	143	ab'ye
24	mıyız	64	2008	104	500	144	adana
25	mi	65	21	105	51	145	adem
26	midir	66	22	106	52	146	adnan
27	misin	67	23	107	53	147	afganistan
28	misiniz	68	2-3	108	54	148	afrika
29	miydi	69	24	109	55	149	afrika'da
30	miyim	70	25	110	56	150	ağar
31	miyiz	71	250	111	5-6	151	ahmed
32	mu	72	26	112	57	152	ahmet
33	mudur	73	27	113	58	153	akdeniz
34	musun	74	28	114	6	154	akif
35	musunuz	75	29	115	60	155	akp
36	muydu	76	3	116	600	156	akp'nin
37	muyum	77	30	117	64	157	ali
38	muyuz	78	300	118	65	158	ali'nin
39	mü	79	31	119	68	159	alman
40	müdür	80	32	120	7	160	almanca

#		#		#		#	
161	almanlar	202	asım	243	bine	284	çankaya
162	almanya	203	asya	244	bini	285	çiğdem
163	almanya'da	204	asya'da	245	binlerce	286	çiller'in
164	almanya'dan	205	atatürk	246	bir	287	çin
165	almanya'nın	206	atatürk'ün	247	bir-iki	288	çin'de
166	almanya'ya	207	atilla	248	bizans	289	çin'in
167	altan	208	attila	249	bm	290	çoğunlukla
168	altı	209	avrupa	250	boğaziçi	291	d
169	altıncı	210	avrupa'da	251	bozkurt	292	david
170	altmış	211	avrupa'daki	252	brezilya	293	demirel
171	amerika	212	avrupa'dan	253	britanya	294	demirel'in
172	amerika'da	213	avrupalı	254	bt	295	denizli
173	amerika'daki	214	avrupa'nın	255	bulgar	296	denktaş
174	amerika'dan	215	avrupa'ya	256	bulgaristan	297	dicle
175	amerikalı	216	avusturya	257	burak	298	diyarbakır
176	amerikalılar	217	ayhan	258	burcu	299	dna
177	amerikan	218	aylin	259	burhan	300	doç
178	amerika'nın	219	ayşe	260	bursa	301	dokuz
179	amerika'ya	220	azerbaycan	261	bursa'da	302	dördüncü
180	anadolu	221	b	262	bush	303	dört
181	anadolu'da	222	bağdat	263	bush'un	304	dörtte
182	anadolu'daki	223	baka	264	bülent	305	dp
183	anadolu'nun	224	balkan	265	c	306	dr
184	anadolu'ya	225	bank	266	cahil	307	dsp
185	anap	226	barzani	267	cahit	308	dyp
186	anap'ın	227	basa	268	canan	309	e
187	ankara	228	bayar	269	cd	310	ebru
188	ankara'da	229	baykal	270	cem	311	ebu
189	ankara'daki	230	baykal'ın	271	cemal	312	ecevit
190	ankara'dan	231	behçet	272	cemil	313	ecevit'in
191	ankara'nın	232	bekir	273	cenab-ı	314	eee
192	ankara'ya	233	belçika	274	cengiz	315	efes
193	annan	234	berlin	275	cevat	316	ege
194	antalya	235	beş	276	cevdet	317	ekrem
195	apo	236	beşiktaş	277	chp	318	elif
196	arabistan	237	beşiktaş'ın	278	chp'nin	319	elli
197	arap	238	beşinci	279	cıa	320	emel
198	arapça	239	beyoğlu	280	clinton	321	emine
199	ar-ge	240	bin	281	cm	322	enver
200	arif	241	binbir	282	cüneyt	323	e-posta
201	arnavut	242	binde	283	çanakkale	324	erbakan

325	ercan	366	gsm	407	ikinci	448	italya
326	erdal	367	gül'ün	408	ikincisi	449	italya'da
327	erdoğan	368	h	409	ikisi	450	italyan
328	erdoğan'ın	369	hakan	410	ikisini	451	izmir
329	ergenekon	370	halil	411	ikisinin	452	izmir'de
330	erkan	371	halit	412	ikişer	453	izmir'e
331	ermeni	372	haluk	413	ikiye	454	izzet
332	ermeniler	373	hamdi	414	ilhan	455	j
333	erol	374	hamit	415	in	456	james
334	ertuğrul	375	handan	416	ingiliz	457	japon
335	erzurum	376	hasan	417	ingilizce	458	japonya
336	eskişehir	377	hatice	418	ingilizler	459	japonya'da
337	esra	378	haydar	419	ingiltere	460	jazz
338	eşref	379	hayri	420	ingiltere'de	461	jean
339	euro	380	hilmi	421	ingiltere'nin	462	john
340	f	381	hindistan	422	ingiltere'ye	463	k
341	faik	382	hindistan'da	423	inönü	464	kadıköy
342	faruk	383	hint	424	iran	465	kadir
343	fatih	384	hollanda	425	iran'a	466	kamil
344	fatma	385	hülya	426	iran'da	467	kanada
345	fazıl	386	hüseyin	427	iran'ın	468	kanuni
346	fenerbahçe	387	hüseyin'in	428	is	469	karadeniz
347	ferit	388	hüsnü	429	isa	470	karagöz
348	fethullah	389	hz	430	isa'nın	471	katrilyon
349	fevzi	390	ı	431	ismail	472	kayseri
350	fırat	391	ıı	432	ismet	473	kazım
351	fikret	392	ııı	433	ispanya	474	kdv
352	filistin	393	ımf	434	ispanyol	475	kemal
353	filistinli	394	ınternet	435	israil	476	kemal'in
354	fransa	395	ırak	436	israil'e	477	kemalist
355	fransa'da	396	ırak'a	437	israil'in	478	kenan
356	fransa'nın	397	ırak'ın	438	istanbul	479	kerem
357	fransız	398	ıraklı	439	istanbul'a	480	kerim
358	fransızca	399	ırak'ta	440	istanbul'da	481	kg
359	fuat	400	ırak'taki	441	istanbul'daki	482	kıbrıs
360	g	401	ıv	442	istanbul'dan	483	kıbrıs'ın
361	galatasaray	402	i	443	istanbullu	484	kıbrıslı
362	george	403	ibn	444	istanbul'u	485	kıbrıs'ta
363	global	404	ibrahim	445	istanbul'un	486	kırk
364	gökhan	405	iki	446	isveç	487	kktc
365	gr	406	ikide	447	isviçre	488	km

489	konya	530	milyonlarca	571	okan	612	reaksiyon
490	kore	531	milyonu	572	oktay	613	recep
491	korkut	532	mit	573	on	614	reel
492	kosova	533	mithat	574	onlarca	615	refik
493	kürdistan	534	mkemal	575	online	616	reşit
494	kürt	535	ml	576	oral	617	rıza
495	kürtçe	536	mm	577	orhan	618	robert
496	kürtler	537	moskova	578	ortaçağ	619	rock
497	kürtlerin	538	mö	579	ortadoğu	620	roma
498	l	539	muhammed	580	ortadoğu'da	621	rp
499	latin	540	muhammed'i n	581	ortodoks	622	rp'nin
500	levent	541	muhsin	582	osman	623	rum
501	leyla	542	mumcu	583	osmanlı	624	rumlar
502	londra	543	murat	584	osmanlılar	625	rus
503	londra'da	544	musa	585	osmanlı'nın	626	rusya
504	m	545	mustafa	586	otuz	627	rusya'da
505	mahmud	546	n	587	öcalan	628	rusya'nın
506	mahmut	547	naciye	588	ömer	629	rusya'ya
507	manço	548	namık	589	özal	630	s
508	mark	549	nasreddin	590	özal'ın	631	sabahattin
509	marmara	550	nato	591	özcan	632	sabancı
510	marx	551	nazım	592	özdemir	633	sabri
511	mehmed	552	necati	593	özer	634	saddam
512	mehmet	553	necdet	594	özkan	635	sait
513	meksika	554	necip	595	p	636	sakarya
514	melih	555	necmettin	596	pakistan	637	salih
515	menderes	556	nedim	597	paris	638	sallallahu
516	meral	557	nermin	598	paris'e	639	sami
517	mersin	558	nevzat	599	paris'te	640	samsun
518	meryem	559	new	600	paul	641	sedat
519	mevlana	560	nihat	601	peter	642	sekiz
520	mg	561	nilüfer	602	ph	643	selahattin
521	mhp	562	no	603	pişkinsüt	644	selçuk
522	mhp'nin	563	nobel	604	pkk	645	selçuklu
523	mısır'da	564	nokia	605	pkk'nın	646	selim
524	michael	565	nolu	606	pkk'ya	647	selma
525	microsoft	566	nurettin	607	polonya	648	sema
526	milyar	567	nuri	608	primer	649	semra
527	milyarlarca	568	oecd	609	prof	650	serdar
528	milyon	569	oğuz	610	r	651	serkan
529	milyona	570	ok	611	rauf	652	sevim

653	seyit	691	tl	729	v	767	zeynep
654	sezen	692	to	730	van	768	ziya
655	sezer	693	tokat	731	vb	769	a
656	shp	694	tony	732	vedat	770	aa
657	sıfır	695	trabzon	733	veysel	771	aferin
658	sırp	696	trabzonspor	734	viyana	772	ah
659	sinan	697	trilyon	735	vs	773	alo
660	sivas	698	trt	736	w	774	aman
661	sovyet	699	tuğrul	737	washington	775	bari
662	sovyetler	700	tuna	738	web	776	be
663	ssk	701	turgut	739	windows	777	eh
664	star	702	turkcell	740	x	778	ey
665	suat	703	türk	741	y	779	eyvah
666	suna	704	türkan	742	yahya	780	ha
667	suriye	705	türkçe	743	yakup	781	haa
668	susurluk	706	türkiye	744	yalçın	782	hadi
669	suudi	707	türkiye'de	745	yasemin	783	hah
670	suzan	708	türkiye'deki	746	yavuz	784	hani
671	süleyman	709	türkiye'den	747	yedi	785	hay
672	şemsi	710	türkiye'nin	748	yedinci	786	he
673	şener	711	türkiye'ye	749	yetmiş	787	hey
674	şevket	712	türkiye'yi	750	yılmaz	788	keşke
675	şükran	713	türkler	751	yılmaz'ın	789	ki
676	şükrü	714	türklere	752	yirmi	790	merhaba
677	t	715	türkleri	753	yirminci	791	of
678	tahir	716	türklerin	754	york	792	oh
679	tak	717	türkmen	755	york'ta	793	ola
680	talat	718	türk'ün	756	yök	794	sakın
681	tansu	719	tv	757	ytl	795	ulan
682	tanzimat	720	u	758	yunan	796	vah
683	tarık	721	uefa	759	yunanistan	797	valla
684	tayyip	722	ü	760	yunanistan'a	798	vallahi
685	tbmm	723	üç	761	yunanistan'ın	799	vay
686	tc	724	üçlü	762	yunus	800	yaa
687	tefek	725	üçte	763	yusuf	801	yahu
688	terakki	726	üçü	764	yücel	802	yoo
689	tevfik	727	üçüncü	765	yüzde		
690	the	728	üçüncüsü	766	yüzlerce		

Appendix 2 The Academic Turkish Corpus Composition

	Discipline				
	Humanities	**Education**	**Economy (Commerce)**	**Law**	**Science**
Running Words	937999	778,799	726038	764482	825124
Textbooks	11	10	10	10	16
Subject Areas	Antropoloji	Anne Baba Eğitimi	Bankacılık ve Sigortacılığa Giriş	Anayasa Hukuku	Aile Sağlığı
	Epistemoloji	Eğitim Bilimine Giriş	Çalışma Ekonomisi	Borçlar Hukuku	Elektromekanik Kumanda Sistemleri
	Etik	Eğitim Psikolojisi	Finansal Ekonomi	Ceza Hukuku	Enerji Ve Çevre
	Felsefe	Eğitim Sosyolojisi	Genel İşletme	Hukuk Dili ve Adli Yazışmalar	Ev teknolojisi
	Güzel Sanatlar	Okul Öncesi Eğitime Giriş	Genel Muhasebe	Hukukun Temel Kavramları	Genel Biyoloji
	Halkla İlişkiler	Öğretim İlke Ve Yöntemleri	İktisada Giriş	İnsan Hakları Ve Kamu Özgürlükleri	Genel Kimya
	İnsan ve Toplum	Öğretmenlikte Mesleki Gelişim	İşletme Yönetimi	Medeni Hukuk - 1	Genel Matematik
	Psikoloji	Özel Eğitim	Para Politikası	Medeni Hukuk - 2	Güç Sistemleri Analizi
	Psikolojiye Giriş	Rehberlik	Türkiye Ekonomisi	Ticaret Hukuku	Güvenlik Sistemleri
	Sosyoloji 1	Sınıf Yönetimi	Uluslararası İktisat Politikası	Türk Anayasa Hukuku	Sayısal Fotoğrafa Giriş
	Yakınçağ Avrupa Tarihi				Sosyal Medya
					Teknolojinin Bilimsel İlkeleri - 1
					Temel Bilgi Teknolojileri
					Temel Klinik Bilgisi
					Yeni İletişim Teknolojileri
					Yenilenebilir Enerji Kaynakları

Appendix 3 Details of the Academic Turkish Corpus Composition

The Academic Turkish Corpus

	TOKENS/%	TYPES/%
Total	4032442/100.00	182285/100.00

Table 1 All raw data corpus

	TOKENS/%	TYPES/%
Total	3415378/100.00	181980/100.00

Table 2 Excluded words corpus

1) Economy
   Number of words: 726038

   | Types: 50,325 | Tokens: 726,038 |

2) Education
   Number of words: 778799

   | Types: 57,529 | Tokens: 778,799 |

3) Humanities
   Number of words: 937999

   | Types: 79,281 | Tokens: 937,962 |

4) Law
   Number of words: 764482

    | Types: 48,687 | Tokens: 764,480 |

5) Science
   Number of words: 825124

   | Types: 77,274 | Tokens: 825,092 |

Appendix 4 Function List (The Second Exclusion List)

1	Acaba	40	Birileri	79	bunun	118	gereği	157	hiçbir
2	Adeta	41	birilerinin	80	bununla	119	gereğince	158	hiçbiri
3	Aksi	42	birinde	81	burada	120	gereğinden	159	hiçbirini
4	Aksine	43	birinden	82	buradaki	121	gereğini	160	için
5	Ama	44	birine	83	buradan	122	gerek	161	içindir
6	Amma	45	birini	84	buralara	123	gerekse	162	idi
7	Ancak	46	birinin	85	buralarda	124	gibi	163	iken
8	Ayrıca	47	birisi	86	buranın	125	gibidir	164	ikili
9	Azar	48	birisidir	87	burası	126	gibiydi	165	ila
10	Bana	49	birisini	88	buraya	127	gibiyim	166	ile
11	Ben	50	birisinin	89	burayı	128	göre	167	illa
12	Bence	51	biriydi	90	burda	129	hakkında	168	ille
13	Bende	52	biriyle	91	buydu	130	hakkındaki	169	ise
14	Benden	53	biz	92	buyrun	131	hangi	170	işte
15	Beni	54	bizde	93	buyur	132	hangisi	171	kaç
16	Benim	55	bizden	94	buyurun	133	hatta	172	kaçıncı
17	Benimki	56	bize	95	çünkü	134	haydi	173	kadar
18	Benimle	57	bizi	96	dahi	135	hayır	174	kadardır
19	Bense	58	bizim	97	değil	136	hele	175	kadarı
20	Beri	59	bizimki	98	değildi	137	hem	176	kadarını
21	Bile	60	bizimkiler	99	değildim	138	hepimiz	177	kadarıyla
22	Bir arada	61	bizimle	100	değildir	139	hepimizi	178	kah
23	Bir araya	62	bizler	101	değilim	140	hepimizin	179	kanımca
24	Birbiri	63	bizlere	102	değiliz	141	hepiniz	180	karşın
25	birbirimize	64	bizleri	103	değiller	142	hepsi	181	kaydıyla
26	birbirimizi	65	bu	104	değilmiş	143	hepsinde	182	kendi
27	birbirinden	66	bucaksız	105	değilse	144	hepsinden	183	kendileri
28	Birbirine	67	budur	106	değilsin	145	hepsine	184	kendilerinden
29	Birbirini	68	buna	107	değin	146	hepsini	185	kendilerine
30	birbirinin	69	bunda	108	dek	147	hepsinin	186	kendilerini
31	Birbiriyle	70	bundan	109	denli	148	her	187	kendilerinin
32	Birbirleri	71	bunlar	110	dolayı	149	herkes	188	kendiliğinden
33	birbirlerinden	72	bunlara	111	eğer	150	herkese	189	kendim
34	birbirlerine	73	bunlarda	112	en	151	herkesi	190	kendimden
35	birbirlerini	74	bunlardan	113	evet	152	herkesin	191	kendime
36	birbirlerinin	75	bunları	114	fakat	153	herkesten	192	kendimi
37	birbirleriyle	76	bunların	115	falan	154	herşey	193	kendimiz
38	Biri	77	bunlarla	116	filan	155	herşeyi	194	kendimize
39	Biridir	78	bunu	117	gerçi	156	hiç	195	kendimizi

#	Word	#	Word	#	Word	#	Word	#	Word
196	kendin	236	Mesela	276	onlarla	316	öyleyse	356	şu
197	kendince	237	Nasıl	277	onu	317	pahasına	357	şudur
198	kendinde	238	Nasılsa	278	o'nu	318	pekala	358	şuna
199	kendinden	239	Nasılsın	279	onun	319	peki	359	şunlar
200	kendine	240	nazaran	280	o'nun	320	rağmen	360	şunlardır
201	Kendini	241	ne	281	onunla	321	sana	361	şunları
202	kendinize	242	neden	282	orada	322	sen	362	şunu
203	kendinizi	243	nedir	283	oradaki	323	sende	363	şurada
204	Kendisi	244	neler	284	oradan	324	senden	364	üzere
205	kendisinde	245	nelerdir	285	oralarda	325	seni	365	üzereydi
206	kendisinden	246	neleri	286	orası	326	senin	366	ve
207	kendisine	247	nerde	287	oraya	327	seninle	367	veya
208	kendisini	248	nerden	288	orda	328	sensin	368	ya
209	kendisinin	249	nerede	289	ordan	329	seve	369	yada
210	kendisiyle	250	nereden	290	orta	330	siz	370	yahut
211	Kim	251	neresi	291	ortada	331	sizce	371	yalnızca
212	Kimbilir	252	nereye	292	ortadan	332	sizde	372	yan
213	Kimden	253	nesi	293	oysa	333	sizden	373	yana
214	Kimdi	254	neydi	294	önce	334	size	374	yanda
215	Kimdir	255	neye	295	önceden	335	sizi	375	yandan
216	Kime	256	neyi	296	önceki	336	sizin	376	yanı
217	Kimi	257	neyin	297	önceleri	337	sizinle	377	yanıma
218	Kimileri	258	neymiş	298	öncesi	338	sizler	378	yanımda
219	Kimin	259	neyse	299	öncesinde	339	sizlere	379	yanımıza
220	Kimisi	260	nezdinde	300	öncesinden	340	sizleri	380	yanımızda
221	Kimler	261	niçin	301	öncesine	341	şayet	381	yanına
222	Kimlerin	262	niye	302	öte	342	şey	382	yanında
223	Kimse	263	o	303	ötede	343	şeyden	383	yanındaki
224	kimseden	264	odur	304	öteden	344	şeydi	384	yanından
225	Kimseler	265	olsa	305	öteki	345	şeydir	385	yanını
226	kimselere	266	ona	306	ötekiler	346	şeye	386	yanısıra
227	kimselerin	267	o'na	307	ötekine	347	şeyi	387	yani
228	kimsenin	268	onda	308	ötesi	348	şeyin	388	yanları
229	Kimseye	269	ondan	309	ötesinde	349	şeyle	389	yanlarına
230	Kimseyi	270	onlar	310	ötesine	350	şeyler	390	yanlarında
231	kimseyle	271	onlara	311	öteye	351	şeylerden	391	yoksa
232	Lakin	272	onlarda	312	ötürü	352	şeylere	392	zira
233	Layık	273	onlardan	313	öyle	353	şeyleri	393	ziyade
234	Madem	274	onları	314	öylece	354	şeylerin		
235	Meğer	275	onların	315	öylesine	355	şeylerle		

Appendix 5 Academic Turkish Vocabulary Test

AKADEMİK TÜRKÇE KELİME TESTİ

Adınız (Optional):	**Ne kadar süredir Türkçe öğrenmektesiniz?**
Yaşınız:　　　　**Cinsiyetiniz:**	
Uyruğunuz:	**Ne kadar süredir Türkiye'de yaşamaktasınız?:**
Bölümünüz:	
Ana Diliniz:	**Sizce Türkçe bilginiz hangi seviyede?**
Bildiğiniz Yabancı Diller (Hangi düzeyde):	**a)** Konuşma:　　　Başlangıç – Orta – İleri Düzey
…..　　　Başlangıç – Orta – İleri Düzey	**b)** Dinleme:　　　Başlangıç – Orta – İleri Düzey
…..　　　Başlangıç – Orta – İleri Düzey	**c)** Okuma:　　　Başlangıç – Orta – İleri Düzey
…..　　　Başlangıç – Orta – İleri Düzey	**d)** Yazma:　　　Başlangıç – Orta – İleri Düzey

Aşağıdaki cümlelerde boş bırakılan yerlere gelebilecek uygun seçeneği işaretleyiniz.

1) Değerler işletmelerin iç ve dış çevreye karşı _____ konusundaki kararlarını etkiler.
A) ürünleri　　B) tutarları　　C) öğretimleri
D) zorlukları　　**E) sorumlulukları**

8) Güvenlik: Toplum yaşamında yasal düzenin aksamadan _____ , kişilerin korkusuzca yaşayabilmesi durumudur.
A) tanımlanması　　B) yayımlanması　　C) kurtulması
D) yürütülmesi　　　　E) tüketilmesi

2) Yapılan son _____ her dört kişiden üçünün kendilerini, ailelerini, evlerini ve iş yerlerini hırsızlığa karşı koruyabilmek için daha fazla şey yapabileceklerini düşündüklerini göstermektedir.
A) yerleşimler　　B) maliyetler　　**C) araştırmalar**
D) işaretler　　　　E) ortamlar

9) Korunma Kanunu, savaş koşullarında ekonominin üzerindeki baskıları _____ altına alabilmek, özellikle gıda zorluklarını gidermek üzere çıkartılmıştır.
A) denetim　　B) ulaşım　　　　C) etkileşim
D) girişim　　　　E) değişim

3) Yaşamak için _____ karşılamaya, bunun için ise çalışmaya mecburuz.
A) uyumlarımızı　　**B) ihtiyaçlarımızı**　　C) tehditlerimizi
D) etkileşimlerimizi　　E) takdirlerimizi

10) Avcılık yapan topluluklar yüksek protein _____ ile sağlıklı bir hayat sürerler.
A) tüketimi　　B) topluluğu　　C) yerleşimi
D) ölçümü　　　　E) konumu

4) Veteriner hekim hasta sahibinin üzüntü ve duygusallığını _____ karşılamak, hasta hayvana gerekli özeni göstermek ve onu tedavi etmek zorundadır.
A) anlayışla　　B) toplulukla　　C) tehditle
D) katılımla　　　　E) tepkiyle

11) Anlatım yöntemi, öğretmen merkezli olan ve okullarda sıklıkla kullanılan _____ bir yöntemdir.
A) çevresel　　**B) geleneksel**　　C) hariç
D) yükümlü　　　　E) yaklaşık

5) Firmanın şantiye dışına çıkardığı işçiler _____ yapmaya başladı.
A) yetki　　B) yaklaşım　　C) konum
D) kavram　　　**E) eylem**

12) Teknolojinin sunduğu _____ sayesinde kol gücü ile yapılan birçok işlem, artık insan müdahalesi olmaksızın otomatik olarak yapılabilir hale gelmiştir.
A) bünyeler　　B) gerekçeler　　C) onaylar
D) imkanlar　　　　E) tutarlar

6) Kendi ilgi, yetenek ve _____ özelliklerine uygun bir meslek sahibi birey, hem kendi ruh sağlığını koruyabilir, hem de daha kaliteli ve çok mal/hizmet üretebilir.
A) ürün　　B) takdir　　C) yarar
D) yerel　　　**E) kişilik**

13) 5 Mayıs 1926'da İngiltere ile Türkiye arasında yapılan _____ ile Irak sınırına son şekli verilmiştir.
A) işaret　　B) malzeme　　C) tehlike
D) teşkil　　　　**E) anlaşma**

7) Çocuğun çalışması ve "çocuk emeği" sorunu bütün ülkelerde yaşanmakta olan evrensel bir _____ hâline gelmiştir
A) takip　　**B) olgu**　　C) ortam
D) fırsat　　　E) imkan

14) Aileler çocuklarının eğitimi için yeni roller _____ durumunda kalmışlardır.
A) toplamak　　B) kurtarmak　　**C) üstlenmek**
D) saklamak　　　E) tüketmek

113

15) Bu noktada dikkatin önemini _____ fayda vardır. Bazen bir şeye baktığımız halde onu göremeyebiliriz. Zihnimizden bambaşka şeyler geçiyordur…
A) vurgulamamızda B) yüklememizde C) sarmamızda
D) yaygınlaşmamızda   E) ulaştırmamızda

16) Almanya ise 1733 uçağını ve bir o kadar da pilotunu kaybetmiştir. Bu başarısızlık karşısında Hitler kış mevsiminin gelmesini de _____ göstererek İngiltere ile savaşa son vermiştir.
A) kayıp   B) emek   **C) gerekçe**
D) tepki   E) takdir

17) Almanya, İtalya ve Japonya'nın bu amaca yönelik çalışmalarda _____ yapmaları benimsenmiştir.
A) tehlike   B) yaklaşım   C) kimlik
D) işbirliği   E) olanak

18) Artık günlük yaşantımızın _____ bir parçası olan bilgisayarların işlem yapma hızının gittikçe artması zor hesapların kısa zamanda yapılmasını sağlar.
A) küresel   B) toplu   C) aşırı
D) karşılıklı   **E) vazgeçilmez**

19) Çok sert bir kışın ardından gelen ve yağış getiren ilkbaharın ardından sıcak bir yaz yaşanması Kuzey ikliminin bir _____ .
A) göstergesidir   B) taklididir   C) bünyesidir
D) işbirliğidir   E) tesisidir

20) Etkileşim, karşılıklı hareket etme, birbirini etkileme anlamında kullanılan bir _____ .
A) emirdir   B) kişiliktir   C) takımdır
D) yayındır   **E) terimdir**

21) Sokak hayvanlarıyla değil, hastalıklarla _____ etmek gerekiyor.
A) tahmin   **B) mücadele**   C) takdir
D) teşvik   E) tehdit

22) Hem daha hızlı, hem daha çevreci araçları _____ aslında Mercedes'in mühendislerinin hayal güçleridir.
A) kapsayan   **B) tasarlayan**   C) ilgilendiren
D) bağlayan   E) çatışan

23) Sınıftaki _____ düzeni öğrencinin öğrenmesini etkilemektedir.
A) yerleşim   B) kitle   C) gözlem
D) mekanizma   E) suret

24) Japonya'nın deprem ve nükleer kazadan büyük ölçüde etkilenen turizmini tekrar canlandırmak için 10000 adet uçak biletini bedavaya vereceğini açıklaması turizmdeki sıkıntıyı _____ .
A) kanıtlamaktadır B) kaydetmektedir C) gözlemektedir
D) korumaktadır   E) yıldırmaktadır

25) Eğitim ortamının _____ iyi olması, çocuğun duygusal ve sosyal gelişimini desteklemektedir.
A) vadesinin   B) işaretinin   **C) ikliminin**
D) şartının   E) tehlikesinin

26) Türkiye Cumhuriyeti kurulduğunda nüfusun büyük bir kısmı eğitimden _____ idi.
A) kesin   B) yetkili   C) aşırı
D) yoksun   E) ahlaki

27) Hedefleri net olmayan bir kurumun çalışmalarının neye yönelik olacağı konusunda bir _____ yaşanır.
A) bilinç   **B) belirsizlik**   C) işleyiş
D) akış   E) nitelik

28) Çalışanlar işe alınırken _____ alanlarına göre çeşitli birimlerde görevlendirilmektedir.
A) uzmanlık   B) etken   C) terim
D) kitle   E) yaygın

29) Öğrendiklerimiz, _____ olarak anladıklarımızdan daha çoktur.
A) başlıca   **B) bilinçli**   C) eksik
D) yükümlü   E) etken

30) Anne babalar çocukların _____ ve duygusal problemleri ile başa çıkmakta güçlükler yaşayabilmektedirler.
A) uyumlu   B) yükümlü   C) değerli
D) etken   **E) davranışsal**

Appendix 6 Academic Turkish Interview Form (English Version)

ACADEMIC TURKISH INTERVIEW FORM

Dear participant,

This form aims to discover opinions of foreign students studying at universities offering Turkish medium instruction on their academic Turkish vocabulary knowledge

Name: Age: Gender: Nationality: Department: Which year: First Language: Foreign Languages (and levels) ….. Beginner – Preintermediate - Advanced ….. Beginner – Preintermediate - Advanced ….. Beginner – Preintermediate - Advanced	How long have you been living in Turkey?: When and How did you learn Turkish?: How well do you use Turkish? a) Speaking: Beginner – Preintermediate - Advanced b) Listening: Beginner – Preintermediate - Advanced c) Reading: Beginner – Preintermediate - Advanced d) Writing: Beginner – Preintermediate - Advanced

Please, try to give more detailed answers instead of saying only "yes" or "no" for the following questions:

1) Do you understand what you read in Turkish? (short stories, novels, newspapers)

2) Do you understand your textbooks? How well? Do you understand all of the words in the textbooks?

3) Do you understand your lecturers?

4) Do you communicate in class using Turkish? (with friends and lecturers about the lesson)

5) Do you have any difficulties in preparing projects or completing the exam paper (mid-term or final exam paper) in Turkish? (in terms of language)

6) Do you communicate in Turkish with your friends in social contexts? (in a café, restaurant etc.)

7) What kind of difficulties do you encounter as a student studying within the scope of Turkish medium instruction?

Appendix 7 Academic Turkish Interview Form (Turkish Version)

AKADEMİK TÜRKÇE GÖRÜŞME FORMU

Sayın katılımcı,

Bu form Türkçe öğretim yapılan üniversitelerdeki yabancı uyruklu öğrencilerin akademik Türkçe kelime bilgileri hakkındaki düşüncelerini ortaya çıkarmayı amaçlamaktadır.

Adınız:	Ne kadar süredir Türkiye'de yaşamaktasınız?
Yaşınız: Cinsiyetiniz:	
Uyruğunuz:	Ne zaman ve Nasıl Türkçe öğrendiniz?
Bölümünüz:	
Kaçıncı sınıftasınız:	
Ana Diliniz:	Sizce Türkçe bilginiz hangi seviyede?
Bildiğiniz Yabancı Diller (Hangi düzeyde):	a) Konuşma: Başlangıç – Orta – İleri Düzey
….. Başlangıç – Orta – İleri Düzey	b) Dinleme: Başlangıç – Orta – İleri Düzey
….. Başlangıç – Orta – İleri Düzey	c) Okuma: Başlangıç – Orta – İleri Düzey
….. Başlangıç – Orta – İleri Düzey	d) Yazma: Başlangıç – Orta – İleri Düzey

Aşağıdaki sorulara sadece "evet/hayır" olarak değil daha ayrıntılı cevaplar vermeye çalışınız!

1) Okuduğunuz Türkçe kitapları anlayabiliyor musunuz? (genel kitaplar, hikaye, roman, gazete)

2) Üniversitedeki ders kitaplarınızı anlayabiliyor musunuz? Ne derecede? Kitaplarınızdaki bütün kelimelerin anlamlarını biliyor musunuz?

3) Üniversitede ders aldığınız öğretim üyelerinin konuşmalarını anlayabiliyor musunuz?

4) Öğretmeninizle veya arkadaşlarınızla dersle alakalı sınıf içinde Türkçe iletişim kurabiliyor musunuz?

5) Türkçe olarak proje/ödev hazırlamada veya sınav kağıdını doldurmada (vize/final) herhangi bir zorlukla karşılaşıyor musunuz? (Dil açısından)

6) Sosyal ortamlarda Türkçe kullanarak iletişim kurabiliyor musunuz? (Okul dışında, kafede vb.)

7) Öğretim dili Türkçe olan bir okulda öğrenim görmekte olan bir öğrenci olarak ne tür zorluklarla karşılaşıyorsunuz?

Appendix 8 Participants Learning Turkish at The Turkish Teaching Centers

Countries	Number of Participants
Afghanistan	17
African	1
Arabic	7
Albania	3
Bangladesh	3
Bosnia	6
Bulgaria	1
Burundi	1
Djibouti	1
Chad	1
Chechnya	11
China	5
Indonesia	5
Ethiopia	1
République de Côte d'Ivoire	3
Philippines	1
Palestine	4
Gabon	2
Gambia	1
Ghana	3
Guinea	1
Georgia	1
Iraq	9
Iran	2
Kazakhstan	6
Kenya	1
Colombia	2
Congo	1
Korea	1
Liberia	1
Libya	2
Lebanon	1
Madagascar	2
Macedonia	1
Malawi	1
Malaysia	1
Egypt	1
Mongolia	2
Mauritania	1
Uzbekistan	2
Pakistan	1
Russia	1
Somalia	6
Sudan	1

Syria	12
Tajikistan	2
Tunisia	1
Turkmenistan	4
Ukraine	1
Jordan	2
Vanuatu	1
Yemen	5
Unknown	43
Total	**196 (52 different countries)**

Appendix 9 Composition of The Written Component of Turkish National Corpus (Aksan et al, 2013)

Domain	%	Medium	%
Imaginative	19	Book	58
Social Science	16	Periodicals	32
Art	7	Miscellaneous published	5
Commerce/ finance	8	Miscellaneous unpublished	3
Belief and thought	4	To-be-spoken	2
World affairs	20		
Applied science	8		
Natural science	4		
Leisure	14		

TNC-Domains
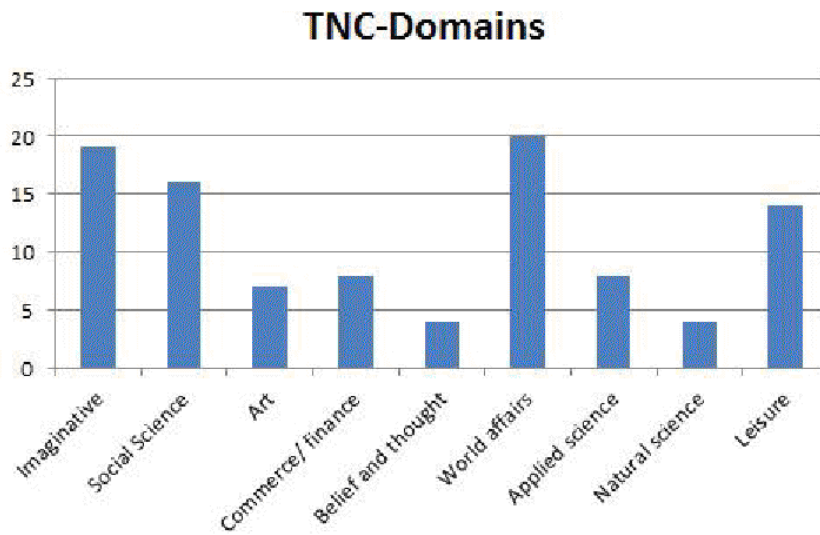
Appendix 10 The General Turkish Wordlist (2000 words)

No	Words	Type	Frequency	‰	No	Words	Type	Frequency	‰
1	ve	Conj	1004455	20.926	35	el	Noun	98790	2.058
2	olmak	Verb	999204	20.817	36	istemek	Verb	97275	2.027
3	bu	Det	798660	16.639	37	görmek	Verb	95798	1.996
4	o	Det	364829	7.601	38	ya	Conj	84608	1.763
5	etmek	Verb	351474	7.322	39	ise	Conj	83686	1.743
6	için	Postp	254849	5.309	40	sen	Pron	83663	1.743
7	yapmak	Verb	217208	4.525	41	son	Adj	83278	1.735
8	ben	Pron	216237	4.505	42	önce	Postp	83164	1.733
9	çok	Adv	209647	4.368	43	çıkmak	Verb	83129	1.732
10	demek	Verb	201303	4.194	44	büyük	Adj	82621	1.721
11	gibi	Postp	192756	4.016	45	iş	Noun	82281	1.714
12	daha	Adv	185489	3.864	46	gitmek	Verb	82079	1.710
13	kendi	Pron	183790	3.829	47	yan	Noun	81096	1.690
14	ne	Pron	176632	3.680	48	üzeri	Noun	80876	1.685
15	ile	Postp	176164	3.670	49	konu	Noun	80408	1.675
16	almak	Verb	152783	3.183	50	baş	Noun	76222	1.588
17	vermek	Verb	149745	3.120	51	diye	Postp	75667	1.576
18	iç	Noun	146397	3.050	52	bilmek	Verb	75150	1.566
19	sonra	Postp	144437	3.009	53	yol	Noun	74910	1.561
20	yer	Noun	141364	2.945	54	böyle	Adj	74183	1.545
21	her	Det	136938	2.853	55	yok	Noun	74008	1.542
22	ama	Conj	128015	2.667	56	ülke	Noun	71694	1.494
23	gelmek	Verb	127079	2.647	57	karşı	Postp	71465	1.489
24	kadar	Postp	124612	2.596	58	neden	Noun	71273	1.485
25	en	Adv	124386	2.591	59	kalmak	Verb	71020	1.480
26	yıl	Noun	123589	2.575	60	çocuk	Noun	70683	1.473
27	ara	Noun	122831	2.559	61	dünya	Noun	70277	1.464
28	zaman	Noun	121848	2.539	62	söylemek	Verb	69915	1.457
29	insan	Noun	118897	2.477	63	yeni	Adj	68915	1.436
30	değil	Conj	116632	2.430	64	göre	Postp	68147	1.420
31	var	Noun	115963	2.416	65	başlamak	Verb	67279	1.402
32	şey	Noun	109465	2.281	66	biri	Pron	66854	1.393
33	biz	Pron	102569	2.137	67	ilk	Adj	65952	1.374
34	gün	Noun	99462	2.072	68	kadın	Noun	65923	1.373
					69	orta	Noun	65826	1.371

70	durum	Noun	63823	1.330		107	oluşmak	Verb	43621	0.909
71	göz	Noun	63390	1.321		108	sıra	Noun	43045	0.897
72	iyi	Adj	62774	1.308		109	göstermek	Verb	42664	0.889
73	ancak	Conj	62438	1.301		110	hayat	Noun	42654	0.889
74	hal	Noun	61422	1.280		111	getirmek	Verb	42325	0.882
75	bakmak	Verb	60361	1.258		112	geçmek	Verb	42095	0.877
76	alan	Noun	59435	1.238		113	tek	Adj	41924	0.873
77	ev	Noun	59167	1.233		114	birlikte	Postp	41646	0.868
78	önemli	Adj	59093	1.231		115	bazı	Adj	41110	0.856
79	hiç	Adv	58360	1.216		116	çünkü	Conj	40913	0.852
80	bütün	Adj	58344	1.216		117	yaşamak	Verb	40746	0.849
81	aynı	Adj	57364	1.195		118	girmek	Verb	40575	0.845
82	söz	Noun	56346	1.174		119	şimdi	Adv	40488	0.844
83	bile	Adv	54411	1.134		120	gerekmek	Verb	40316	0.840
84	başka	Adj	53776	1.120		121	tüm	Det	40206	0.838
85	siz	Pron	52679	1.097		122	artık	Adv	40170	0.837
86	doğru	Adj	52572	1.095		123	düşünmek	Verb	39867	0.831
87	şekil	Noun	52185	1.087		124	çalışma	Noun	39730	0.828
88	devlet	Noun	52167	1.087		125	kız	Noun	39650	0.826
89	veya	Conj	51538	1.074		126	tarafından	Postp	39302	0.819
90	ön	Adj	51508	1.073		127	yine	Adv	39267	0.818
91	diğer	Adj	51339	1.070		128	tarih	Noun	38874	0.810
92	bulunmak	Verb	50570	1.054		129	yani	Conj	38054	0.793
93	kişi	Noun	49817	1.038		130	açmak	Verb	37907	0.790
94	hem	Conj	49763	1.037		131	genel	Adj	37789	0.787
95	nasıl	Adv	49502	1.031		132	ad	Noun	37784	0.787
96	alt	Adj	48948	1.020		133	fazla	Adj	37784	0.787
97	bura	Noun	48115	1.002		134	çalışmak	Verb	37498	0.781
98	bulmak	Verb	47660	0.993		135	baba	Noun	37455	0.780
99	dış	Adj	46566	0.970		136	ilgili	Postp	37453	0.780
100	şu	Det	45984	0.958		137	öyle	Adv	37192	0.775
101	dönem	Noun	45941	0.957		138	kim	Pron	36905	0.769
102	su	Noun	45549	0.949		139	varmak	Verb	36713	0.765
103	kullanmak	Verb	45524	0.948		140	üst	Adj	36506	0.761
104	sonuç	Noun	45511	0.948		141	anne	Noun	36469	0.760
105	sahip	Noun	44947	0.936		142	bilgi	Noun	36310	0.756
106	adam	Noun	43874	0.914		143	ses	Noun	36114	0.752

144	sadece	Adv	35885	0.748		181	bugün	Noun	28631	0.596
145	az	Adj	34981	0.729		182	yüksek	Adj	28409	0.592
146	gelen	Noun	34584	0.721		183	bölge	Noun	28167	0.587
147	sorun	Noun	34530	0.719		184	gece	Noun	28114	0.586
148	karar	Noun	34431	0.717		185	genç	Adj	27903	0.581
149	hiçbir	Det	34313	0.715		186	kapı	Noun	27761	0.578
150	örnek	Noun	34024	0.709		187	ora	Noun	27759	0.578
151	sağlamak	Verb	33873	0.706		188	biçim	Noun	27757	0.578
152	hemen	Adv	33759	0.703		189	saat	Noun	27727	0.578
153	küçük	Adj	33654	0.701		190	yaşam	Noun	27585	0.575
154	sistem	Noun	33322	0.694		191	grup	Noun	27563	0.574
155	ilişki	Noun	33253	0.693		192	belki	Adv	27555	0.574
156	süre	Noun	33190	0.691		193	kitap	Noun	27493	0.573
157	güzel	Adj	33085	0.689		194	değer	Noun	27419	0.571
158	para	Noun	32364	0.674		195	geri	Adv	27414	0.571
159	toplum	Noun	32301	0.673		196	açık	Adj	27250	0.568
160	uzun	Adj	32209	0.671		197	eski	Adj	27025	0.563
161	güç	Noun	32175	0.670		198	hep	Adv	26800	0.558
162	yüz	Noun	32123	0.669		199	arkadaş	Noun	26646	0.555
163	an	Noun	31956	0.666		200	kurmak	Verb	26635	0.555
164	ay	Noun	31271	0.651		201	özel	Adj	26581	0.554
165	olay	Noun	31141	0.649		202	kimse	Noun	26514	0.552
166	birbiri	Pron	30823	0.642		203	erkek	Noun	26451	0.551
167	dil	Noun	30770	0.641		204	pek	Adv	26328	0.549
168	eğitim	Noun	30693	0.639		205	herkes	Noun	26271	0.547
169	durmak	Verb	30342	0.632		206	sayı	Noun	26259	0.547
170	tam	Adj	30273	0.631		207	yapı	Noun	26197	0.546
171	biraz	Adj	30141	0.628		208	oran	Noun	26131	0.544
172	işte	Postp	30044	0.626		209	parti	Noun	25847	0.538
173	hareket	Noun	29976	0.625		210	özellikle	Adv	25803	0.538
174	hak	Noun	29842	0.622		211	dönmek	Verb	25790	0.537
175	halk	Noun	29637	0.617		212	aile	Noun	25671	0.535
176	kabul	Noun	29385	0.612		213	anlamak	Verb	25654	0.534
177	kez	Noun	29355	0.612		214	konuşmak	Verb	25623	0.534
178	devam	Noun	29044	0.605		215	temel	Noun	25600	0.533
179	farklı	Noun	28856	0.601		216	belirtmek	Verb	25558	0.532
180	gerçek	Noun	28805	0.600		217	tür	Noun	25280	0.527

218	çekmek	Verb	25234	0.526		255	ayrıca	Adv	21009	0.438
219	anlatmak	Verb	25060	0.522		256	belli	Adj	20952	0.437
220	yeniden	Adv	24730	0.515		257	koymak	Verb	20822	0.434
221	yakın	Noun	24716	0.515		258	yemek	Noun	20728	0.432
222	yaş	Noun	24657	0.514		259	tutmak	Verb	20611	0.429
223	bey	Noun	24524	0.511		260	mümkün	Adj	20580	0.429
224	gelir	Noun	24246	0.505		261	akıl	Noun	20568	0.429
225	amaç	Noun	24169	0.504		262	çevre	Noun	20457	0.426
226	yazmak	Verb	24082	0.502		263	haber	Noun	20428	0.426
227	yönetim	Noun	24030	0.501		264	özellik	Noun	20365	0.424
228	bırakmak	Verb	23999	0.500		265	üretim	Noun	20298	0.423
229	beklemek	Verb	23915	0.498		266	öte	Noun	20287	0.423
230	ifade	Noun	23904	0.498		267	birkaç	Det	20285	0.423
231	başkan	Noun	23881	0.498		268	görev	Noun	20252	0.422
232	dikkat	Noun	23372	0.487		269	nokta	Noun	20181	0.420
233	üzere	Postp	23338	0.486		270	art	Noun	20147	0.420
234	okul	Noun	23284	0.485		271	taraf	Noun	19955	0.416
235	savaş	Noun	23210	0.484		272	uygun	Adj	19944	0.416
236	açı	Noun	23128	0.482		273	düşünce	Noun	19912	0.415
237	fakat	Conj	23038	0.480		274	tabi	Conj	19666	0.410
238	gerek	Postp	22873	0.477		275	gerekmek	Verb	19566	0.408
239	içeri	Noun	22622	0.471		276	duymak	Verb	19526	0.407
240	düşmek	Verb	22376	0.466		277	okumak	Verb	19469	0.406
241	etki	Noun	22102	0.460		278	olsa	Conj	19468	0.406
242	hatta	Conj	21799	0.454		279	ekonomik	Adj	19405	0.404
243	hizmet	Noun	21758	0.453		280	sosyal	Adj	19278	0.402
244	süreç	Noun	21623	0.450		281	bağlı	Postp	19258	0.401
245	anlam	Noun	21600	0.450		282	allah	Noun	19156	0.399
246	soru	Noun	21347	0.445		283	zaten	Adv	19029	0.396
247	atmak	Verb	21254	0.443		284	geçen	Adj	19028	0.396
248	madde	Noun	21247	0.443		285	kısa	Adj	19017	0.396
249	hepsi	Pron	21216	0.442		286	hava	Noun	18709	0.390
250	merkez	Noun	21212	0.442		287	bölüm	Noun	18628	0.388
251	çıkarmak	Verb	21183	0.441		288	yön	Noun	18615	0.388
252	şöyle	Adv	21092	0.439		289	oturmak	Verb	18586	0.387
253	kaynak	Noun	21088	0.439		290	öğrenci	Noun	18533	0.386
254	toprak	Noun	21063	0.439		291	gelecek	Noun	18526	0.386

292	oda	Noun	18525	0.386
293	arka	Noun	18449	0.384
294	sormak	Verb	18366	0.383
295	evet	Adv	18299	0.381
296	aslında	Adv	18262	0.380
297	program	Noun	18224	0.380
298	birçok	Det	18180	0.379
299	düzey	Noun	18096	0.377
300	eğer	Conj	18022	0.375
301	taşımak	Verb	17884	0.373
302	nerede	Pron	17843	0.372
303	kültür	Noun	17812	0.371
304	artmak	Verb	17464	0.364
305	oyun	Noun	17289	0.360
306	sevmek	Verb	17198	0.358
307	araştırma	Verb	17132	0.357
308	ürün	Noun	17082	0.356
309	deniz	Noun	17077	0.356
310	fark	Noun	17014	0.354
311	zor	Adj	16998	0.354
312	din	Noun	16949	0.353
313	gelişme	Verb	16658	0.347
314	seçim	Noun	16611	0.346
315	hakkında	Postp	16580	0.345
316	film	Noun	16507	0.344
317	sabah	Noun	16312	0.340
318	gazete	Noun	16310	0.340
319	ileri	Noun	16275	0.339
320	ölüm	Noun	16184	0.337
321	çeşitli	Adj	16174	0.337
322	sanki	Adv	16047	0.334
323	hangi	Adj	15950	0.332
324	uygulama	Noun	15895	0.331
325	varlık	Noun	15771	0.329
326	sınıf	Noun	15695	0.327
327	rağmen	Postp	15692	0.327
328	politika	Noun	15532	0.324
329	hükümet	Noun	15406	0.321
330	yazı	Noun	15374	0.320
331	önem	Noun	15306	0.319
332	şirket	Noun	15261	0.318
333	ana	Noun	15209	0.317
334	kazanmak	Verb	15173	0.316
335	oğul	Noun	15173	0.316
336	yardım	Noun	15173	0.316
337	hala	Adv	15070	0.314
338	kurum	Noun	15050	0.314
339	belirlemek	Verb	15043	0.313
340	sürmek	Verb	15019	0.313
341	uzak	Adj	14984	0.312
342	bakan	Noun	14980	0.312
343	aramak	Verb	14870	0.310
344	sanat	Noun	14833	0.309
345	banka	Noun	14816	0.309
346	türlü	Adj	14807	0.308
347	ayrı	Adj	14790	0.308
348	üniversite	Noun	14764	0.308
349	ihtiyaç	Noun	14754	0.307
350	yabancı	Noun	14740	0.307
351	çalışan	Noun	14697	0.306
352	toplumsal	Adj	14693	0.306
353	can	Noun	14687	0.306
354	üye	Noun	14558	0.303
355	hız	Noun	14481	0.302
356	sektör	Noun	14458	0.301
357	kötü	Adj	14399	0.300
358	ait	Postp	14373	0.299
359	ilişkin	Postp	14336	0.299
360	oysa	Conj	14283	0.298
361	derece	Noun	14253	0.297
362	birlik	Noun	14239	0.297
363	sokak	Noun	14135	0.294
364	sürekli	Adv	14036	0.292
365	sık	Adv	13988	0.291

366	aşağı	Noun	13981	0.291
367	yoksa	Conj	13939	0.290
368	ışık	Noun	13908	0.290
369	uluslararası	Adj	13877	0.289
370	yukarı	Noun	13873	0.289
371	yaratmak	Verb	13867	0.289
372	ayak	Noun	13802	0.288
373	yazar	Noun	13799	0.287
374	gerekli	Adj	13783	0.287
375	yalnız	Adj	13767	0.287
376	ortam	Noun	13738	0.286
377	duygu	Noun	13698	0.285
378	akşam	Noun	13646	0.284
379	kent	Noun	13638	0.284
380	kısım	Noun	13593	0.283
381	benzer	Adj	13538	0.282
382	telefon	Noun	13442	0.280
383	bilim	Noun	13335	0.278
384	gerçekten	Adv	13224	0.276
385	geçmiş	Adj	13213	0.275
386	koşul	Noun	13168	0.274
387	cevap	Noun	13152	0.274
388	dolayısı	Conj	13140	0.274
389	kavram	Noun	13107	0.273
390	hanım	Noun	13102	0.273
391	batı	Noun	13079	0.272
392	bitmek	Verb	13061	0.272
393	siyasi	Adj	12975	0.270
394	geçirmek	Verb	12970	0.270
395	tekrar	Adv	12888	0.269
396	dışarı	Noun	12867	0.268
397	milli	Adj	12859	0.268
398	yalnızca	Adv	12844	0.268
399	kamu	Noun	12723	0.265
400	ortak	Noun	12717	0.265
401	karşılık	Noun	12650	0.264
402	bazen	Adv	12641	0.263

403	yönelik	Postp	12574	0.262
404	köy	Noun	12539	0.261
405	toplam	Noun	12521	0.261
406	ağır	Adj	12514	0.261
407	mal	Noun	12421	0.259
408	geniş	Adj	12417	0.259
409	doğal	Adj	12375	0.258
410	ekonomi	Noun	12258	0.255
411	işçi	Noun	12172	0.254
412	öğretmen	Noun	12154	0.253
413	gerçekleşmek	Verb	12128	0.253
414	kolay	Adj	12109	0.252
415	siyasal	Adj	12087	0.252
416	hafta	Noun	12047	0.251
417	görüş	Noun	12044	0.251
418	araç	Noun	12023	0.250
419	yarı	Noun	12012	0.250
420	müzik	Noun	11957	0.249
421	boyunca	Postp	11933	0.249
422	tanımak	Verb	11930	0.249
423	islam	Noun	11928	0.249
424	yöntem	Noun	11908	0.248
425	sebep	Noun	11895	0.248
426	şehir	Noun	11809	0.246
427	öğrenmek	Verb	11745	0.245
428	parça	Noun	11732	0.244
429	mücadele	Noun	11712	0.244
430	ders	Noun	11686	0.243
431	zorunda	Adj	11675	0.243
432	açıklama	Noun	11607	0.242
433	ölçü	Noun	11562	0.241
434	herhangi	Adj	11555	0.241
435	dolu	Adj	11539	0.240
436	yüzyıl	Noun	11511	0.240
437	yanlış	Adj	11492	0.239
438	hayır	Noun	11446	0.238
439	peki	Adv	11302	0.235

125

440	birden	Adv	11275	0.235
441	kalkmak	Verb	11237	0.234
442	paşa	Noun	11194	0.233
443	kanun	Noun	11131	0.232
444	katılmak	Verb	11106	0.231
445	miktar	Noun	11090	0.231
446	sevgi	Noun	11086	0.231
447	acı	Noun	11085	0.231
448	adım	Noun	11079	0.231
449	faaliyet	Noun	11063	0.230
450	meydana	Noun	11057	0.230
451	henüz	Adv	11029	0.230
452	efendi	Noun	11016	0.230
453	sol	Noun	10981	0.229
454	teknik	Adj	10977	0.229
455	beraber	Adj	10968	0.229
456	çıkar	Noun	10964	0.228
457	şiir	Noun	10939	0.228
458	asıl	Adj	10926	0.228
459	plan	Noun	10885	0.227
460	aşk	Noun	10880	0.227
461	bakımından	Postp	10855	0.226
462	ağaç	Noun	10845	0.226
463	görünmek	Verb	10809	0.225
464	örgüt	Noun	10789	0.225
465	ciddi	Adj	10775	0.224
466	ilgi	Noun	10770	0.224
467	güneş	Noun	10753	0.224
468	hukuk	Noun	10748	0.224
469	kural	Noun	10713	0.223
470	polis	Noun	10664	0.222
471	destek	Noun	10644	0.222
472	isim	Noun	10574	0.220
473	yüzden	Conj	10551	0.220
474	hasta	Noun	10547	0.220
475	yasa	Noun	10517	0.219
476	asker	Noun	10466	0.218

477	cumhuriyet	Noun	10452	0.218
478	beri	Postp	10431	0.217
479	adlı	Adj	10419	0.217
480	proje	Noun	10365	0.216
481	kurul	Noun	10350	0.216
482	renk	Noun	10302	0.215
483	güvenlik	Noun	10286	0.214
484	davranış	Noun	10235	0.213
485	yavaş	Adj	10216	0.213
486	beyaz	Adj	10199	0.212
487	büyümek	Verb	10184	0.212
488	yardımcı	Noun	10161	0.212
489	vergi	Noun	10098	0.210
490	başkası	Noun	10092	0.210
491	işlem	Noun	10069	0.210
492	kan	Noun	10069	0.210
493	hedef	Noun	10051	0.209
494	ayrılmak	Verb	10002	0.208
495	çözüm	Noun	9933	0.207
496	yaklaşık	Adj	9906	0.206
497	düzen	Noun	9902	0.206
498	yıllık	Adj	9891	0.206
499	kuruluş	Noun	9882	0.206
500	hazırlamak	Verb	9841	0.205
501	hastalık	Noun	9815	0.204
502	izlemek	Verb	9815	0.204
503	izin	Noun	9814	0.204
504	çerçeve	Noun	9793	0.204
505	güçlü	Adj	9780	0.204
506	rol	Noun	9755	0.203
507	yürümek	Verb	9750	0.203
508	hesap	Noun	9749	0.203
509	günümüz	Noun	9747	0.203
510	değişik	Adj	9693	0.202
511	ticaret	Noun	9659	0.201
512	ulaşmak	Verb	9648	0.201
513	toplantı	Noun	9632	0.201

514	ölmek	Verb	9605	0.200
515	askeri	Adj	9579	0.200
516	fikir	Noun	9569	0.199
517	uygulamak	Verb	9545	0.199
518	sanayi	Noun	9537	0.199
519	defa	Noun	9529	0.199
520	piyasa	Noun	9500	0.198
521	değişiklik	Noun	9487	0.198
522	elbette	Adv	9486	0.198
523	ulusal	Adj	9465	0.197
524	birey	Noun	9447	0.197
525	düşük	Adj	9420	0.196
526	mevcut	Adj	9419	0.196
527	kontrol	Noun	9407	0.196
528	ağız	Noun	9383	0.195
529	demokrasi	Noun	9378	0.195
530	anlayış	Noun	9351	0.195
531	doktor	Noun	9325	0.194
532	derin	Adj	9308	0.194
533	boş	Adj	9279	0.193
534	hızlı	Adj	9256	0.193
535	teknoloji	Noun	9252	0.193
536	düzenlemek	Verb	9217	0.192
537	enerji	Noun	9205	0.192
538	sermaye	Noun	9197	0.192
539	başarı	Noun	9194	0.192
540	genellikle	Adv	9171	0.191
541	konuşma	Noun	9166	0.191
542	pazar	Noun	9159	0.191
543	millet	Noun	9142	0.190
544	dün	Noun	9141	0.190
545	sağ	Noun	9135	0.190
546	basın	Noun	9126	0.190
547	dolmak	Verb	9101	0.190
548	inanmak	Verb	9087	0.189
549	oynamak	Verb	9075	0.189
550	doğmak	Verb	9072	0.189

551	takım	Noun	9068	0.189
552	yatırım	Noun	8964	0.187
553	koca	Noun	8927	0.186
554	tartışma	Verb	8902	0.185
555	açıklamak	Verb	8895	0.185
556	dakika	Noun	8845	0.184
557	firma	Noun	8842	0.184
558	anlam	Noun	8839	0.184
559	sağlık	Noun	8832	0.184
560	kaç	Adj	8773	0.183
561	ruh	Noun	8749	0.182
562	sınır	Noun	8712	0.182
563	işaret	Noun	8708	0.181
564	kafa	Noun	8703	0.181
565	resmi	Adj	8637	0.180
566	masa	Noun	8635	0.180
567	ters	Adj	8621	0.180
568	gül	Noun	8592	0.179
569	bitki	Noun	8590	0.179
570	tanrı	Noun	8536	0.178
571	artış	Noun	8533	0.178
572	başbakan	Noun	8518	0.177
573	ilke	Noun	8500	0.177
574	hayvan	Noun	8497	0.177
575	korumak	Verb	8478	0.177
576	değişmek	Verb	8464	0.176
577	iktidar	Noun	8441	0.176
578	nitelik	Noun	8441	0.176
579	toplamak	Verb	8385	0.175
580	üstelik	Adv	8373	0.174
581	yaşayan	Noun	8333	0.174
582	resim	Noun	8328	0.174
583	kullanım	Noun	8313	0.173
584	başarılı	Adj	8311	0.173
585	anayasa	Noun	8309	0.173
586	fiyat	Noun	8301	0.173
587	eş	Noun	8287	0.173

588	iyice	Adv	8276	0.172	625	tercih	Noun	7845	0.163
589	sanmak	Verb	8276	0.172	626	eser	Noun	7844	0.163
590	kardeş	Noun	8257	0.172	627	talep	Noun	7825	0.163
591	acaba	Adv	8250	0.172	628	tamamen	Adv	7799	0.162
592	bakış	Noun	8243	0.172	629	ceza	Noun	7795	0.162
593	çiçek	Noun	8236	0.172	630	değişim	Noun	7748	0.161
594	veri	Noun	8235	0.172	631	yönetici	Noun	7726	0.161
595	satmak	Verb	8233	0.172	632	karanlık	Noun	7723	0.161
596	konu	Noun	8194	0.171	633	içermek	Verb	7677	0.160
597	yeterli	Adj	8194	0.171	634	etkili	Adj	7670	0.160
598	iletişim	Noun	8192	0.171	635	merak	Noun	7598	0.158
599	ateş	Noun	8159	0.170	636	değerlendirmek	Verb	7585	0.158
600	duvar	Noun	8157	0.170	637	kesim	Noun	7571	0.158
601	taş	Noun	8143	0.170	638	neredeyse	Adv	7543	0.157
602	geliştirmek	Verb	8127	0.169	639	tip	Noun	7527	0.157
603	gene	Adv	8109	0.169	640	suç	Noun	7524	0.157
604	kara	Adj	8109	0.169	641	belirli	Adj	7516	0.157
605	kol	Noun	8088	0.169	642	bahçe	Noun	7497	0.156
606	kuş	Noun	8063	0.168	643	karşın	Postp	7471	0.156
607	boy	Noun	8061	0.168	644	sıcak	Adj	7465	0.156
608	doğrudan	Adj	8055	0.168	645	dinlemek	Verb	7453	0.155
609	tamam	Noun	8051	0.168	646	hoş	Adj	7432	0.155
610	tane	Noun	8050	0.168	647	yaz	Noun	7431	0.155
611	model	Noun	8044	0.168	648	uç	Noun	7404	0.154
612	mutlu	Adj	8035	0.167	649	dahil	Postp	7390	0.154
613	yanıt	Noun	8031	0.167	650	karı	Noun	7385	0.154
614	dolayı	Postp	8006	0.167	651	dağ	Noun	7375	0.154
615	televizyon	Noun	7974	0.166	652	yatak	Noun	7361	0.153
616	maç	Noun	7959	0.166	653	yerel	Adj	7351	0.153
617	ince	Adj	7950	0.166	654	yaşlı	Adj	7350	0.153
618	müslüman	Noun	7908	0.165	655	mahkeme	Noun	7342	0.153
619	ünlü	Adj	7897	0.165	656	kat	Noun	7326	0.153
620	idi	Verb	7896	0.165	657	niye	Adv	7326	0.153
621	itibaren	Postp	7864	0.164	658	dost	Noun	7305	0.152
622	lira	Noun	7864	0.164	659	mutlaka	Adv	7298	0.152
623	hoca	Noun	7852	0.164	660	gizli	Adj	7294	0.152
624	şart	Noun	7851	0.164	661	kesmek	Verb	7263	0.151

662	yükselmek	Verb	7262	0.151		699	sonraki	Adj	6700	0.140
663	fotoğraf	Noun	7252	0.151		700	olumsuz	Adj	6665	0.139
664	yoğun	Adj	7222	0.150		701	kesin	Adj	6663	0.139
665	eylem	Noun	7220	0.150		702	sunmak	Verb	6659	0.139
666	doğu	Noun	7160	0.149		703	saymak	Verb	6648	0.139
667	dayanmak	Verb	7154	0.149		704	normal	Adj	6636	0.138
668	kar	Noun	7127	0.148		705	silah	Noun	6628	0.138
669	doğa	Noun	7122	0.148		706	imkan	Noun	6605	0.138
670	kredi	Noun	7108	0.148		707	hazır	Adj	6599	0.137
671	ilginç	Adj	7105	0.148		708	canlı	Noun	6596	0.137
672	tespit	Noun	7088	0.148		709	serbest	Adj	6587	0.137
673	siyaset	Noun	7086	0.148		710	rekabet	Noun	6580	0.137
674	yıldız	Noun	7084	0.148		711	çay	Noun	6567	0.137
675	işletme	Noun	7079	0.147		712	sanatçı	Noun	6558	0.137
676	nüfus	Noun	7051	0.147		713	lazım	Noun	6556	0.137
677	tedavi	Noun	7048	0.147		714	yerinde	Noun	6550	0.136
678	ortalama	Noun	7039	0.147		715	katkı	Noun	6515	0.136
679	gündem	Noun	7034	0.147		716	kimlik	Noun	6502	0.135
680	mektup	Noun	6962	0.145		717	bakanlık	Noun	6500	0.135
681	esas	Noun	6956	0.145		718	dolar	Noun	6486	0.135
682	savunmak	Verb	6932	0.144		719	kağıt	Noun	6480	0.135
683	pay	Noun	6915	0.144		720	hele	Conj	6450	0.134
684	olumlu	Adj	6914	0.144		721	sigara	Noun	6450	0.134
685	sevgili	Noun	6910	0.144		722	boyut	Noun	6393	0.133
686	erken	Adv	6863	0.143		723	adet	Noun	6368	0.133
687	saye	Postp	6863	0.143		724	iddia	Noun	6364	0.133
688	sözcük	Noun	6828	0.142		725	mesela	Postp	6362	0.133
689	kriz	Noun	6809	0.142		726	şiddet	Noun	6360	0.133
690	lider	Noun	6806	0.142		727	yaklaşım	Noun	6356	0.132
691	korku	Noun	6798	0.142		728	kırmızı	Adj	6350	0.132
692	meclis	Noun	6780	0.141		729	barış	Noun	6341	0.132
693	gelişmek	Verb	6776	0.141		730	sahne	Noun	6334	0.132
694	olanak	Noun	6772	0.141		731	koşmak	Verb	6333	0.132
695	özgürlük	Noun	6762	0.141		732	asla	Adv	6329	0.132
696	kişilik	Noun	6735	0.140		733	hayal	Noun	6324	0.132
697	kelime	Noun	6722	0.140		734	kalp	Noun	6324	0.132
698	günlük	Adj	6721	0.140		735	fırsat	Noun	6310	0.131

736	zarar	Noun	6291	0.131
737	aday	Noun	6274	0.131
738	hücre	Noun	6260	0.130
739	soğuk	Adj	6248	0.130
740	oyuncu	Noun	6237	0.130
741	eylül	Noun	6236	0.130
742	belediye	Noun	6233	0.130
743	değiştirmek	Verb	6224	0.130
744	yayın	Noun	6223	0.130
745	kaldırmak	Verb	6218	0.130
746	rapor	Noun	6184	0.129
747	gören	Noun	6171	0.129
748	yüzünden	Postp	6158	0.128
749	internet	Noun	6153	0.128
750	abi	Noun	6146	0.128
751	gider	Noun	6142	0.128
752	kültürel	Adj	6106	0.127
753	köşe	Noun	6105	0.127
754	emin	Noun	6101	0.127
755	satış	Noun	6099	0.127
756	milletvekili	Noun	6097	0.127
757	rahat	Adj	6086	0.127
758	güven	Noun	6053	0.126
759	kere	Noun	6051	0.126
760	düzenleme	Noun	6047	0.126
761	temsil	Noun	6045	0.126
762	geç	Adj	6012	0.125
763	yeşil	Adj	5996	0.125
764	vakit	Noun	5991	0.125
765	seçmek	Verb	5946	0.124
766	altın	Noun	5942	0.124
767	öncelik	Noun	5933	0.124
768	tepki	Noun	5927	0.123
769	beyin	Noun	5924	0.123
770	gülmek	Verb	5910	0.123
771	şarkı	Noun	5892	0.123
772	unutmak	Verb	5866	0.122

773	görüşme	Noun	5851	0.122
774	baskı	Noun	5849	0.122
775	ordu	Noun	5841	0.122
776	metin	Noun	5838	0.122
777	ağlamak	Verb	5832	0.122
778	sayın	Adj	5817	0.121
779	tarım	Noun	5805	0.121
780	aşırı	Adj	5804	0.121
781	oy	Noun	5795	0.121
782	dahi	Conj	5789	0.121
783	gazeteci	Noun	5785	0.121
784	üretmek	Verb	5785	0.121
785	kuzey	Noun	5779	0.120
786	modern	Adj	5747	0.120
787	götürmek	Verb	5714	0.119
788	inmek	Verb	5705	0.119
789	malzeme	Noun	5696	0.119
790	demokratik	Adj	5691	0.119
791	dek	Postp	5687	0.118
792	herhalde	Adv	5667	0.118
793	kaçmak	Verb	5666	0.118
794	ilaç	Noun	5643	0.118
795	kaybetmek	Verb	5640	0.118
796	adeta	Adv	5619	0.117
797	bağımsız	Adj	5613	0.117
798	bilimsel	Adj	5613	0.117
799	aşama	Noun	5591	0.116
800	mali	Adj	5580	0.116
801	kısaca	Adv	5579	0.116
802	insanlık	Noun	5577	0.116
803	etraf	Noun	5550	0.116
804	anlaşma	Noun	5535	0.115
805	çaba	Noun	5526	0.115
806	başlangıç	Noun	5518	0.115
807	roman	Noun	5512	0.115
808	bilgisayar	Noun	5509	0.115
809	birim	Noun	5505	0.115

810	ekmek	Noun	5502	0.115	847	hikaye	Noun	5209	0.109
811	sayılı	Adj	5483	0.114	848	öykü	Noun	5206	0.108
812	iz	Noun	5475	0.114	849	faktör	Noun	5184	0.108
813	müşteri	Noun	5472	0.114	850	nitekim	Adv	5175	0.108
814	akmak	Verb	5463	0.114	851	sorumlu	Noun	5161	0.108
815	zengin	Adj	5459	0.114	852	iken	Adv	5158	0.107
816	basit	Adj	5437	0.113	853	bol	Adj	5154	0.107
817	kaynaklanmak	Verb	5435	0.113	854	hissetmek	Verb	5148	0.107
818	sürdürmek	Verb	5430	0.113	855	vurmak	Verb	5135	0.107
819	dair	Postp	5429	0.113	856	edebiyat	Noun	5129	0.107
820	terk	Noun	5429	0.113	857	hüküm	Noun	5126	0.107
821	meslek	Noun	5421	0.113	858	mesaj	Noun	5123	0.107
822	salon	Noun	5418	0.113	859	öğretim	Noun	5110	0.106
823	sivil	Adj	5416	0.113	860	açıkça	Adv	5103	0.106
824	istek	Noun	5405	0.113	861	pencere	Noun	5091	0.106
825	teslim	Noun	5394	0.112	862	ihracat	Noun	5089	0.106
826	geçerli	Adj	5370	0.112	863	geleneksel	Adj	5083	0.106
827	kapsam	Noun	5363	0.112	864	tıpkı	Noun	5082	0.106
828	engel	Noun	5361	0.112	865	giriş	Noun	5077	0.106
829	at	Noun	5360	0.112	866	sınırlı	Noun	5074	0.106
830	ziyaret	Noun	5351	0.111	867	medya	Noun	5070	0.106
831	zorunlu	Adj	5350	0.111	868	bağırmak	Verb	5036	0.105
832	dava	Noun	5330	0.111	869	cinsel	Adj	5023	0.105
833	siyah	Adj	5329	0.111	870	yaygın	Adj	5017	0.105
834	yatmak	Verb	5315	0.111	871	çekilmek	Verb	5015	0.104
835	kuvvet	Noun	5304	0.111	872	şair	Noun	5014	0.104
836	sorumluluk	Noun	5299	0.110	873	haklı	Noun	5013	0.104
837	balık	Noun	5295	0.110	874	unsur	Noun	4999	0.104
838	yarın	Noun	5285	0.110	875	aydın	Adj	4997	0.104
839	etkilemek	Verb	5284	0.110	876	farklılık	Noun	4991	0.104
840	sinema	Noun	5284	0.110	877	çeşit	Noun	4953	0.103
841	yarım	Adj	5272	0.110	878	problem	Noun	4941	0.103
842	çift	Adj	5260	0.110	879	etkin	Adj	4935	0.103
843	takip	Noun	5258	0.110	880	bina	Noun	4929	0.103
844	tiyatro	Noun	5246	0.109	881	petrol	Noun	4928	0.103
845	sert	Adj	5239	0.109	882	mayıs	Noun	4917	0.102
846	konum	Noun	5211	0.109	883	idare	Noun	4913	0.102

884	tavır	Noun	4908	0.102		921	gelişmiş	Adj	4625	0.096
885	yetkili	Noun	4908	0.102		922	niçin	Adv	4625	0.096
886	kapalı	Adj	4897	0.102		923	müdür	Noun	4618	0.096
887	uzman	Noun	4892	0.102		924	kenar	Noun	4598	0.096
888	zira	Conj	4892	0.102		925	dayalı	Postp	4594	0.096
889	dergi	Noun	4875	0.102		926	başkanlık	Noun	4590	0.096
890	düşman	Noun	4872	0.102		927	devrim	Noun	4588	0.096
891	teşekkür	Noun	4861	0.101		928	memur	Noun	4587	0.096
892	çağdaş	Adj	4857	0.101		929	ocak	Noun	4585	0.096
893	ücret	Noun	4840	0.101		930	görevli	Noun	4584	0.096
894	etkinlik	Adj	4831	0.101		931	aracılık	Noun	4580	0.095
895	sene	Noun	4809	0.100		932	uzanmak	Verb	4577	0.095
896	cep	Noun	4808	0.100		933	içmek	Verb	4563	0.095
897	güney	Noun	4797	0.100		934	vücut	Noun	4560	0.095
898	ağırlık	Noun	4782	0.100		935	ilan	Noun	4559	0.095
899	gelişim	Noun	4778	0.100		936	kulak	Noun	4538	0.095
900	saç	Noun	4775	0.099		937	yağmur	Noun	4533	0.094
901	uyum	Noun	4769	0.099		938	itibariyle	Noun	4524	0.094
902	tatlı	Adj	4768	0.099		939	peş	Noun	4521	0.094
903	belge	Noun	4766	0.099		940	elektrik	Noun	4520	0.094
904	yürek	Noun	4757	0.099		941	aralık	Noun	4519	0.094
905	göndermek	Verb	4753	0.099		942	borç	Noun	4504	0.094
906	tahmin	Noun	4744	0.099		943	doğrultu	Noun	4493	0.094
907	hakim	Noun	4738	0.099		944	çabuk	Adj	4487	0.093
908	saygı	Noun	4731	0.099		945	kaydetmek	Verb	4473	0.093
909	görüntü	Noun	4727	0.098		946	yazılı	Noun	4472	0.093
910	sağlıklı	Noun	4697	0.098		947	orman	Noun	4467	0.093
911	öz	Noun	4692	0.098		948	felsefe	Noun	4464	0.093
912	kişisel	Adj	4687	0.098		949	heyecan	Noun	4457	0.093
913	emir	Noun	4683	0.098		950	yaprak	Noun	4455	0.093
914	eşit	Adj	4667	0.097		951	işbirliği	Noun	4440	0.093
915	koku	Noun	4636	0.097		952	dışişleri	Noun	4427	0.092
916	tablo	Noun	4633	0.097		953	kalabalık	Noun	4423	0.092
917	ertesi	Adj	4631	0.096		954	risk	Noun	4412	0.092
918	beden	Noun	4627	0.096		955	galiba	Adv	4410	0.092
919	hafif	Adj	4626	0.096		956	yasal	Adj	4410	0.092
920	çağ	Noun	4625	0.096		957	sevinç	Noun	4408	0.092

958	rahatsız	Adj	4380	0.091
959	kış	Noun	4357	0.091
960	cumhurbaşkanı	Noun	4348	0.091
961	mavi	Adj	4340	0.090
962	sarı	Adj	4340	0.090
963	kuran	Noun	4338	0.090
964	incelemek	Verb	4336	0.090
965	rüzgar	Noun	4335	0.090
966	deprem	Noun	4317	0.090
967	gönül	Noun	4298	0.090
968	yeter	Adj	4293	0.089
969	birtakım	Adj	4281	0.089
970	kalite	Noun	4271	0.089
971	garip	Adj	4263	0.089
972	klasik	Adj	4250	0.089
973	sefer	Noun	4242	0.088
974	sultan	Noun	4242	0.088
975	spor	Noun	4239	0.088
976	sıkı	Adj	4235	0.088
977	yargı	Noun	4231	0.088
978	tarihsel	Adj	4222	0.088
979	faiz	Noun	4221	0.088
980	kamuoyu	Noun	4209	0.088
981	yetki	Noun	4200	0.088
982	mart	Noun	4199	0.087
983	mesele	Noun	4190	0.087
984	yazık	Noun	4188	0.087
985	maddi	Adj	4186	0.087
986	değerli	Noun	4183	0.087
987	enflasyon	Noun	4181	0.087
988	bugünkü	Noun	4175	0.087
989	politik	Adj	4165	0.087
990	denge	Noun	4156	0.087
991	çalmak	Verb	4151	0.086
992	il	Noun	4146	0.086
993	sırt	Noun	4142	0.086
994	kesinlikle	Adv	4137	0.086

995	gerektirmek	Verb	4131	0.086
996	nisan	Noun	4125	0.086
997	yalan	Noun	4120	0.086
998	emek	Noun	4110	0.086
999	demir	Noun	4100	0.085
1000	yerli	Adj	4091	0.085
1001	kasım	Noun	4084	0.085
1002	çizgi	Noun	4077	0.085
1003	bilinç	Noun	4076	0.085
1004	aylık	Noun	4072	0.085
1005	kayıt	Noun	4072	0.085
1006	renkli	Noun	4068	0.085
1007	önlem	Noun	4062	0.085
1008	sıkıntı	Noun	4058	0.085
1009	kocaman	Adj	4051	0.084
1010	bağ	Noun	4050	0.084
1011	otel	Noun	4050	0.084
1012	gerekse	Conj	4043	0.084
1013	kuşkusuz	Noun	4039	0.084
1014	kapatmak	Verb	4021	0.084
1015	kuru	Adj	4021	0.084
1016	top	Noun	4013	0.084
1017	neyse	Adv	4005	0.083
1018	falan	Postp	4000	0.083
1019	denli	Adj	3999	0.083
1020	dönüşmek	Verb	3999	0.083
1021	özgür	Adj	3978	0.083
1022	vatandaş	Noun	3977	0.083
1023	yeterince	Adv	3972	0.083
1024	maliyet	Noun	3956	0.082
1025	futbol	Noun	3950	0.082
1026	özgü	Adj	3936	0.082
1027	sabit	Adj	3930	0.082
1028	net	Adj	3929	0.082
1029	tehdit	Noun	3928	0.082
1030	kalem	Noun	3927	0.082
1031	ticari	Adj	3927	0.082

1032	kayıp	Adj	3916	0.082		1069	tarz	Noun	3705	0.077
1033	tehlike	Noun	3915	0.082		1070	zemin	Noun	3703	0.077
1034	bebek	Noun	3893	0.081		1071	husus	Noun	3701	0.077
1035	takdirde	Postp	3888	0.081		1072	tüketim	Noun	3701	0.077
1036	karşılıklı	Adj	3886	0.081		1073	dede	Noun	3698	0.077
1037	adalet	Noun	3881	0.081		1074	çıkış	Noun	3693	0.077
1038	yorum	Noun	3872	0.081		1075	hastane	Noun	3687	0.077
1039	devir	Noun	3871	0.081		1076	teklif	Noun	3687	0.077
1040	egemen	Adj	3868	0.081		1077	deyiş	Noun	3686	0.077
1041	boyun	Noun	3858	0.080		1078	yük	Noun	3665	0.076
1042	denetim	Noun	3854	0.080		1079	servis	Noun	3663	0.076
1043	zayıf	Adj	3850	0.080		1080	hata	Noun	3662	0.076
1044	analiz	Noun	3838	0.080		1081	tuhaf	Adj	3657	0.076
1045	dar	Adj	3836	0.080		1082	ekim	Noun	3652	0.076
1046	evlilik	Noun	3832	0.080		1083	yaramak	Verb	3642	0.076
1047	işlev	Noun	3832	0.080		1084	bireysel	Adj	3636	0.076
1048	vurgulamak	Verb	3822	0.080		1085	yanmak	Verb	3636	0.076
1049	çoğunluk	Noun	3816	0.080		1086	anlamlı	Adj	3634	0.076
1050	yürütmek	Verb	3797	0.079		1087	teşvik	Noun	3634	0.076
1051	benzemek	Verb	3789	0.079		1088	eksik	Noun	3624	0.076
1052	terör	Noun	3778	0.079		1089	temmuz	Noun	3622	0.075
1053	aşık	Noun	3775	0.079		1090	uyku	Noun	3619	0.075
1054	mutluluk	Adj	3774	0.079		1091	dolaşmak	Verb	3617	0.075
1055	sonsuz	Noun	3765	0.078		1092	işsizlik	Noun	3606	0.075
1056	ufak	Adj	3753	0.078		1093	bakım	Noun	3595	0.075
1057	döviz	Noun	3745	0.078		1094	usta	Noun	3593	0.075
1058	ek	Noun	3745	0.078		1095	topluluk	Noun	3581	0.075
1059	sonradan	Adv	3738	0.078		1096	sessiz	Adj	3575	0.074
1060	bardak	Noun	3734	0.078		1097	ilave	Noun	3568	0.074
1061	umut	Noun	3734	0.078		1098	katı	Noun	3550	0.074
1062	gülümsemek	Verb	3733	0.078		1099	işlemek	Verb	3544	0.074
1063	kitle	Noun	3730	0.078		1100	mekan	Noun	3544	0.074
1064	davet	Noun	3725	0.078		1101	istihdam	Noun	3530	0.074
1065	yağ	Noun	3725	0.078		1102	sağlam	Adj	3515	0.073
1066	güzellik	Adj	3720	0.078		1103	geçiş	Noun	3500	0.073
1067	koruma	Noun	3718	0.077		1104	temiz	Adj	3493	0.073
1068	gaz	Noun	3707	0.077		1105	boşluk	Noun	3478	0.072

1106	haziran	Noun	3476	0.072	1143	komutan	Noun	3299	0.069
1107	geçici	Adj	3474	0.072	1144	fiziksel	Adj	3293	0.069
1108	dal	Noun	3472	0.072	1145	nihayet	Noun	3291	0.069
1109	farkında	Noun	3472	0.072	1146	bayram	Noun	3290	0.069
1110	sürmek	Verb	3471	0.072	1147	filan	Pron	3286	0.068
1111	karşılamak	Verb	3466	0.072	1148	ömür	Noun	3286	0.068
1112	yurt	Noun	3460	0.072	1149	komisyon	Noun	3281	0.068
1113	ayakta	Noun	3458	0.072	1150	köpek	Noun	3260	0.068
1114	inanç	Noun	3456	0.072	1151	süt	Noun	3254	0.068
1115	cami	Noun	3455	0.072	1152	seviye	Noun	3248	0.068
1116	köylü	Noun	3455	0.072	1153	dalga	Noun	3244	0.068
1117	kalın	Adj	3454	0.072	1154	ahlak	Noun	3243	0.068
1118	tat	Noun	3447	0.072	1155	dizi	Noun	3240	0.068
1119	kral	Noun	3438	0.072	1156	bütçe	Noun	3224	0.067
1120	aykırı	Adj	3436	0.072	1157	ucuz	Adj	3216	0.067
1121	ödeme	Noun	3433	0.072	1158	aksine	Adv	3215	0.067
1122	tuz	Noun	3417	0.071	1159	kıyı	Noun	3215	0.067
1123	desteklemek	Verb	3409	0.071	1160	öğle	Noun	3215	0.067
1124	olgu	Noun	3407	0.071	1161	memleket	Noun	3199	0.067
1125	eskiden	Adv	3395	0.071	1162	yasak	Noun	3199	0.067
1126	somut	Adj	3383	0.070	1163	cümle	Noun	3192	0.067
1127	lütfen	Adv	3380	0.070	1164	özen	Noun	3190	0.066
1128	şans	Noun	3377	0.070	1165	olağanüstü	Adj	3186	0.066
1129	personel	Noun	3362	0.070	1166	kutsal	Adj	3183	0.066
1130	sıcaklık	Noun	3360	0.070	1167	tutum	Noun	3183	0.066
1131	çoktan	Adv	3357	0.070	1168	huzur	Noun	3182	0.066
1132	kaşık	Noun	3357	0.070	1169	psikolojik	Adj	3179	0.066
1133	aktif	Adj	3355	0.070	1170	devrimci	Adj	3178	0.066
1134	saha	Noun	3347	0.070	1171	gümrük	Noun	3178	0.066
1135	işgal	Noun	3338	0.070	1172	daima	Adv	3173	0.066
1136	sosyalist	Noun	3335	0.069	1173	laf	Noun	3159	0.066
1137	azalmak	Verb	3333	0.069	1174	şeker	Noun	3154	0.066
1138	düzenli	Noun	3322	0.069	1175	öbür	Adj	3152	0.066
1139	sözleşme	Noun	3316	0.069	1176	tüketici	Noun	3151	0.066
1140	deli	Adj	3309	0.069	1177	elektronik	Adj	3150	0.066
1141	gerekçe	Noun	3304	0.069	1178	şubat	Noun	3141	0.065
1142	nefes	Noun	3301	0.069	1179	aşkın	Noun	3139	0.065

1180	işgücü	Noun	3134	0.065	1217	tanım	Noun	2993	0.062
1181	doğum	Noun	3127	0.065	1218	emniyet	Noun	2992	0.062
1182	uçak	Noun	3125	0.065	1219	silahlı	Adj	2992	0.062
1183	bağlam	Noun	3121	0.065	1220	sayfa	Noun	2987	0.062
1184	öneri	Noun	3106	0.065	1221	iktisadi	Adj	2986	0.062
1185	yaklaşmak	Verb	3094	0.064	1222	belirgin	Adj	2984	0.062
1186	halen	Adv	3093	0.064	1223	haydi	Interj	2984	0.062
1187	islami	Adj	3086	0.064	1224	harcama	Noun	2974	0.062
1188	metre	Noun	3075	0.064	1225	evrensel	Adj	2973	0.062
1189	ideolojik	Adj	3074	0.064	1226	eşya	Noun	2968	0.062
1190	yetenek	Noun	3073	0.064	1227	azınlık	Noun	2966	0.062
1191	müdahale	Noun	3072	0.064	1228	lise	Noun	2966	0.062
1192	teknolojik	Adj	3067	0.064	1229	eğilim	Noun	2964	0.062
1193	teşkil	Noun	3063	0.064	1230	yumuşak	Adj	2959	0.062
1194	ısrar	Noun	3058	0.064	1231	sürü	Noun	2956	0.062
1195	bizzat	Adv	3054	0.064	1232	içerik	Noun	2955	0.062
1196	sakin	Adj	3052	0.064	1233	parmak	Noun	2953	0.062
1197	gemi	Noun	3048	0.064	1234	mahkum	Adj	2951	0.061
1198	dev	Noun	3046	0.063	1235	gerçi	Adv	2949	0.061
1199	gençlik	Noun	3046	0.063	1236	üstün	Adj	2949	0.061
1200	öldürmek	Verb	3038	0.063	1237	cins	Noun	2946	0.061
1201	arz	Noun	3034	0.063	1238	imza	Noun	2944	0.061
1202	ortalık	Noun	3031	0.063	1239	korkunç	Adj	2942	0.061
1203	suretiyle	Noun	3026	0.063	1240	kavga	Noun	2935	0.061
1204	emekli	Noun	3021	0.063	1241	idari	Adj	2930	0.061
1205	kılmak	Verb	3015	0.063	1242	imparatorluk	Noun	2930	0.061
1206	küresel	Adj	3014	0.063	1243	kurban	Noun	2929	0.061
1207	yapım	Noun	3013	0.063	1244	ak	Adj	2927	0.061
1208	saldırı	Noun	3010	0.063	1245	potansiyel	Adj	2925	0.061
1209	artırmak	Verb	3009	0.063	1246	kaçınılmaz	Adj	2922	0.061
1210	toplu	Adj	3007	0.063	1247	koltuk	Noun	2921	0.061
1211	ağustos	Noun	3004	0.063	1248	kolayca	Adv	2920	0.061
1212	yolculuk	Noun	3004	0.063	1249	inşa	Noun	2917	0.061
1213	arzu	Noun	3000	0.063	1250	ada	Noun	2914	0.061
1214	gıda	Noun	2999	0.062	1251	transfer	Noun	2913	0.061
1215	makine	Noun	2999	0.062	1252	anı	Noun	2909	0.061
1216	not	Noun	2998	0.062	1253	mutfak	Noun	2908	0.061

1254	ibaret	Postp	2907	0.061		1291	başlıca	Adj	2736	0.057
1255	eşlik	Noun	2904	0.061		1292	mensup	Noun	2736	0.057
1256	ölü	Noun	2898	0.060		1293	dua	Noun	2733	0.057
1257	bunca	Adj	2894	0.060		1294	cam	Noun	2728	0.057
1258	test	Noun	2893	0.060		1295	namaz	Noun	2725	0.057
1259	saptamak	Verb	2884	0.060		1296	ihtimal	Noun	2718	0.057
1260	yerleşmek	Verb	2878	0.060		1297	standart	Noun	2710	0.056
1261	soluk	Noun	2875	0.060		1298	aç	Adj	2706	0.056
1262	gelenek	Noun	2874	0.060		1299	manevi	Adj	2701	0.056
1263	parlak	Adj	2870	0.060		1300	uğramak	Verb	2700	0.056
1264	ilerlemek	Verb	2862	0.060		1301	usul	Noun	2695	0.056
1265	olasılık	Noun	2860	0.060		1302	sohbet	Noun	2689	0.056
1266	tehlikeli	Adj	2860	0.060		1303	hayli	Adv	2682	0.056
1267	yayılmak	Verb	2850	0.059		1304	katılım	Noun	2682	0.056
1268	sigorta	Noun	2848	0.059		1305	kurtuluş	Noun	2677	0.056
1269	diş	Noun	2843	0.059		1306	kalkınma	Noun	2674	0.056
1270	bayan	Noun	2841	0.059		1307	kapsamak	Verb	2674	0.056
1271	inceleme	Noun	2835	0.059		1308	yurtdışı	Noun	2673	0.056
1272	memnun	Adj	2834	0.059		1309	mahalle	Noun	2670	0.056
1273	bağlantı	Noun	2816	0.059		1310	öngörmek	Verb	2667	0.056
1274	amca	Noun	2809	0.059		1311	yararlı	Adj	2667	0.056
1275	bağlamak	Verb	2809	0.059		1312	yapısal	Adj	2664	0.056
1276	aynen	Adv	2804	0.058		1313	burun	Noun	2663	0.055
1277	tıp	Noun	2802	0.058		1314	kahve	Noun	2662	0.055
1278	çatışma	Noun	2785	0.058		1315	radyo	Noun	2659	0.055
1279	değişen	Adj	2779	0.058		1316	tatil	Noun	2652	0.055
1280	temsilci	Noun	2771	0.058		1317	olası	Adj	2648	0.055
1281	ödül	Noun	2770	0.058		1318	reklam	Noun	2646	0.055
1282	kart	Noun	2762	0.058		1319	nükleer	Adj	2642	0.055
1283	daire	Noun	2761	0.058		1320	ulus	Noun	2637	0.055
1284	saray	Noun	2758	0.057		1321	karmaşık	Adj	2635	0.055
1285	başlık	Noun	2748	0.057		1322	trafik	Noun	2633	0.055
1286	intihar	Noun	2748	0.057		1323	iri	Adj	2630	0.055
1287	yeryüzü	Noun	2745	0.057		1324	karşılaşmak	Verb	2630	0.055
1288	kimyasal	Adj	2744	0.057		1325	mutlak	Adj	2629	0.055
1289	rejim	Noun	2744	0.057		1326	önlemek	Verb	2629	0.055
1290	nesne	Noun	2742	0.057		1327	hazırlık	Noun	2628	0.055

1328	göç	Noun	2620	0.055		1365	kulüp	Noun	2511	0.052
1329	demokrat	Adj	2613	0.054		1366	tanımlamak	Verb	2509	0.052
1330	hat	Noun	2612	0.054		1367	müthiş	Adj	2505	0.052
1331	finansal	Adj	2603	0.054		1368	yetersiz	Adj	2505	0.052
1332	saf	Noun	2602	0.054		1369	aksi	Adj	2504	0.052
1333	basmak	Verb	2599	0.054		1370	döner	Noun	2496	0.052
1334	gökyüzü	Noun	2599	0.054		1371	beklenti	Noun	2490	0.052
1335	kanal	Noun	2586	0.054		1372	korkmak	Verb	2489	0.052
1336	strateji	Noun	2584	0.054		1373	taze	Noun	2486	0.052
1337	turizm	Noun	2584	0.054		1374	ilim	Noun	2484	0.052
1338	yahudi	Noun	2584	0.054		1375	halbuki	Conj	2483	0.052
1339	yıllarca	Adv	2584	0.054		1376	nefret	Noun	2483	0.052
1340	duygusal	Adj	2580	0.054		1377	organ	Noun	2480	0.052
1341	hıristiyan	Noun	2577	0.054		1378	dans	Noun	2477	0.052
1342	muhalefet	Noun	2576	0.054		1379	isimli	Adj	2477	0.052
1343	ikna	Noun	2574	0.054		1380	yumurta	Noun	2475	0.052
1344	oluşum	Noun	2571	0.054		1381	hacı	Noun	2473	0.052
1345	büyüklük	Noun	2568	0.054		1382	yorgun	Adj	2465	0.051
1346	keyif	Noun	2566	0.053		1383	çıplak	Adj	2464	0.051
1347	hukuki	Adj	2565	0.053		1384	düş	Noun	2463	0.051
1348	civar	Noun	2562	0.053		1385	tayin	Noun	2461	0.051
1349	fayda	Noun	2554	0.053		1386	bağımsızlık	Noun	2459	0.051
1350	paralel	Adj	2550	0.053		1387	toz	Noun	2458	0.051
1351	tesis	Noun	2550	0.053		1388	kahraman	Noun	2457	0.051
1352	dönük	Adj	2545	0.053		1389	vadeli	Adj	2451	0.051
1353	yüce	Adj	2541	0.053		1390	dip	Noun	2446	0.051
1354	pratik	Noun	2538	0.053		1391	koyu	Adj	2439	0.051
1355	paket	Noun	2537	0.053		1392	verici	Noun	2432	0.051
1356	vali	Noun	2536	0.053		1393	verimli	Adj	2432	0.051
1357	planlama	Noun	2534	0.053		1394	gayet	Adv	2422	0.050
1358	abla	Noun	2531	0.053		1395	hazreti	Noun	2420	0.050
1359	gittikçe	Adv	2529	0.053		1396	akıllı	Adj	2412	0.050
1360	çevirmek	Verb	2527	0.053		1397	meyve	Noun	2408	0.050
1361	karışmak	Verb	2527	0.053		1398	alternatif	Noun	2406	0.050
1362	dürüst	Adj	2523	0.053		1399	bildirmek	Verb	2403	0.050
1363	bedel	Noun	2522	0.053		1400	tarif	Noun	2401	0.050
1364	peygamber	Noun	2520	0.053		1401	yararlanmak	Verb	2394	0.050

1402	kapitalizm	Noun	2392	0.050	1439	gerçeklik	Noun	2284	0.048
1403	ayrıntılı	Adj	2390	0.050	1440	birleşmek	Verb	2280	0.048
1404	buyurmak	Verb	2386	0.050	1441	gösterge	Noun	2278	0.047
1405	içki	Noun	2376	0.050	1442	kurumsal	Adj	2278	0.047
1406	kaza	Noun	2371	0.049	1443	inşaat	Noun	2276	0.047
1407	netice	Noun	2370	0.049	1444	eklemek	Verb	2273	0.047
1408	çil	Noun	2366	0.049	1445	öfke	Noun	2273	0.047
1409	vakıf	Noun	2366	0.049	1446	etik	Noun	2268	0.047
1410	anahtar	Noun	2361	0.049	1447	sessizlik	Noun	2259	0.047
1411	nice	Adj	2357	0.049	1448	temin	Noun	2254	0.047
1412	sır	Noun	2352	0.049	1449	bilinçli	Adj	2252	0.047
1413	salt	Adv	2347	0.049	1450	batılı	Adj	2239	0.047
1414	karşıt	Adj	2346	0.049	1451	mükemmel	Adj	2239	0.047
1415	nedense	Postp	2342	0.049	1452	yönetmen	Noun	2235	0.047
1416	stratejik	Adj	2338	0.049	1453	evren	Noun	2232	0.047
1417	padişah	Noun	2332	0.049	1454	kedi	Noun	2231	0.046
1418	yüzey	Noun	2332	0.049	1455	temas	Noun	2230	0.046
1419	dikkatli	Adj	2327	0.048	1456	başlıklı	Adj	2228	0.046
1420	gök	Noun	2327	0.048	1457	tepe	Noun	2228	0.046
1421	ihmal	Noun	2327	0.048	1458	dağılım	Noun	2226	0.046
1422	ağrı	Noun	2326	0.048	1459	pazarlama	Noun	2224	0.046
1423	bahar	Noun	2326	0.048	1460	sekreter	Noun	2219	0.046
1424	gölge	Noun	2325	0.048	1461	refah	Noun	2217	0.046
1425	zavallı	Adj	2323	0.048	1462	buz	Noun	2216	0.046
1426	özür	Noun	2322	0.048	1463	eleman	Noun	2216	0.046
1427	girişim	Noun	2321	0.048	1464	gözlemek	Verb	2213	0.046
1428	mısır	Noun	2316	0.048	1465	hatırlamak	Verb	2210	0.046
1429	bölgesel	Adj	2315	0.048	1466	mezun	Adj	2210	0.046
1430	zeki	Noun	2313	0.048	1467	ahlaki	Adj	2206	0.046
1431	hüzün	Noun	2311	0.048	1468	kök	Noun	2201	0.046
1432	mantık	Noun	2306	0.048	1469	esna	Noun	2196	0.046
1433	saniye	Noun	2305	0.048	1470	belirsiz	Adj	2194	0.046
1434	tedbir	Noun	2302	0.048	1471	düz	Adj	2194	0.046
1435	dönüş	Noun	2299	0.048	1472	ümit	Noun	2194	0.046
1436	çap	Noun	2295	0.048	1473	otobüs	Noun	2193	0.046
1437	anında	Noun	2287	0.048	1474	akım	Noun	2192	0.046
1438	aziz	Adj	2284	0.048	1475	verimlilik	Noun	2190	0.046

1476	etnik	Adj	2185	0.046	1513	devre	Noun	2102	0.044
1477	yavru	Noun	2179	0.045	1514	matematik	Noun	2102	0.044
1478	okur	Noun	2178	0.045	1515	dolaylı	Adj	2100	0.044
1479	birdenbire	Adv	2174	0.045	1516	tahta	Noun	2100	0.044
1480	asit	Noun	2172	0.045	1517	birşey	Pron	2099	0.044
1481	kapitalist	Noun	2172	0.045	1518	kolaylık	Adj	2095	0.044
1482	sarılmak	Verb	2172	0.045	1519	çekici	Adj	2091	0.044
1483	bozmak	Verb	2169	0.045	1520	bankacılık	Noun	2090	0.044
1484	puan	Noun	2169	0.045	1521	bütünlük	Adj	2087	0.043
1485	kökenli	Noun	2168	0.045	1522	komşu	Noun	2084	0.043
1486	zevk	Noun	2166	0.045	1523	etken	Noun	2083	0.043
1487	çözmek	Verb	2162	0.045	1524	yoksun	Adj	2082	0.043
1488	sinir	Noun	2162	0.045	1525	organik	Adj	2076	0.043
1489	vatan	Noun	2158	0.045	1526	yaratıcı	Adj	2076	0.043
1490	eleştiri	Noun	2155	0.045	1527	gündüz	Noun	2075	0.043
1491	ağ	Noun	2152	0.045	1528	hazine	Noun	2072	0.043
1492	biyolojik	Adj	2150	0.045	1529	yıldırım	Noun	2072	0.043
1493	hapis	Noun	2150	0.045	1530	besin	Noun	2070	0.043
1494	kilo	Noun	2150	0.045	1531	evvel	Postp	2069	0.043
1495	alev	Noun	2149	0.045	1532	ilçe	Noun	2067	0.043
1496	performans	Noun	2149	0.045	1533	egemenlik	Adj	2063	0.043
1497	inanılmaz	Adj	2142	0.045	1534	tanık	Noun	2059	0.043
1498	dinsel	Adj	2141	0.045	1535	ekran	Noun	2050	0.043
1499	ziyade	Adj	2139	0.045	1536	şimdilik	Adv	2049	0.043
1500	çelik	Noun	2137	0.045	1537	eşitlik	Adj	2048	0.043
1501	direktör	Noun	2130	0.044	1538	ideal	Noun	2044	0.043
1502	kuvvetli	Adj	2130	0.044	1539	üyelik	Noun	2042	0.043
1503	uyumak	Verb	2130	0.044	1540	yahut	Conj	2034	0.042
1504	alkol	Noun	2129	0.044	1541	yalnızlık	Adj	2033	0.042
1505	yoksul	Noun	2127	0.044	1542	kör	Adj	2032	0.042
1506	oluş	Noun	2124	0.044	1543	başkent	Noun	2028	0.042
1507	darbe	Noun	2118	0.044	1544	fon	Noun	2026	0.042
1508	tarımsal	Adj	2111	0.044	1545	kurtulmak	Verb	2020	0.042
1509	itiraz	Noun	2110	0.044	1546	pozitif	Adj	2015	0.042
1510	ithal	Noun	2108	0.044	1547	şiddetli	Adj	2013	0.042
1511	şaşkın	Adj	2108	0.044	1548	şikayet	Noun	2008	0.042
1512	ağırlıklı	Adj	2106	0.044	1549	onca	Adj	2007	0.042

1550	kararlı	Adj	2005	0.042	1587	uyuşturucu	Noun	1928	0.040
1551	kilise	Noun	2002	0.042	1588	derhal	Adv	1925	0.040
1552	gereksinim	Noun	1998	0.042	1589	konut	Noun	1925	0.040
1553	uçmak	Verb	1996	0.042	1590	ötürü	Postp	1924	0.040
1554	yetmek	Verb	1995	0.042	1591	sessizce	Adv	1922	0.040
1555	hediye	Noun	1993	0.042	1592	karışık	Adj	1921	0.040
1556	yetiştirmek	Verb	1993	0.042	1593	maalesef	Adv	1921	0.040
1557	yürürlük	Noun	1990	0.041	1594	katmak	Verb	1919	0.040
1558	yiyecek	Noun	1989	0.041	1595	karakter	Noun	1917	0.040
1559	takılmak	Verb	1984	0.041	1596	ideoloji	Noun	1915	0.040
1560	ihraç	Noun	1979	0.041	1597	sınav	Noun	1915	0.040
1561	mekanizma	Noun	1978	0.041	1598	ikili	Adj	1913	0.040
1562	kanat	Noun	1975	0.041	1599	sıvı	Adj	1913	0.040
1563	dik	Adj	1974	0.041	1600	öylece	Adv	1912	0.040
1564	istifa	Noun	1968	0.041	1601	birikim	Noun	1903	0.040
1565	futbolcu	Noun	1964	0.041	1602	şüphesiz	Adv	1903	0.040
1566	maruz	Adj	1963	0.041	1603	suçlu	Adj	1900	0.040
1567	kalıcı	Adj	1961	0.041	1604	mesleki	Adj	1898	0.040
1568	komünist	Noun	1961	0.041	1605	burjuva	Adj	1895	0.039
1569	dudak	Noun	1960	0.041	1606	keskin	Noun	1895	0.039
1570	istikrar	Noun	1960	0.041	1607	rakip	Adj	1893	0.039
1571	maliye	Noun	1955	0.041	1608	gereksiz	Adj	1892	0.039
1572	ton	Noun	1954	0.041	1609	yön	Noun	1891	0.039
1573	kapsamlı	Adj	1953	0.041	1610	uymak	Verb	1890	0.039
1574	kısmen	Adv	1950	0.041	1611	kapasite	Noun	1889	0.039
1575	zararlı	Adj	1950	0.041	1612	küçücük	Adj	1889	0.039
1576	estetik	Noun	1947	0.041	1613	zafer	Noun	1887	0.039
1577	pırıl	Adj	1947	0.041	1614	fonksiyon	Noun	1886	0.039
1578	kurtarmak	Verb	1944	0.041	1615	uyumlu	Adj	1886	0.039
1579	aşmak	Verb	1942	0.040	1616	haksız	Adj	1885	0.039
1580	operasyon	Noun	1940	0.040	1617	başlatmak	Verb	1884	0.039
1581	uğruna	Noun	1937	0.040	1618	kongre	Noun	1880	0.039
1582	şaşırmak	Verb	1936	0.040	1619	cezaevi	Noun	1879	0.039
1583	alışveriş	Noun	1935	0.040	1620	laik	Adj	1878	0.039
1584	lig	Noun	1930	0.040	1621	uzatmak	Verb	1878	0.039
1585	imalat	Noun	1929	0.040	1622	cuma	Noun	1876	0.039
1586	rüya	Noun	1928	0.040	1623	sahte	Adj	1876	0.039

1624	içten	Adj	1869	0.039	1661	yanısıra	Adv	1818	0.038
1625	kaba	Adj	1869	0.039	1662	zorunluluk	Adj	1813	0.038
1626	ayrım	Noun	1868	0.039	1663	insanoğlu	Noun	1812	0.038
1627	pembe	Noun	1867	0.039	1664	teşkilat	Noun	1811	0.038
1628	aslan	Noun	1866	0.039	1665	kirli	Adj	1810	0.038
1629	özgün	Adj	1865	0.039	1666	basınç	Noun	1805	0.038
1630	tasarım	Noun	1859	0.039	1667	bayrak	Noun	1802	0.038
1631	ruhsal	Adj	1856	0.039	1668	kırık	Noun	1801	0.038
1632	istihbarat	Noun	1853	0.039	1669	işkence	Noun	1800	0.038
1633	ödemek	Verb	1853	0.039	1670	kırsal	Adj	1800	0.038
1634	razı	Adj	1853	0.039	1671	çare	Noun	1796	0.037
1635	kuşak	Noun	1852	0.039	1672	hariç	Noun	1796	0.037
1636	hikmet	Noun	1851	0.039	1673	soruşturma	Noun	1796	0.037
1637	numara	Noun	1851	0.039	1674	olanaksız	Adj	1793	0.037
1638	ilahi	Adj	1850	0.039	1675	şarap	Noun	1790	0.037
1639	ayırmak	Verb	1847	0.038	1676	tavsiye	Noun	1789	0.037
1640	pahalı	Adj	1844	0.038	1677	güvenli	Adj	1788	0.037
1641	çocukluk	Noun	1843	0.038	1678	otomobil	Noun	1787	0.037
1642	plastik	Noun	1843	0.038	1679	popüler	Adj	1787	0.037
1643	selam	Noun	1843	0.038	1680	genelkurmay	Noun	1784	0.037
1644	liberal	Adj	1841	0.038	1681	fakülte	Noun	1783	0.037
1645	fena	Adj	1838	0.038	1682	cesaret	Noun	1781	0.037
1646	tasarruf	Noun	1835	0.038	1683	birleşik	Adj	1779	0.037
1647	şampiyon	Noun	1834	0.038	1684	sırf	Adv	1779	0.037
1648	yayımlamak	Verb	1833	0.038	1685	bulgu	Noun	1775	0.037
1649	açlık	Noun	1831	0.038	1686	güçlük	Noun	1773	0.037
1650	gerçekçi	Adj	1831	0.038	1687	eğlence	Noun	1772	0.037
1651	merdiven	Noun	1831	0.038	1688	hassas	Adj	1769	0.037
1652	meydan	Noun	1829	0.038	1689	gayri	Adv	1767	0.037
1653	kaya	Noun	1828	0.038	1690	amaçlı	Adj	1766	0.037
1654	heyet	Noun	1826	0.038	1691	evlat	Noun	1766	0.037
1655	alay	Noun	1824	0.038	1692	hayret	Noun	1765	0.037
1656	reform	Noun	1823	0.038	1693	epey	Adv	1764	0.037
1657	ithalat	Noun	1819	0.038	1694	soyut	Adj	1764	0.037
1658	susmak	Verb	1819	0.038	1695	düzgün	Adj	1763	0.037
1659	süreli	Adj	1819	0.038	1696	hafifçe	Adv	1763	0.037
1660	terim	Noun	1818	0.038	1697	gönüllü	Adj	1756	0.037

1698	tren	Noun	1755	0.037	1735	bilhassa	Adv	1703	0.035
1699	boşuna	Adv	1749	0.036	1736	çorba	Noun	1703	0.035
1700	ağa	Noun	1746	0.036	1737	finans	Noun	1703	0.035
1701	muhteşem	Adj	1746	0.036	1738	melek	Noun	1700	0.035
1702	marka	Noun	1744	0.036	1739	ekip	Noun	1697	0.035
1703	itiraf	Noun	1743	0.036	1740	kiraz	Noun	1697	0.035
1704	rakam	Noun	1743	0.036	1741	hayali	Adj	1692	0.035
1705	kurucu	Noun	1742	0.036	1742	taban	Noun	1690	0.035
1706	göğüs	Noun	1740	0.036	1743	evli	Adj	1689	0.035
1707	kanlı	Noun	1738	0.036	1744	duyarlı	Adj	1688	0.035
1708	ortaklık	Noun	1738	0.036	1745	alıcı	Noun	1687	0.035
1709	yarar	Noun	1738	0.036	1746	yüklü	Adj	1687	0.035
1710	tatmin	Noun	1735	0.036	1747	buçuk	Adj	1682	0.035
1711	arazi	Noun	1734	0.036	1748	çığlık	Noun	1682	0.035
1712	gözyaşı	Noun	1734	0.036	1749	vazgeçmek	Verb	1682	0.035
1713	kapak	Noun	1731	0.036	1750	sadık	Adj	1681	0.035
1714	şeref	Noun	1731	0.036	1751	disiplin	Noun	1680	0.035
1715	gümüş	Noun	1730	0.036	1752	tur	Noun	1677	0.035
1716	başlı	Postp	1728	0.036	1753	sosyalizm	Noun	1676	0.035
1717	yerleşim	Noun	1727	0.036	1754	araştırmacı	Noun	1674	0.035
1718	hürriyet	Noun	1725	0.036	1755	çirkin	Adj	1674	0.035
1719	şoför	Noun	1721	0.036	1756	insani	Adj	1674	0.035
1720	sergi	Noun	1720	0.036	1757	iptal	Noun	1674	0.035
1721	teker	Noun	1720	0.036	1758	deri	Noun	1672	0.035
1722	çetin	Noun	1719	0.036	1759	şimdiki	Adj	1667	0.035
1723	gol	Noun	1719	0.036	1760	gösteri	Noun	1657	0.035
1724	teori	Noun	1716	0.036	1761	motor	Noun	1656	0.035
1725	bulut	Noun	1714	0.036	1762	ihale	Noun	1655	0.034
1726	ani	Adj	1711	0.036	1763	isyan	Noun	1655	0.034
1727	ayna	Noun	1710	0.036	1764	parasal	Adj	1654	0.034
1728	eksiklik	Noun	1708	0.036	1765	imkansız	Adj	1651	0.034
1729	tamamlamak	Verb	1708	0.036	1766	boylu	Adj	1648	0.034
1730	tutar	Noun	1707	0.036	1767	varolmak	Verb	1647	0.034
1731	adil	Adj	1706	0.036	1768	işadamı	Noun	1645	0.034
1732	uyanmak	Verb	1706	0.036	1769	müdürlük	Noun	1644	0.034
1733	yetişmek	Verb	1706	0.036	1770	özlem	Noun	1644	0.034
1734	kritik	Adj	1705	0.036	1771	fabrika	Noun	1637	0.034

143

1772	giysi	Noun	1634	0.034	1809	otomatik	Adj	1580	0.033
1773	imam	Noun	1634	0.034	1810	metal	Noun	1578	0.033
1774	ayrıntı	Noun	1633	0.034	1811	sezon	Noun	1578	0.033
1775	denk	Noun	1633	0.034	1812	apartman	Noun	1576	0.033
1776	acil	Adj	1632	0.034	1813	göl	Noun	1575	0.033
1777	park	Noun	1630	0.034	1814	kaygı	Noun	1574	0.033
1778	mesut	Noun	1629	0.034	1815	güvenilir	Adj	1573	0.033
1779	teorik	Adj	1628	0.034	1816	değin	Postp	1572	0.033
1780	hitap	Noun	1625	0.034	1817	meşhur	Noun	1570	0.033
1781	niyet	Noun	1624	0.034	1818	aydınlık	Adj	1569	0.033
1782	radikal	Adj	1621	0.034	1819	kurşun	Noun	1566	0.033
1783	genetik	Noun	1612	0.034	1820	imar	Noun	1565	0.033
1784	aniden	Adv	1611	0.034	1821	çağrı	Noun	1562	0.033
1785	anmak	Verb	1608	0.034	1822	şeyh	Noun	1562	0.033
1786	sözkonusu	Adj	1605	0.033	1823	akademik	Adj	1561	0.033
1787	koç	Noun	1603	0.033	1824	heykel	Noun	1560	0.033
1788	celal	Noun	1602	0.033	1825	ihtiyar	Noun	1559	0.032
1789	federal	Adj	1602	0.033	1826	madem	Conj	1557	0.032
1790	rica	Noun	1601	0.033	1827	profesyonel	Adj	1557	0.032
1791	birara	Adj	1600	0.033	1828	faydalı	Adj	1556	0.032
1792	öğe	Noun	1600	0.033	1829	kaptan	Noun	1556	0.032
1793	bağımlı	Adj	1598	0.033	1830	masum	Adj	1555	0.032
1794	herşey	Pron	1598	0.033	1831	acele	Adj	1554	0.032
1795	sallamak	Verb	1598	0.033	1832	ressam	Noun	1554	0.032
1796	yapay	Adj	1594	0.033	1833	organize	Adj	1553	0.032
1797	fizik	Noun	1593	0.033	1834	mühendislik	Noun	1552	0.032
1798	sabır	Noun	1592	0.033	1835	coğrafi	Adj	1550	0.032
1799	dilim	Noun	1591	0.033	1836	nur	Noun	1549	0.032
1800	gazi	Noun	1590	0.033	1837	delikanlı	Noun	1548	0.032
1801	alım	Noun	1589	0.033	1838	katlı	Noun	1548	0.032
1802	profesör	Noun	1586	0.033	1839	bozuk	Adj	1547	0.032
1803	turist	Noun	1586	0.033	1840	deneyim	Noun	1546	0.032
1804	sandık	Noun	1585	0.033	1841	takdir	Noun	1546	0.032
1805	cephe	Noun	1584	0.033	1842	şef	Noun	1545	0.032
1806	mobil	Adj	1584	0.033	1843	moda	Noun	1543	0.032
1807	beton	Noun	1581	0.033	1844	damla	Noun	1542	0.032
1808	laiklik	Adj	1581	0.033	1845	tipik	Adj	1540	0.032

1846	kamusal	Adj	1539	0.032	1883	ilkel	Adj	1488	0.031
1847	kucak	Noun	1539	0.032	1884	dişi	Noun	1486	0.031
1848	şaşkınlık	Adj	1538	0.032	1885	protesto	Noun	1486	0.031
1849	anlamsız	Adj	1537	0.032	1886	zorluk	Noun	1486	0.031
1850	dostluk	Noun	1534	0.032	1887	dağıtım	Noun	1485	0.031
1851	kanser	Noun	1534	0.032	1888	muhafaza	Noun	1484	0.031
1852	minik	Adj	1534	0.032	1889	söylem	Noun	1483	0.031
1853	ısı	Noun	1533	0.032	1890	yatay	Adj	1482	0.031
1854	dinamik	Noun	1532	0.032	1891	hisse	Noun	1478	0.031
1855	şehit	Noun	1532	0.032	1892	rahatlıkla	Adv	1478	0.031
1856	beslenme	Noun	1531	0.032	1893	konferans	Noun	1474	0.031
1857	çarpıcı	Verb	1531	0.032	1894	örtülü	Adj	1474	0.031
1858	jandarma	Noun	1530	0.032	1895	perde	Noun	1474	0.031
1859	kilometre	Noun	1527	0.032	1896	fakir	Noun	1473	0.031
1860	yayınlamak	Verb	1527	0.032	1897	içişleri	Noun	1472	0.031
1861	üstünlük	Noun	1525	0.032	1898	şükür	Noun	1472	0.031
1862	uğur	Noun	1521	0.032	1899	adlandırmak	Verb	1471	0.031
1863	lan	Slang	1520	0.032	1900	yangın	Noun	1471	0.031
1864	meğer	Conj	1520	0.032	1901	video	Noun	1470	0.031
1865	general	Noun	1518	0.032	1902	çatı	Noun	1465	0.031
1866	kaliteli	Adj	1514	0.032	1903	nitelikli	Adj	1465	0.031
1867	teyze	Noun	1514	0.032	1904	evlenmek	Verb	1461	0.030
1868	çaresiz	Adj	1513	0.032	1905	görünüm	Noun	1457	0.030
1869	samimi	Adj	1510	0.031	1906	onay	Noun	1456	0.030
1870	kampanya	Noun	1509	0.031	1907	sendika	Noun	1456	0.030
1871	federasyon	Noun	1504	0.031	1908	tören	Noun	1456	0.030
1872	uzay	Noun	1504	0.031	1909	boya	Noun	1453	0.030
1873	gayret	Noun	1503	0.031	1910	liste	Noun	1452	0.030
1874	sarmak	Verb	1499	0.031	1911	emperyalist	Adj	1449	0.030
1875	gövde	Noun	1498	0.031	1912	ayıp	Noun	1446	0.030
1876	zihinsel	Adj	1498	0.031	1913	zihin	Noun	1446	0.030
1877	engellemek	Verb	1496	0.031	1914	ilkokul	Noun	1445	0.030
1878	asgari	Adj	1495	0.031	1915	telaş	Noun	1445	0.030
1879	konmak	Verb	1495	0.031	1916	alışkanlık	Noun	1444	0.030
1880	kader	Noun	1490	0.031	1917	altyapı	Noun	1439	0.030
1881	yuvarlak	Adj	1490	0.031	1918	negatif	Noun	1438	0.030
1882	yolcu	Noun	1489	0.031	1919	güncel	Adj	1434	0.030

1920	felsefi	Adj	1433	0.030		1957	resmen	Adv	1389	0.029
1921	defter	Noun	1431	0.030		1958	garanti	Noun	1388	0.029
1922	yazılım	Noun	1428	0.030		1959	inkar	Noun	1387	0.029
1923	belirmek	Verb	1426	0.030		1960	meraklı	Adj	1387	0.029
1924	hacim	Noun	1426	0.030		1961	yoğunluk	Adj	1386	0.029
1925	konuk	Noun	1426	0.030		1962	mühendis	Noun	1385	0.029
1926	lüks	Noun	1423	0.030		1963	sayısal	Adj	1385	0.029
1927	nesnel	Adj	1423	0.030		1964	zorlu	Adj	1383	0.029
1928	sözde	Noun	1423	0.030		1965	yepyeni	Adj	1382	0.029
1929	sebze	Noun	1420	0.030		1966	satır	Noun	1378	0.029
1930	şahıs	Noun	1418	0.030		1967	esmek	Verb	1375	0.029
1931	iktisat	Noun	1413	0.029		1968	gündelik	Adj	1375	0.029
1932	azaltmak	Verb	1412	0.029		1969	menkul	Adj	1375	0.029
1933	cumartesi	Noun	1412	0.029		1970	kuşku	Noun	1374	0.029
1934	deneme	Noun	1411	0.029		1971	pamuk	Noun	1374	0.029
1935	sevimli	Adj	1411	0.029		1972	kum	Noun	1373	0.029
1936	oksijen	Noun	1405	0.029		1973	sevk	Noun	1372	0.029
1937	sayısız	Adj	1405	0.029		1974	sulama	Noun	1372	0.029
1938	süper	Adj	1404	0.029		1975	harp	Noun	1367	0.028
1939	tuvalet	Noun	1403	0.029		1976	komik	Noun	1365	0.028
1940	özelleştirme	Noun	1400	0.029		1977	endişe	Noun	1363	0.028
1941	yavaşça	Adv	1400	0.029		1978	kahverengi	Adj	1363	0.028
1942	finansman	Noun	1399	0.029		1979	masal	Noun	1363	0.028
1943	kusur	Noun	1398	0.029		1980	protein	Noun	1363	0.028
1944	güneydoğu	Noun	1397	0.029		1981	kızıl	Adj	1358	0.028
1945	iyilik	Noun	1396	0.029		1982	üretici	Noun	1356	0.028
1946	bay	Noun	1395	0.029		1983	harika	Adj	1354	0.028
1947	iklim	Noun	1395	0.029		1984	medeni	Adj	1352	0.028
1948	numaralı	Adj	1395	0.029		1985	zeka	Noun	1352	0.028
1949	bacak	Noun	1393	0.029		1986	ahşap	Adj	1351	0.028
1950	icra	Noun	1392	0.029		1987	albüm	Noun	1350	0.028
1951	yara	Noun	1392	0.029		1988	boru	Noun	1350	0.028
1952	inşallah	Interj	1391	0.029		1989	has	Adj	1350	0.028
1953	görsel	Adj	1390	0.029		1990	yaralı	Adj	1349	0.028
1954	köklü	Adj	1390	0.029		1991	ipek	Adj	1347	0.028
1955	beyefendi	Noun	1389	0.029		1992	tedirgin	Adj	1345	0.028
1956	iman	Noun	1389	0.029		1993	büyükşehir	Noun	1344	0.028

1994	sarhoş	Noun	1344	0.028		1999	şeytan	Noun	1338	0.028
1995	burjuvazi	Noun	1342	0.028		2000	gözaltı	Noun	1337	0.028
1996	gurur	Noun	1342	0.028						
1997	seçenek	Noun	1342	0.028						
1998	boyutlu	Adj	1339	0.028						

Appendix 11 The First 100 Words According to Part of Speech (The General Turkish Wordlist)

	Verbs	Nouns	Adjectives	Adverbs	Others
1	olmak	iç	son	çok	ve
2	etmek	yer	büyük	daha	bu
3	yapmak	yıl	böyle	en	o
4	demek	ara	yeni	hiç	için
5	almak	zaman	ilk	bile	ben
6	vermek	insan	iyi	nasıl	gibi
7	gelmek	var	önemli	şimdi	kendi
8	istemek	şey	bütün	artık	ne
9	görmek	gün	aynı	yine	ile
10	çıkmak	el	başka	öyle	sonra
11	gitmek	iş	doğru	sadece	her
12	bilmek	yan	ön	hemen	ama
13	kalmak	üzeri	diğer	belki	kadar
14	söylemek	konu	alt	geri	değil
15	başlamak	baş	dış	hep	biz
16	bakmak	yol	tek	pek	ya
17	bulunmak	yok	bazı	özellikle	ise
18	bulmak	ülke	genel	yeniden	sen
19	kullanmak	neden	fazla	şöyle	önce
20	oluşmak	çocuk	üst	ayrıca	diye
21	göstermek	dünya	az	zaten	karşı
22	getirmek	kadın	küçük	evet	göre
23	geçmek	orta	güzel	aslında	biri
24	yaşamak	durum	uzun	sanki	ancak
25	girmek	göz	tam	hala	siz
26	gerekmek	hal	biraz	sürekli	veya
27	düşünmek	alan	yüksek	sık	hem
28	açmak	ev	genç	gerçekten	şu
29	çalışmak	söz	açık	tekrar	birlikte
30	varmak	şekil	eski	yalnızca	çünkü
31	sağlamak	devlet	özel	bazen	tüm
32	durmak	kişi	belli	peki	tarafından
33	kurmak	bura	mümkün	birden	yani
34	dönmek	dönem	uygun	henüz	ilgili
35	anlamak	su	ekonomik	elbette	kim
36	konuşmak	sonuç	sosyal	genellikle	hiçbir
37	belirtmek	sahip	geçen	üstelik	birbiri

148

38	çekmek	adam	kısa	iyice	işte
39	anlatmak	sıra	zor	acaba	üzere
40	yazmak	hayat	çeşitli	gene	fakat
41	bırakmak	çalışma	hangi	tamamen	gerek
42	beklemek	kız	uzak	neredeyse	hatta
43	düşmek	tarih	türlü	niye	hepsi
44	atmak	ad	ayrı	mutlaka	birkaç
45	çıkarmak	baba	toplumsal	erken	tabi
46	koymak	anne	kötü	asla	olsa
47	tutmak	bilgi	uluslararası	herhalde	bağlı
48	gerekmek	ses	gerekli	adeta	birçok
49	duymak	gelen	yalnız	kısaca	eğer
50	okumak	sorun	benzer	nitekim	nerede
51	oturmak	karar	geçmiş	iken	hakkında
52	sormak	örnek	siyasi	açıkça	rağmen
53	taşımak	sistem	milli	niçin	ait
54	artmak	ilişki	ağır	galiba	ilişkin
55	sevmek	süre	geniş	kesinlikle	oysa
56	araştırma	para	doğal	neyse	yoksa
57	gelişme	toplum	kolay	yeterince	dolayısı
58	kazanmak	güç	siyasal	sonradan	yönelik
59	belirlemek	yüz	zorunda	eskiden	boyunca
60	sürmek	an	herhangi	lütfen	bakımından
61	aramak	ay	dolu	çoktan	yüzden
62	yaratmak	olay	yanlış	aksine	beri
63	bitmek	dil	teknik	daima	dolayı
64	geçirmek	eğitim	beraber	halen	itibaren
65	gerçekleşmek	hareket	asıl	bizzat	karşın
66	tanımak	hak	ciddi	gerçi	dahil
67	öğrenmek	halk	adlı	kolayca	saye
68	kalkmak	kabul	yavaş	aynen	hele
69	katılmak	kez	beyaz	hayli	mesela
70	görünmek	devam	yaklaşık	yıllarca	yüzünden
71	büyümek	farklı	yıllık	gittikçe	dahi
72	ayrılmak	gerçek	güçlü	gayet	dek
73	hazırlamak	bugün	değişik	salt	dair
74	izlemek	bölge	askeri	birdenbire	zira
75	yürümek	gece	ulusal	şimdilik	dayalı
76	ulaşmak	kapı	düşük	kısmen	gerekse
77	ölmek	ora	mevcut	derhal	falan

78	uygulamak	biçim	derin	sessizce	takdirde
79	düzenlemek	saat	boş	maalesef	filan
80	dolmak	yaşam	hızlı	öylece	haydi
81	inanmak	grup	kaç	şüphesiz	ibaret
82	oynamak	kitap	resmi	yanısıra	halbuki
83	doğmak	değer	ters	sırf	nedense
84	tartışma	arkadaş	başarılı	gayri	birşey
85	açıklamak	kimse	yeterli	epey	evvel
86	korumak	erkek	kara	hafifçe	yahut
87	değişmek	herkes	doğrudan	boşuna	ötürü
88	toplamak	sayı	mutlu	bilhassa	başlı
89	sanmak	yapı	ince	aniden	herşey
90	satmak	oran	ünlü	rahatlıkla	değin
91	geliştirmek	parti	etkili	yavaşça	madem
92	idi	aile	belirli	resmen	lan
93	içermek	temel	sıcak		meğer
94	değerlendirmek	tür	hoş		inşallah
95	dinlemek	yakın	yerel		
96	kesmek	yaş	yaşlı		
97	yükselmek	bey	gizli		
98	dayanmak	gelir	yoğun		
99	savunmak	amaç	ilginç		
100	gelişmek	yönetim	olumlu		

Appendix 12 The Academic Turkish Wordlist (1010 Words)

*in 10 000 words

No	Words (Lemmas)	Type	Freq	%*
1	öğrenmek	Verb	7913	24.00
2	davranış	Noun	7313	22.00
3	belirlemek	Verb	7174	22.00
4	işletme	Noun	7164	22.00
5	birey	Noun	6945	21.00
6	gelişim	Noun	6840	21.00
7	kavram	Noun	6331	19.00
8	sözleşme	Noun	5950	18.00
9	aşağı	Noun	5912	18.00
10	işlem	Noun	5858	18.00
11	oluşturmak	Verb	5806	18.00
12	yanıt	Noun	5635	17.00
13	kanun	Noun	5580	17.00
14	geçirmek	Verb	5380	16.00
15	kurum	Noun	5148	16.00
16	açıklamak	Verb	5091	16.00
17	uygulama	Noun	4997	15.00
18	varlık	Noun	4927	15.00
19	araç	Noun	4867	15.00
20	uygulamak	Verb	4804	15.00
21	değerlendirmek	Verb	4754	14.00
22	ilişkin	Postp	4736	14.00
23	suç	Noun	4718	14.00
24	meslek	Noun	4640	14.00
25	iletişim	Noun	4495	14.00
26	mal	Noun	4336	13.00
27	toplumsal	Adj	4190	13.00
28	ilke	Noun	4173	13.00
29	kural	Noun	4079	12.00
30	gerçekleştirmek	Verb	4071	12.00
31	ortam	Noun	4068	12.00
32	etkinlik	Noun	4046	12.00
33	tanımlamak	Verb	3878	12.00
34	geliştirmek	Verb	3829	12.00
35	hüküm	Noun	3826	12.00
36	koşul	Noun	3821	12.00
37	artmak	Verb	3812	11.00
38	politika	Noun	3783	11.00
39	nitelik	Noun	3770	11.00
40	sınıf	Noun	3704	11.00
41	yöntem	Noun	3669	11.00
42	kurulmak	Verb	3584	11.00
43	yapılanmak	Verb	3517	11.00
44	öğretim	Noun	3508	11.00
45	faaliyet	Noun	3504	11.00
46	yaklaşım	Noun	3472	10.00
47	ulaşmak	Verb	3423	10.00
48	talep	Noun	3412	10.00
49	korumak	Verb	3408	10.00
50	ürün	Noun	3391	10.00
51	sermaye	Noun	3274	10.00
52	gerçekleşmek	Verb	3259	10.00
53	yönelik	Postp	3221	10.00
54	açısından	Postp	3148	9.00
55	ekonomi	Noun	3147	9.00
56	düzenlemek	Verb	3109	9.00
57	kuram	Noun	3088	9.00
58	araştırma	Noun	3087	9.00
59	kültür	Noun	3062	9.00
60	taşımak	Verb	3052	9.00
61	tanımak	Verb	3027	9.00
62	çeşitli	Adj	3025	9.00
63	görüş	Noun	3014	9.00
64	zarar	Noun	2991	9.00
65	hesap	Noun	2973	9.00

66	yaratmak	Verb	2973	9.00		102	teknik	Adj	2452	7.00
67	ihtiyaç	Noun	2969	9.00		103	izlemek	Verb	2449	7.00
68	ceza	Noun	2923	9.00		104	etkili	Adj	2440	7.00
69	değişmek	Verb	2910	9.00		105	değiştirmek	Verb	2429	7.00
70	model	Noun	2904	9.00		106	medya	Noun	2414	7.00
71	hedef	Noun	2874	9.00		107	rol	Noun	2410	7.00
72	gelişme	Noun	2870	9.00		108	aşama	Noun	2403	7.00
73	gerekli	Adj	2859	8.00		109	bilim	Noun	2396	7.00
74	teknoloji	Noun	2853	8.00		110	kuruluş	Noun	2395	7.00
75	kazanmak	Verb	2836	8.00		111	nesne	Noun	2388	7.00
76	toplam	Noun	2832	8.00		112	finansal	Adj	2387	7.00
77	sunmak	Verb	2813	8.00		113	sürmek	Verb	2386	7.00
78	sanat	Noun	2775	8.00		114	ticaret	Noun	2374	7.00
79	miktar	Noun	2770	8.00		115	birim	Noun	2372	7.00
80	yetki	Noun	2770	8.00		116	ortak	Adj	2350	7.00
81	üretmek	Verb	2766	8.00		117	pay	Noun	2346	7.00
82	fiyat	Noun	2760	8.00		118	dolayısıyla	Adv	2334	7.00
83	kamu	Noun	2742	8.00		119	felsefe	Noun	2310	7.00
84	üye	Noun	2730	8.00		120	örgüt	Noun	2307	7.00
85	değişim	Noun	2709	8.00		121	gereksinim	Noun	2301	7.00
86	içermek	Verb	2691	8.00		122	dayanmak	Verb	2278	7.00
87	kullanım	Noun	2686	8.00		123	değişiklik	Noun	2273	7.00
88	beceri	Noun	2650	8.00		124	unsur	Noun	2268	7.00
89	şirket	Noun	2642	8.00		125	içerik	Noun	2241	7.00
90	sorumluluk	Noun	2640	8.00		126	işlemek	Verb	2240	7.00
91	ödemek	Verb	2637	8.00		127	azalmak	Verb	2237	7.00
92	risk	Noun	2630	8.00		128	işlev	Noun	2235	7.00
93	önem	Noun	2623	8.00		129	fonksiyon	Noun	2225	6.00
94	denge	Noun	2592	8.00		130	doğmak	Verb	2213	6.00
95	hız	Noun	2581	8.00		131	plan	Noun	2197	6.00
96	belirli	Adj	2576	8.00		132	sebep	Noun	2194	6.00
97	karşılık	Noun	2559	8.00		133	anlayış	Noun	2174	6.00
98	kontrol	Noun	2537	7.00		134	benzer	Adj	2149	6.00
99	kısım	Noun	2531	7.00		135	seçmek	Verb	2148	6.00
100	etkilemek	Verb	2478	7.00		136	ücret	Noun	2134	6.00
101	incelemek	Verb	2469	7.00		137	yardım	Noun	2130	6.00

138	ait	Postp	2125	6.00		174	planlamak	Verb	1834	5.00
139	doğal	Adj	2105	6.00		175	sigorta	Noun	1824	5.00
140	yürütmek	Verb	2097	6.00		176	devralmak	Verb	1818	5.00
141	yasa	Noun	2097	6.00		177	düzen	Noun	1818	5.00
142	öngörmek	Verb	2097	6.00		178	yargı	Noun	1812	5.00
143	fiziksel	Adj	2093	6.00		179	sabit	Adj	1786	5.00
144	yatırım	Noun	2087	6.00		180	giriş	Noun	1755	5.00
145	artış	Noun	2079	6.00		181	genellikle	Adv	1743	5.00
146	sektör	Noun	2058	6.00		182	devre	Noun	1741	5.00
147	şart	Noun	2045	6.00		183	karşılamak	Verb	1738	5.00
148	faktör	Noun	2028	6.00		184	artırmak	Verb	1736	5.00
149	olanak	Noun	2027	6.00		185	ölçü	Noun	1716	5.00
150	veri	Noun	2026	6.00		186	sağlık	Noun	1709	5.00
151	düzenleme	Noun	2023	6.00		187	doğa	Noun	1703	5.00
152	ileri	Noun	2021	6.00		188	kayıt	Noun	1699	5.00
153	maliyet	Noun	2021	6.00		189	düşük	Adj	1688	5.00
154	fark	Noun	2018	6.00		190	kişilik	Noun	1688	5.00
155	katılmak	Verb	2018	6.00		191	satış	Noun	1675	5.00
156	birlik	Noun	2017	6.00		192	çözüm	Noun	1667	5.00
157	sayılmak	Verb	2016	6.00		193	akım	Noun	1659	5.00
158	sürekli	Adv	2010	6.00		194	strateji	Noun	1658	5.00
159	meydan	Noun	1993	6.00		195	algılamak	Verb	1637	5.00
160	esas	Adj	1978	6.00		196	ayrılmak	Verb	1626	5.00
161	derece	Noun	1970	6.00		197	etkileşim	Noun	1624	5.00
162	yönetici	Noun	1956	6.00		198	boyut	Noun	1617	5.00
163	inanç	Noun	1953	6.00		199	kar	Noun	1614	5.00
164	güvenlik	Noun	1950	6.00		200	problem	Noun	1597	5.00
165	farklılık	Noun	1934	6.00		201	sınır	Noun	1597	5.00
166	kapsam	Noun	1912	6.00		202	gerektirmek	Verb	1593	5.00
167	yararlanmak	Verb	1910	6.00		203	temsil	Noun	1587	4.00
168	doğrudan	Adj	1904	5.00		204	konum	Noun	1586	4.00
169	eylem	Noun	1904	5.00		205	emek	Noun	1585	4.00
170	organ	Noun	1902	5.00		206	firma	Noun	1576	4.00
171	satmak	Verb	1879	5.00		207	ilgi	Noun	1575	4.00
172	bireysel	Adj	1865	5.00		208	paylaşmak	Verb	1570	4.00
173	anahtar	Noun	1862	5.00		209	yasal	Adj	1553	4.00

210	gelişmek	Verb	1547	4.00		246	yeterli	Adj	1412	4.00
211	büyümek	Verb	1542	4.00		247	ticari	Adj	1410	4.00
212	yaşanmak	Verb	1542	4.00		248	bağ	Noun	1403	4.00
213	hastalık	Noun	1540	4.00		249	gözlemek	Verb	1398	4.00
214	anlaşılmak	Verb	1538	4.00		250	analiz	Noun	1397	4.00
215	hesaplamak	Verb	1534	4.00		251	parça	Noun	1396	4.00
216	kriz	Noun	1528	4.00		252	tanım	Noun	1395	4.00
217	makine	Noun	1521	4.00		253	başarı	Noun	1393	4.00
218	çerçeve	Noun	1520	4.00		254	duygu	Noun	1392	4.00
219	olgu	Noun	1515	4.00		255	kolay	Adj	1392	4.00
220	tip	Noun	1513	4.00		256	itibar	Noun	1388	4.00
221	uluslararası	Adj	1513	4.00		257	öncelik	Noun	1381	4.00
222	dayalı	Adj	1511	4.00		258	kültürel	Adj	1380	4.00
223	denetim	Noun	1509	4.00		259	çıkar	Noun	1379	4.00
224	kaldırmak	Verb	1508	4.00		260	olumlu	Adj	1375	4.00
225	sürdürmek	Verb	1503	4.00		261	bakımından	Postp	1371	4.00
226	milli	Adj	1493	4.00		262	arttırmak	Verb	1367	4.00
227	işçi	Noun	1486	4.00		263	geniş	Adj	1364	4.00
228	bağımsız	Adj	1484	4.00		264	edinmek	Verb	1360	4.00
229	kitle	Noun	1479	4.00		265	hazırlanmak	Verb	1354	4.00
230	aramak	Verb	1475	4.00		266	etkin	Adj	1349	4.00
231	tutulmak	Verb	1472	4.00		267	standart	Noun	1347	4.00
232	yardımcı	Noun	1472	4.00		268	mevcut	Adj	1336	4.00
233	katkı	Noun	1471	4.00		269	ölçmek	Verb	1336	4.00
234	seçim	Noun	1468	4.00		270	belge	Noun	1325	4.00
235	tamamlamak	Verb	1464	4.00		271	canlı	Noun	1321	4.00
236	ışık	Noun	1457	4.00		272	yazılı	Adj	1320	4.00
237	harcama	Noun	1454	4.00		273	uyarmak	Verb	1317	4.00
238	ayrı	Adj	1449	4.00		274	fikir	Noun	1316	4.00
239	sıralamak	Verb	1446	4.00		275	yabancı	Noun	1316	4.00
240	destek	Noun	1436	4.00		276	bağlanmak	Verb	1310	4.00
241	psikolojik	Adj	1436	4.00		277	eş	Noun	1308	4.00
242	zorunlu	Adj	1435	4.00		278	tercih	Noun	1302	4.00
243	adlandırmak	Verb	1434	4.00		279	eşitlik	Adj	1299	4.00
244	altın	Noun	1426	4.00		280	teori	Noun	1295	4.00
245	olumsuz	Adj	1413	4.00		281	nüfus	Noun	1290	4.00

282	azaltmak	Verb	1289	4.00
283	geçerli	Adj	1283	4.00
284	bilimsel	Adj	1282	4.00
285	yayın	Noun	1282	4.00
286	alacak	Noun	1278	4.00
287	kurulu	Noun	1268	3.00
288	konmak	Verb	1260	3.00
289	eşit	Adj	1258	3.00
290	aracılık	Noun	1257	3.00
291	yetenek	Noun	1256	3.00
292	kapsamak	Verb	1254	3.00
293	siyasal	Adj	1250	3.00
294	tartışma	Noun	1245	3.00
295	taşınmak	Verb	1243	3.00
296	tüketici	Noun	1240	3.00
297	itibaren	Postp	1237	3.00
298	desteklemek	Verb	1222	3.00
299	gider	Noun	1221	3.00
300	kılmak	Verb	1220	3.00
301	karşılaşmak	Verb	1215	3.00
302	kaynaklanmak	Verb	1211	3.00
303	deneyim	Noun	1200	3.00
304	kişisel	Adj	1200	3.00
305	ölüm	Noun	1198	3.00
306	maddi	Adj	1188	3.00
307	uyum	Noun	1188	3.00
308	sorumlu	Noun	1188	3.00
309	deyiş	Noun	1184	3.00
310	büyüklük	Noun	1176	3.00
311	normal	Adj	1173	3.00
312	sınırlı	Adj	1166	3.00
313	bağlam	Noun	1165	3.00
314	kalan	Adj	1160	3.00
315	katılım	Noun	1150	3.00
316	modern	Adj	1150	3.00
317	toplanmak	Verb	1147	3.00

318	ortalamak	Verb	1145	3.00
319	savunmak	Verb	1139	3.00
320	rekabet	Noun	1128	3.00
321	çaba	Noun	1128	3.00
322	resmi	Adj	1124	3.00
323	hükumet	Noun	1121	3.00
324	tepki	Noun	1119	3.00
325	site	Noun	1117	3.00
326	göndermek	Verb	1116	3.00
327	zorunda	Adj	1113	3.00
328	cinsel	Adj	1111	3.00
329	hızlı	Adj	1110	3.00
330	geleneksel	Adj	1105	3.00
331	tüketim	Noun	1097	3.00
332	ters	Adj	1087	3.00
333	doğrultu	Noun	1086	3.00
334	görünmek	Verb	1084	3.00
335	aktarmak	Verb	1080	3.00
336	bakım	Noun	1080	3.00
337	ayrıntılı	Adj	1079	3.00
338	yaygın	Adj	1077	3.00
339	başlangıç	Noun	1075	3.00
340	kazandırmak	Verb	1070	3.00
341	tekrar	Adv	1065	3.00
342	değişken	Adj	1058	3.00
343	şiddet	Noun	1057	3.00
344	tutar	Noun	1056	3.00
345	norm	Noun	1056	3.00
346	hakim	Adj	1053	3.00
347	engellemek	Verb	1051	3.00
348	basit	Adj	1048	3.00
349	günlük	Adj	1036	3.00
350	uymak	Verb	1035	3.00
351	sanayi	Noun	1033	3.00
352	alıcı	Noun	1028	3.00
353	tamamen	Adv	1027	3.00

354	uzak	Adj	1024	3.00
355	hasta	Noun	1020	3.00
356	ulusal	Adj	1020	3.00
357	topluluk	Noun	1019	3.00
358	tutum	Noun	1014	3.00
359	bakış	Noun	1009	3.00
360	kötü	Adj	1009	3.00
361	eser	Noun	1007	3.00
362	istek	Noun	1003	3.00
363	müşteri	Noun	1002	3.00
364	benimsemek	Verb	993	3.00
365	sayısal	Adj	992	3.00
366	yetiştirmek	Verb	992	3.00
367	yıllık	Adj	988	3.00
368	eklemek	Verb	985	3.00
369	sınıflandırmak	Verb	982	3.00
370	izin	Noun	977	3.00
371	aktif	Adj	967	3.00
372	uzman	Noun	962	3.00
373	çatışmak	Verb	962	3.00
374	cumhuriyet	Noun	960	3.00
375	takip	Noun	960	3.00
376	vade	Noun	960	3.00
377	doğurmak	Verb	949	2.00
378	ayırmak	Verb	944	2.00
379	hazırlamak	Verb	944	2.00
380	aşırı	Adj	939	2.00
381	kesin	Adj	936	2.00
382	eşya	Noun	935	2.00
383	yaklaşık	Adj	932	2.00
384	hedeflemek	Verb	931	2.00
385	beklenti	Noun	930	2.00
386	birleşmek	Verb	927	2.00
387	dahil	Postp	926	2.00
388	imkan	Noun	924	2.00
389	birden	Adv	923	2.00
390	malzeme	Noun	923	2.00
391	güçlü	Adj	919	2.00
392	toplamak	Verb	919	2.00
393	ek	Noun	915	2.00
394	bakan	Noun	914	2.00
395	çağ	Noun	913	2.00
396	kaydetmek	Verb	910	2.00
397	açıklama	Noun	909	2.00
398	öz	Noun	909	2.00
399	hat	Noun	908	2.00
400	özgü	Adj	906	2.00
401	kurumsal	Adj	899	2.00
402	türlü	Adj	895	2.00
403	yönetmek	Verb	894	2.00
404	tarım	Noun	893	2.00
405	ermek	Verb	887	2.00
406	yük	Noun	887	2.00
407	işaret	Noun	887	2.00
408	klasik	Adj	881	2.00
409	tartışmak	Verb	881	2.00
410	proje	Noun	880	2.00
411	davranmak	Verb	878	2.00
412	eğilim	Noun	878	2.00
413	inanmak	Verb	878	2.00
414	yükselmek	Verb	875	2.00
415	anlaşma	Noun	874	2.00
416	başvurmak	Verb	872	2.00
417	güven	Noun	871	2.00
418	fırsat	Noun	870	2.00
419	serbest	Adj	869	2.00
420	bileşen	Noun	864	2.00
421	zor	Adj	863	2.00
422	çözmek	Verb	855	2.00
423	aralık	Noun	854	2.00
424	doğum	Noun	854	2.00
425	kayıp	Adj	854	2.00

426	düşünür	Noun	853	2.00	462	yansıtmak	Verb	783	2.00
427	geçmiş	Noun	853	2.00	463	toplantı	Noun	780	2.00
428	kent	Noun	845	2.00	464	vurgulamak	Verb	778	2.00
429	sözcük	Noun	845	2.00	465	ilan	Noun	777	2.00
430	siyasi	Adj	845	2.00	466	bütçe	Noun	776	2.00
431	net	Adj	845	2.00	467	kapasite	Noun	769	2.00
432	dönüştürmek	Verb	843	2.00	468	ilerlemek	Verb	768	2.00
433	değişik	Adj	842	2.00	469	sağlıklı	Adj	768	2.00
434	kimlik	Noun	842	2.00	470	güvence	Noun	766	2.00
435	bağlantı	Noun	841	2.00	471	gelişmiş	Adj	764	2.00
436	ağırlık	Noun	839	2.00	472	marka	Noun	762	2.00
437	karşılaştırmak	Verb	834	2.00	473	tespit	Noun	760	2.00
438	merkezi	Adj	831	2.00	474	asıl	Adj	759	2.00
439	öneri	Noun	830	2.00	475	statü	Noun	755	2.00
440	mesaj	Noun	829	2.00	476	erken	Adv	754	2.00
441	bölmek	Verb	826	2.00	477	oynamak	Verb	754	2.00
442	kalkınma	Noun	825	2.00	478	eleman	Noun	751	2.00
443	üstlenmek	Verb	814	2.00	479	sağ	Noun	749	2.00
444	rapor	Noun	813	2.00	480	küresel	Adj	748	2.00
445	önlem	Noun	810	2.00	481	özet	Noun	745	2.00
446	oluşum	Noun	809	2.00	482	tehdit	Noun	742	2.00
447	tarz	Noun	807	2.00	483	yazı	Noun	742	2.00
448	bakanlık	Noun	805	2.00	484	müdahale	Noun	741	2.00
449	sınırlamak	Verb	805	2.00	485	çevresel	Adj	741	2.00
450	taraflı	Adj	804	2.00	486	temsilci	Noun	740	2.00
451	başvuru	Noun	801	2.00	487	yoğun	Adj	739	2.00
452	sık	Adv	800	2.00	488	dal	Noun	738	2.00
453	boyunca	Postp	799	2.00	489	devretmek	Verb	737	2.00
454	kaçınmak	Verb	798	2.00	490	ölçüt	Noun	737	2.00
455	mekanizma	Noun	797	2.00	491	amaçlamak	Verb	735	2.00
456	performans	Noun	794	2.00	492	görevli	Noun	734	2.00
457	başarılı	Adj	793	2.00	493	uğramak	Verb	731	2.00
458	seviye	Noun	792	2.00	494	vücut	Noun	730	2.00
459	yetersizlik	Noun	792	2.00	495	denetlemek	Verb	729	2.00
460	kısaca	Adv	790	2.00	496	takdir	Noun	729	2.00
461	cinsiyet	Noun	789	2.00	497	yarar	Noun	729	2.00

498	yerleştirmek	Verb	729	2.00		534	yetkili	Noun	666	2.00
499	bildirmek	Verb	728	2.00		535	millet	Noun	665	2.00
500	yarı	Adj	726	2.00		536	baskı	Noun	663	2.00
501	gerekçe	Noun	725	2.00		537	çift	Adj	662	2.00
502	suret	Noun	724	2.00		538	engel	Noun	661	2.00
503	gidermek	Verb	723	2.00		539	husus	Noun	661	2.00
504	geçiş	Noun	722	2.00		540	yönlendirmek	Verb	659	2.00
505	aşmak	Verb	720	2.00		541	açıkça	Adj	656	2.00
506	kalite	Noun	717	2.00		542	girdi	Noun	654	2.00
507	çıkış	Noun	716	2.00		543	kavramak	Verb	654	2.00
508	evlilik	Noun	715	2.00		544	bina	Noun	650	2.00
509	yönelmek	Verb	715	2.00		545	olanaklı	Adj	649	2.00
510	seçenek	Noun	711	2.00		546	yöneltmek	Verb	649	2.00
511	kaybetmek	Verb	709	2.00		547	kalkmak	Verb	647	2.00
512	çoğunluk	Noun	709	2.00		548	disiplin	Noun	646	2.00
513	işbirliği	Noun	708	2.00		549	kapalı	Adj	646	2.00
514	çekilmek	Verb	705	2.00		550	gözlem	Noun	644	2.00
515	bedel	Noun	704	2.00		551	katı	Adj	644	2.00
516	inceleme	Noun	699	2.00		552	ciddi	Adj	641	2.00
517	kan	Noun	697	2.00		553	beraber	Adv	640	2.00
518	olasılık	Noun	695	2.00		554	kolaylaştırmak	Verb	638	2.00
519	pozitif	Adj	694	2.00		555	evrensel	Adj	635	1.00
520	yürürlük	Noun	694	2.00		556	düşürmek	Verb	633	1.00
521	etkileyen	Adj	693	2.00		557	kelime	Noun	633	1.00
522	kesim	Noun	687	2.00		558	pazar	Noun	632	1.00
523	tehlike	Noun	687	2.00		559	bozukluk	Noun	631	1.00
524	bazen	Adv	683	2.00		560	dönüşmek	Verb	631	1.00
525	adres	Noun	678	2.00		561	bulundurmak	Verb	630	1.00
526	tarihsel	Adj	678	2.00		562	vazgeçmek	Verb	628	1.00
527	karşılıklı	Adj	677	2.00		563	somut	Adj	626	1.00
528	başlatmak	Verb	671	2.00		564	çeşit	Noun	625	1.00
529	çağdaş	Adj	671	2.00		565	idare	Noun	622	1.00
530	cins	Noun	669	2.00		566	kanal	Noun	622	1.00
531	daire	Noun	668	2.00		567	önermek	Verb	622	1.00
532	değişme	Noun	667	2.00		568	ilişkili	Adj	620	1.00
533	süreli	Adj	667	2.00		569	mutlak	Adj	619	1.00

570	duygusal	Adj	617	1.00	606	dağılım	Noun	567	1.00
571	ahlaki	Adj	616	1.00	607	tahmin	Noun	566	1.00
572	önlemek	Verb	616	1.00	608	yayılmak	Verb	565	1.00
573	çalıştırmak	Verb	615	1.00	609	örgütlenmek	Verb	561	1.00
574	personel	Noun	610	1.00	610	basamak	Noun	560	1.00
575	karmaşık	Adj	609	1.00	611	ayırdetmek	Verb	559	1.00
576	çevirmek	Verb	603	1.00	612	doğruluk	Noun	559	1.00
577	ayrım	Noun	601	1.00	613	bilinç	Noun	558	1.00
578	adet	Noun	599	1.00	614	terim	Noun	558	1.00
579	dışarı	Noun	597	1.00	615	negatif	Adj	555	1.00
580	kapatmak	Verb	594	1.00	616	gündem	Noun	554	1.00
581	geçici	Adj	593	1.00	617	yorum	Noun	549	1.00
582	zorunluluk	Noun	591	1.00	618	potansiyel	Adj	548	1.00
583	öğe	Noun	590	1.00	619	istikrar	Noun	547	1.00
584	varsaymak	Verb	590	1.00	620	devrim	Noun	546	1.00
585	yönlü	Adj	587	1.00	621	etmen	Noun	542	1.00
586	dolaylı	Adj	586	1.00	622	paralel	Adj	542	1.00
587	yalnız	Adj	586	1.00	623	ölçüm	Noun	541	1.00
588	bozulmak	Verb	585	1.00	624	başlıca	Adj	537	1.00
589	saptamak	Verb	585	1.00	625	dinamik	Noun	537	1.00
590	işleyiş	Noun	584	1.00	626	akış	Noun	535	1.00
591	görüşme	Noun	583	1.00	627	gümrük	Noun	532	1.00
592	kesmek	Verb	582	1.00	628	kolaylık	Noun	532	1.00
593	cezalandırmak	Verb	579	1.00	629	nitelikli	Adj	532	1.00
594	ilgilenmek	Verb	579	1.00	630	yarmak	Verb	532	1.00
595	kas	Noun	578	1.00	631	düzenli	Adj	531	1.00
596	hata	Noun	574	1.00	632	araştırmak	Verb	528	1.00
597	beden	Noun	573	1.00	633	sıklık	Noun	528	1.00
598	ağır	Adj	572	1.00	634	iletmek	Verb	527	1.00
599	gösterge	Noun	572	1.00	635	gereç	Noun	526	1.00
600	özetlemek	Verb	572	1.00	636	konut	Noun	526	1.00
601	düzeltmek	Verb	570	1.00	637	dolmak	Verb	525	1.00
602	yönetmelik	Noun	570	1.00	638	ruh	Noun	521	1.00
603	yorumlamak	Verb	569	1.00	639	tedavi	Noun	520	1.00
604	hissetmek	Verb	568	1.00	640	verimlilik	Noun	520	1.00
605	basmak	Verb	567	1.00	641	takım	Noun	519	1.00

642	adalet	Noun	513	1.00		678	yenilik	Noun	480	1.00
643	belirleyici	Adj	513	1.00		679	retetmek	Verb	478	1.00
644	kat	Noun	510	1.00		680	yazar	Noun	476	1.00
645	boy	Noun	509	1.00		681	esnek	Adj	473	1.00
646	bitmek	Verb	509	1.00		682	çizmek	Verb	473	1.00
647	emir	Noun	509	1.00		683	ölçek	Noun	471	1.00
648	güvenmek	Verb	509	1.00		684	alışkanlık	Noun	470	1.00
649	yaklaşmak	Verb	509	1.00		685	dağıtım	Noun	468	1.00
650	durdurmak	Verb	508	1.00		686	eksik	Adj	468	1.00
651	vatandaş	Noun	507	1.00		687	tasarım	Noun	468	1.00
652	genelde	Adv	507	1.00		688	lider	Noun	467	1.00
653	bağımsızlık	Noun	505	1.00		689	bağlamak	Verb	466	1.00
654	amaçlı	Adj	504	1.00		690	kategori	Noun	466	1.00
655	yükümlü	Adj	504	1.00		691	reklam	Noun	465	1.00
656	pratik	Adj	501	1.00		692	verimli	Adj	464	1.00
657	il	Noun	498	1.00		693	çeşitlilik	Noun	464	1.00
658	bileşik	Adj	497	1.00		694	kazanç	Noun	463	1.00
659	endüstri	Noun	496	1.00		695	teşvik	Noun	461	1.00
660	soyut	Adj	495	1.00		696	tasarlamak	Verb	460	1.00
661	kapamak	Verb	494	1.00		697	bütünlük	Noun	458	1.00
662	tedbir	Noun	494	1.00		698	form	Noun	455	1.00
663	savunma	Noun	492	1.00		699	anmak	Verb	454	1.00
664	hazır	Adj	492	1.00		700	birtakım	Adj	454	1.00
665	dönüşüm	Noun	491	1.00		701	nitekim	Adv	454	1.00
666	formül	Noun	491	1.00		702	henüz	Adv	451	1.00
667	etken	Adj	486	1.00		703	kağıt	Noun	448	1.00
668	ölmek	Verb	486	1.00		704	bahsetmek	Verb	446	1.00
669	danışman	Noun	484	1.00		705	harcamak	Verb	446	1.00
670	kol	Noun	484	1.00		706	yeteri	Adj	445	1.00
671	unutmak	Verb	484	1.00		707	zorlamak	Verb	443	1.00
672	görüşmek	Verb	483	1.00		708	yapısal	Adj	441	1.00
673	aday	Noun	482	1.00		709	yükseltmek	Verb	441	1.00
674	gecikmek	Verb	482	1.00		710	eleştirmek	Verb	440	1.00
675	otorite	Noun	481	1.00		711	yapılandırmak	Verb	439	1.00
676	kararlaştırmak	Verb	480	1.00		712	anlamlı	Adj	438	1.00
677	mücadele	Noun	480	1.00		713	objektif	Adj	438	1.00

714	uyarlamak	Verb	438	1.00
715	zayıf	Adj	438	1.00
716	yerleşim	Noun	437	1.00
717	çizgi	Noun	437	1.00
718	süreklilik	Noun	432	1.00
719	odaklanmak	Verb	432	1.00
720	tarihli	Adj	431	1.00
721	organizasyon	Noun	430	1.00
722	evren	Noun	429	1.00
723	bütünleşmek	Verb	428	1.00
724	birleştirmek	Verb	428	1.00
725	pasif	Adj	428	1.00
726	yetersiz	Adj	425	1.00
727	saygı	Noun	424	1.00
728	aylık	Adj	423	1.00
729	denemek	Verb	423	1.00
730	ihmal	Noun	422	1.00
731	kurtulmak	Verb	422	1.00
732	yatmak	Verb	422	1.00
733	sınamak	Verb	421	1.00
734	boş	Adj	419	1.00
735	yerel	Adj	418	1.00
736	güçlük	Noun	417	1.00
737	katmak	Verb	416	1.00
738	hatırlamak	Verb	411	1.00
739	sembol	Noun	411	1.00
740	bilgilendirmek	Verb	410	1.00
741	başarmak	Verb	409	1.00
742	üretici	Adj	409	1.00
743	eşitsizlik	Noun	407	1.00
744	dağıtmak	Verb	406	1.00
745	kanıtlamak	Verb	406	1.00
746	iz	Noun	404	1.00
747	tüketmek	Verb	402	1.00
748	indirmek	Verb	402	1.00
749	elbette	Adv	401	1.00

750	gözlemlemek	Verb	400	1.00
751	etraf	Noun	398	1.00
752	eksiklik	Noun	397	1.00
753	şikayet	Noun	396	1.00
754	kavramsal	Adj	395	1.00
755	görünüm	Noun	394	1.00
756	tesis	Noun	394	1.00
757	kart	Noun	393	1.00
758	sapmak	Verb	391	1.00
759	tane	Noun	390	1.00
760	yararlı	Adj	390	1.00
761	yoğunlaşmak	Verb	390	1.00
762	irdelemek	Verb	389	1.00
763	değinmek	Verb	388	1.00
764	karakter	Noun	388	1.00
765	nitelendirmek	Verb	387	1.00
766	izleyici	Noun	386	1.00
767	dernek	Noun	385	1.00
768	sistematik	Adj	385	1.00
769	taban	Noun	385	1.00
770	yollamak	Verb	385	1.00
771	yoğunluk	Noun	383	1.00
772	değerli	Adj	382	1.00
773	yüklemek	Verb	380	1.00
774	önlenmek	Verb	379	1.00
775	erişim	Noun	378	1.00
776	sınırlandırmak	Verb	378	1.00
777	zincir	Noun	378	1.00
778	çıkartmak	Verb	378	1.00
779	alışveriş	Noun	377	1.00
780	yaptırım	Noun	377	1.00
781	açıklık	Noun	376	1.00
782	rastlamak	Verb	376	1.00
783	girişim	Noun	375	1.00
784	iklim	Noun	374	1.00
785	saymak	Verb	371	1.00

161

786	belirgin	Adj	370	1.00		822	üstün	Adj	336	1.00
787	nesil	Noun	370	1.00		823	insanlık	Noun	332	1.00
788	yoksun	Adj	370	1.00		824	çözülmek	Verb	331	1.00
789	kısmen	Adv	367	1.00		825	inşa	Noun	330	1.00
790	kuşak	Noun	367	1.00		826	götürmek	Verb	329	1.00
791	birikim	Noun	366	1.00		827	tabii	Adj	328	1.00
792	kapsamlı	Adj	366	1.00		828	yetmek	Verb	328	1.00
793	dokunmak	Verb	365	1.00		829	faydalanmak	Verb	327	1.00
794	uyumlu	Adj	364	1.00		830	dizi	Noun	326	1.00
795	anlatım	Noun	363	1.00		831	ulaştırmak	Verb	324	1.00
796	yürümek	Verb	363	1.00		832	mantık	Noun	322	1.00
797	üstünlük	Noun	362	1.00		833	stratejik	Adj	322	1.00
798	saklamak	Verb	362	1.00		834	maruz	Adj	321	1.00
799	yaygınlaşmak	Verb	362	1.00		835	zemin	Noun	321	1.00
800	işyeri	Noun	361	1.00		836	zengin	Adj	321	1.00
801	yüklenmek	Verb	360	1.00		837	ilgilendirmek	Verb	320	1.00
802	dini	Adj	359	1.00		838	not	Noun	317	0.00
803	genişlemek	Verb	358	1.00		839	liste	Noun	315	0.00
804	farklılaşmak	Verb	355	1.00		840	dergi	Noun	314	0.00
805	pozisyon	Noun	355	1.00		841	genişletmek	Verb	313	0.00
806	yayımlamak	Verb	354	1.00		842	belirsizlik	Noun	312	0.00
807	erişmek	Verb	353	1.00		843	doldurmak	Verb	311	0.00
808	hasar	Noun	353	1.00		844	kalıcı	Adj	311	0.00
809	toplu	Adj	353	1.00		845	yakalamak	Verb	311	0.00
810	eleştiri	Noun	352	1.00		846	sınav	Noun	311	0.00
811	alışmak	Verb	349	1.00		847	yanmak	Verb	309	0.00
812	dönüş	Noun	348	1.00		848	şehir	Noun	306	0.00
813	katılımcı	Noun	348	1.00		849	yaymak	Verb	306	0.00
814	temelli	Adj	345	1.00		850	arzu	Noun	303	0.00
815	barındırmak	Verb	344	1.00		851	yansımak	Verb	303	0.00
816	bildirim	Noun	342	1.00		852	ödül	Noun	303	0.00
817	yavaş	Adj	342	1.00		853	hazırlık	Noun	302	0.00
818	güçlendirmek	Verb	341	1.00		854	derin	Adj	301	0.00
819	danışmanlık	Noun	338	1.00		855	yaratıcı	Adj	301	0.00
820	yayınlanmak	Verb	338	1.00		856	bağımlı	Adj	300	0.00
821	dar	Adj	336	1.00		857	sıkı	Adj	300	0.00

858	yenilenmek	Verb	300	0.00	894	oynaklık	Noun	273	0.00
859	gündelik	Adj	299	0.00	895	yapım	Noun	273	0.00
860	kaza	Noun	298	0.00	896	yaramak	Verb	273	0.00
861	duyarlı	Adj	297	0.00	897	güvenli	Adj	272	0.00
862	kalıp	Noun	297	0.00	898	özgür	Adj	271	0.00
863	benzemek	Verb	295	0.00	899	kuşkusuz	Noun	270	0.00
864	dilimlemek	Verb	295	0.00	900	tekrarlamak	Verb	269	0.00
865	ani	Adj	294	0.00	901	iyileştirmek	Verb	269	0.00
866	bağımlılık	Noun	294	0.00	902	kırmak	Verb	268	0.00
867	işlevsel	Adj	292	0.00	903	gözetmek	Verb	265	0.00
868	onaylamak	Verb	289	0.00	904	teşkil	Noun	264	0.00
869	kriter	Noun	288	0.00	905	mektup	Noun	263	0.00
870	sıkıntı	Noun	288	0.00	906	özenmek	Verb	262	0.00
871	yerleşmek	Verb	287	0.00	907	masraf	Noun	262	0.00
872	gereklilik	Noun	285	0.00	908	tanıtmak	Verb	261	0.00
873	kesinlik	Noun	285	0.00	909	harf	Noun	260	0.00
874	işleyen	Adj	283	0.00	910	konferans	Noun	260	0.00
875	benzerlik	Noun	282	0.00	911	sunu	Noun	260	0.00
876	sivil	Adj	282	0.00	912	gönüllü	Adj	258	0.00
877	lisans	Noun	282	0.00	913	vasıta	Noun	257	0.00
878	kısıtlamak	Verb	281	0.00	914	taklit	Noun	256	0.00
879	sınıflamak	Verb	281	0.00	915	yaradan	Noun	256	0.00
880	donanım	Noun	280	0.00	916	zorlanmak	Verb	254	0.00
881	adım	Noun	279	0.00	917	zorluk	Noun	252	0.00
882	istismar	Noun	279	0.00	918	biçimsel	Adj	250	0.00
883	gizli	Adj	278	0.00	919	takmak	Verb	250	0.00
884	uzmanlık	Noun	278	0.00	920	durulmak	Verb	249	0.00
885	yitirmek	Verb	278	0.00	921	soy	Adj	249	0.00
886	motivasyon	Noun	277	0.00	922	zararlı	Adj	248	0.00
887	zamanlamak	Verb	277	0.00	923	dayanak	Noun	247	0.00
888	ağırlıklı	Adj	275	0.00	924	kaçmak	Verb	247	0.00
889	yaptırmak	Verb	275	0.00	925	bağlılık	Noun	246	0.00
890	bilinçli	Noun	274	0.00	926	coğrafi	Adj	246	0.00
891	paydaş	Noun	274	0.00	927	rakip	Noun	246	0.00
892	hayal	Noun	273	0.00	928	ayrıntı	Noun	244	0.00
893	koşmak	Verb	273	0.00	929	koruyucu	Adj	244	0.00

930	uğraşmak	Verb	242	0.00
931	şans	Noun	241	0.00
932	kurtarmak	Verb	239	0.00
933	uzaklaşmak	Verb	237	0.00
934	fabrika	Noun	235	0.00
935	bozmak	Verb	232	0.00
936	düz	Adj	232	0.00
937	bünye	Noun	231	0.00
938	şema	Noun	231	0.00
939	işletmek	Verb	231	0.00
940	bağlantılı	Adj	229	0.00
941	tipik	Adj	229	0.00
942	uzlaşmak	Verb	229	0.00
943	imparatorluk	Noun	229	0.00
944	aktarım	Noun	228	0.00
945	davranışsal	Adj	228	0.00
946	görünüş	Noun	228	0.00
947	yenilemek	Verb	228	0.00
948	arayış	Noun	226	0.00
949	öncelikli	Adj	226	0.00
950	tıpkı	Adv	224	0.00
951	sonuçlanmak	Verb	224	0.00
952	ulaşım	Noun	223	0.00
953	yerleşik	Adj	222	0.00
954	ziyaret	Noun	216	0.00
955	terk	Noun	215	0.00
956	çap	Noun	214	0.00
957	rahat	Adj	212	0.00
958	verici	Noun	212	0.00
959	karışmak	Verb	211	0.00
960	imkansız	Noun	209	0.00
961	dayanışma	Noun	208	0.00
962	rutin	Adj	208	0.00
963	ömür	Noun	208	0.00
964	ihtimal	Noun	206	0.00
965	güncel	Adj	205	0.00
966	dengelemek	Verb	204	0.00
967	karma	Adj	204	0.00
968	katlanmak	Verb	203	0.00
969	gereksiz	Noun	202	0.00
970	inmek	Verb	201	0.00
971	fizyolojik	Adj	199	0.00
972	organize	Adj	198	0.00
973	dinlenmek	Verb	196	0.00
974	gözetim	Noun	196	0.00
975	hafif	Adj	196	0.00
976	altyapı	Noun	195	0.00
977	çoğaltmak	Verb	193	0.00
978	kıyas	Noun	192	0.00
979	odaklı	Adj	190	0.00
980	karıştırmak	Verb	189	0.00
981	vurmak	Verb	186	0.00
982	duyarlılık	Noun	185	0.00
983	mikro	Adj	185	0.00
984	gösteri	Noun	182	0.00
985	kastetmek	Verb	181	0.00
986	sakınca	Noun	178	0.00
987	sokmak	Verb	177	0.00
988	detaylı	Adj	175	0.00
989	uzatmak	Verb	174	0.00
990	kaçırmak	Verb	171	0.00
991	duyurmak	Verb	170	0.00
992	uzamak	Verb	169	0.00
993	hızlandırmak	Verb	168	0.00
994	muhtemel	Adj	166	0.00
995	donatmak	Verb	165	0.00
996	halen	Adv	160	0.00
997	kavuşmak	Verb	159	0.00
998	belediye	Noun	156	0.00
999	üstelik	Adv	156	0.00
1000	beğenmek	Verb	153	0.00
1001	patlamak	Verb	153	0.00

1002	belirsiz	Adj	152	0.00
1003	zamanlı	Adj	149	0.00
1004	denk	Adj	142	0.00
1005	endişe	Noun	139	0.00
1006	yıldırmak	Verb	134	0.00
1007	tecrübe	Noun	128	0.00
1008	kıyaslamak	Verb	126	0.00
1009	sarmak	Verb	115	0.00
1010	vakit	Noun	105	0.00

Appendix 13 The First 100 Words According to Part of Speech (The Academic Turkish Wordlist)

	Verbs	Nouns	Adjectives	Adverbs	Others
1	öğrenmek	davranış	toplumsal	dolayısıyla	ilişkin
2	belirlemek	işletme	çeşitli	sürekli	yönelik
3	oluşturmak	birey	gerekli	genellikle	açısından
4	geçirmek	gelişim	belirli	tekrar	ait
5	açıklamak	kavram	teknik	tamamen	bakımından
6	uygulamak	sözleşme	etkili	birden	itibaren
7	değerlendirmek	aşağı	finansal	sık	dahil
8	gerçekleştirmek	işlem	ortak	kısaca	boyunca
9	tanımlamak	yanıt	benzer	erken	
10	geliştirmek	kanun	doğal	bazen	
11	artmak	kurum	fiziksel	beraber	
12	kurulmak	uygulama	esas	genelde	
13	yapılanmak	varlık	doğrudan	nitekim	
14	ulaşmak	araç	bireysel	henüz	
15	korumak	suç	sabit	elbette	
16	gerçekleşmek	meslek	düşük	kısmen	
17	düzenlemek	iletişim	yasal	tıpkı	
18	taşımak	mal	uluslararası	halen	
19	tanımak	ilke	dayalı	üstelik	
20	yaratmak	kural	milli		
21	değişmek	ortam	bağımsız		
22	kazanmak	etkinlik	ayrı		
23	sunmak	hüküm	psikolojik		
24	üretmek	koşul	zorunlu		
25	içermek	politika	olumsuz		
26	ödemek	nitelik	yeterli		
27	etkilemek	sınıf	ticari		
28	incelemek	yöntem	kolay		
29	izlemek	öğretim	kültürel		
30	değiştirmek	faaliyet	olumlu		
31	sürmek	yaklaşım	geniş		
32	dayanmak	talep	etkin		
33	işlemek	ürün	mevcut		
34	azalmak	sermaye	yazılı		
35	doğmak	ekonomi	eşitlik		
36	seçmek	kuram	geçerli		

37	yürütmek	araştırma	bilimsel
38	öngörmek	kültür	eşit
39	katılmak	görüş	siyasal
40	sayılmak	zarar	kişisel
41	yararlanmak	hesap	maddi
42	satmak	ihtiyaç	normal
43	planlamak	ceza	sınırlı
44	devralmak	model	kalan
45	karşılamak	hedef	modern
46	artırmak	gelişme	resmi
47	algılamak	teknoloji	zorunda
48	ayrılmak	toplam	cinsel
49	gerektirmek	sanat	hızlı
50	paylaşmak	miktar	geleneksel
51	gelişmek	yetki	ters
52	büyümek	fiyat	ayrıntılı
53	yaşanmak	kamu	yaygın
54	anlaşılmak	üye	değişken
55	hesaplamak	değişim	hakim
56	kaldırmak	kullanım	basit
57	sürdürmek	beceri	günlük
58	aramak	şirket	uzak
59	tutulmak	sorumluluk	ulusal
60	tamamlamak	risk	kötü
61	sıralamak	önem	sayısal
62	adlandırmak	denge	yıllık
63	gözlemek	hız	aktif
64	arttırmak	karşılık	aşırı
65	edinmek	kontrol	kesin
66	hazırlanmak	kısım	yaklaşık
67	ölçmek	medya	güçlü
68	uyarmak	rol	özgü
69	bağlanmak	aşama	kurumsal
70	azaltmak	bilim	türlü
71	konmak	kuruluş	klasik
72	kapsamak	nesne	serbest
73	taşınmak	ticaret	zor
74	desteklemek	birim	kayıp
75	kılmak	pay	siyasi
76	karşılaşmak	felsefe	net

77	kaynaklanmak	örgüt	değişik
78	toplanmak	gereksinim	merkezi
79	ortalamak	değişiklik	taraflı
80	savunmak	unsur	başarılı
81	göndermek	içerik	sağlıklı
82	görünmek	işlev	gelişmiş
83	aktarmak	fonksiyon	asıl
84	kazandırmak	plan	küresel
85	engellemek	sebep	çevresel
86	uymak	anlayış	yoğun
87	benimsemek	ücret	yarı
88	yetiştirmek	yardım	pozitif
89	eklemek	yasa	etkileyen
90	sınıflandırmak	yatırım	tarihsel
91	çatışmak	artış	karşılıklı
92	doğurmak	sektör	çağdaş
93	ayırmak	şart	süreli
94	hazırlamak	faktör	çift
95	hedeflemek	olanak	açıkça
96	birleşmek	veri	olanaklı
97	toplamak	düzenleme	kapalı
98	kaydetmek	ileri	katı
99	yönetmek	maliyet	ciddi
100	ermek	fark	evrensel

Appendix 14 Multiple Comparisons of Coursebooks

COURSEBOOKS	COURSEBOOKS	Mean Difference	Std. Error	Sig.
Gazi Üniversitesi TÖMER Yabancılar İçin Türkçe Seti	İstanbul Üniversitesi Dil Merkezi İstanbul Seti	-.21207*	.02939	.000
	Ankara Üniversitesi TÖMER Yeni Hitit Yabancılar İçin Türkçe Öğretim Seti	.02851	.03767	.836
İstanbul Üniversitesi Dil Merkezi İstanbul Seti	Gazi Üniversitesi TÖMER Yabancılar İçin Türkçe Seti	.21207*	.02939	.000
	Ankara Üniversitesi TÖMER Yeni Hitit Yabancılar İçin Türkçe Öğretim Seti	.24058*	.04210	.000
Ankara Üniversitesi TÖMER Yeni Hitit Yabancılar İçin Türkçe Öğretim Seti	Gazi Üniversitesi TÖMER Yabancılar İçin Türkçe Seti	-.02851	.03767	.836
	İstanbul Üniversitesi Dil Merkezi İstanbul Seti	-.24058*	.04210	.000

*. The mean difference is significant at the 0.05 level.

Appendix 15 TrLister Technical Specifications

TrLister Technical Specifications
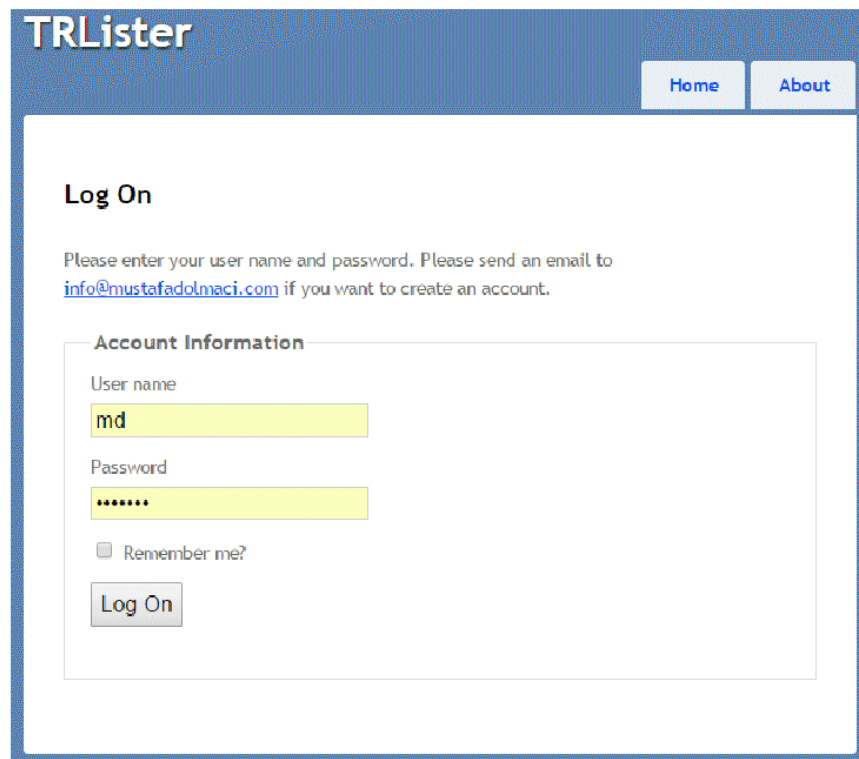
Introduction

TrLister is a web application hosted on http://www.mustafadolmaci.com web address that keeps list of Turkish words acquired from the Academic Corpus which includes 57textbooks and helps users to define lemmas.

The application could be analysed in three different parts.

Authentication

The application has built in authentication mechanism that allows only registered users to operate. Currently, there are 6 users defined in the system.

The user is able to set an authentication cookie on client, which remembers the user logged in for a parametric period of time (currently 7 days).

Adding new user could be achieved by executing SQL statements directly into database. There is no registration page for the sake of simplicity.

Users could visit a simple Home page and About page without logging in. All operations and reports related to words and bases require Users to authenticate.

Word List

This page lists all the words defined in the system and shows detailed information about them. There are 3 main operations users could do in this screen.

Fields

While listing words, detailed information is also displayed. The purpose of this information is to give users details about the words they are operating. The columns are:

- Word: Text of the word
- Root: Lemma of given words
- Total: Total number of occurance in 5 categories
- Frequency: Number of categories the word appears
- Count1 to Count5: Number of occurance of the word in each category respectively.

171

Assigning word to base (lemma)

After user select a set of words, s/he can assign them to a lemma by filling the text box with lemma and pressing "Add To Root" button. All selected words will be detached from their current base and will be assigned to new base(lemma) typed.

Users cannot assign deleted words to lemmas. Only unassigned or already-assigned words could be assigned to a lemma.

Disabling (or removing) words

Users are able to select a number of words and disable them from operation by pressing "x" button. All selected words will be detached from their current root and marked as disabled in the system. Disabled words will not be counted when generating reports.

Enabling words

Enable button marks words are enabled. The button appears next to disabled buttons only. The button allows users to undo disabling selected words.

Report

Users are able to generate reports based on the words and lemmas they previously specified. A number of filters are displayed on "Report Screen" and an excel file is generated when user presses Generate button.

Filters

13 different filters could be entered before generating the output.

The following filters are taken into account when generating the report:

- Minimum Frequency: Only lemmas that have at least given frequency will be included

- Maximum Frequency: Only lemmas less than or equal to given frequency will be included

- Minimum Number of Sub Words: Only lemmas that have at least given number of sub words will be included

- Minimum Total Appearance: Only lemmas that have at least given total appearance will be included

- Maximum Total Appearance: The lemmas that have more than given total appearance will be excluded

- Include Subwords: If checked, individual words attached to lemmas will be displayed in the report. If not checked, only lemmas will be generated.

- Min - Max Appearances in Book #1 - 5: Every lemma in the report sould have at least given minimum appearance on each category respectively and should not exceed the maximum appearance given.

- Has To Appear In At Least X books Y times each: Only the lemmas that have at least Y appearance in at least X different gategories will be included.

- Exclude these words: Given words will be excluded from the result.

When user presses Generate button, the filters above are executed on database and the output file is an excel document containing lemmas, words, frequencies, and percentages.

Excel Format

The output of report screen is named as Report_DATE_TIME.xlsx

Roots Tab

	A	B	C	D	E	F	G	H	I	J	K
1	Id	Root	Frequency	Count 1	Count 2	Count 3	Count 4	Count 5	Total Appearance	# of Subwords	%oo Appearance
2	299	ALMAK	5	3286	3516	3118	3458	4092	17470	513	54,00
3	310	ALAN	5	2092	1314	3566	3061	1319	11352	64	35,00
4	398	ALT	5	986	643	910	611	1040	4190	21	13,00
5	422	AMAÇ	5	1006	1512	1588	3423	1996	9525	58	29,00
6	655	ARA	5	2436	2247	3260	1983	2629	12555	40	39,00
7	797	AŞAĞI	5	1599	879	1004	1135	1295	5912	34	18,00
8	1065	AYNI	5	1371	1139	1409	875	1172	5966	16	18,00
9	1127	AÇMAK	5	1118	1251	772	633	721	4495	265	14,00

# of Subwords column shows the total number of words attached to each lemma.

% Appearance is total appearance of base divided by total appearance of all lemmas.

The rest of the columns are similar to the table in Words screen.

Words Tab

	A	B	C	D	E
1	Id	Root	Word	Total Appearance	% Appearance
2	2144	ALMAK	ALABİBLİRİZ	1	0,01 %
3	2146	ALMAK	ALABİLDİ	1	0,01 %
4	2147	ALMAK	ALABİLDİĞİ	4	0,02 %
5	2148	ALMAK	ALABİLDİĞİMİZ	1	0,01 %
6	2149	ALMAK	ALABİLDİĞİNCE	3	0,02 %
7	2150	ALMAK	ALABİLDİĞİNDEN	1	0,01 %
8	2151	ALMAK	ALABİLDİĞİNE	15	0,09 %
9	2152	ALMAK	ALABİLDİĞİNİ	2	0,01 %

% Appearance column is total appearance of each word divided by total appearance of the lemma it is attached.

Technology

The web application was developed using Microsoft ASP.NET MVC 3. Microsoft SQL Server 2012 was used as database engine and some open source libraries and frameworks were included in each level

Server Side

The following frameworks were included into project on server side:

- System.Web.MVC Version 3: MVC framework allows developers to separate view from data and provides easy to develop and easy to maintain web interfaces.
- System.Web.Razor: Razor view engine is used instead of ASPX in order to simplify code.
- Entity Framework Version 5: The communication between server side MVC application and database is accomplished with Entity Framework. EF provides database server independent implementation and allows developers to easily switch between database engines
- NPOI: All excel related operations in report screen is developed using NPOI open source excel libraries.

Client Side

- jQuery: jQuery framework is a javascript library that enables developers to build user-friendly interfaces.
- MicrosoftMvcValidation: MVC Validation framework is used to validate user inputs on client side.
- All user interface is designed using responsive layout, meaning the application will work on all devices and all screen sizes

GAZİ GELECEKTİR…