THE RELATIONSHIP BETWEEN VOCABULARY SIZE, LEXICAL DIVERSITY, LEXICAL DENSITY AND EFL WRITING SCORES: A CROSS-SECTIONAL STUDY

MUSA TÖMEN

Yüksek Lisans Tezi

Eskişehir, 2016

THE RELATIONSHIP BETWEEN VOCABULARY SIZE, LEXICAL DIVERSITY, LEXICAL DENSITY AND EFL WRITING SCORES: A CROSS-SECTIONAL STUDY

MUSA TÖMEN

MA THESIS

MA in English Language Teaching Program

Program in Foreign Language Education

Supervisor: Prof. Dr. Gül DURMUŞOĞLU KÖSE

Eskişehir

Anadolu University

Graduate School of Educational Sciences

June, 2016

JÜRİ VE ENSTİTÜ ONAYI

(APPROVAL OF JURY AND THE INSTITUTION)

Musa TÖMEN'in The Relationship Between Vocabulary Size, Lexical Diversity, Lexical Density and Efl Writing Scores: A Cross-Sectional Study" başlıklı tezi 27.06.2016 tarihinde, aşağıda belirtilen jüri üyeleri tarafından Anadolu Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin ilgili maddeleri uyarınca Yabancı Diller Eğitimi Anabilim Dalı İngilizce Öğretmenliği programı yüksek lisans tezi olarak değerlendirilerek kabul edilmiştir.

	Adı-Soyadı	**İmza**
Üye (Tez Danışmanı)	: Prof.Dr. Gül DURMUŞOĞLU KÖSE	
Üye	: Prof.Dr. Zülal BALPINAR	
Üye	: Doç.Dr. İlknur SAVAŞKAN	
Üye	: Yard.Doç.Dr. Gonca SUBAŞI	
Üye	: Yard.Doç.Dr. Hasan ÇEKİÇ	

Prof.Dr. Handan DEVECİ
Anadolu Üniversitesi
Eğitim Bilimleri Enstitüsü
Müdür Vekili

ABSTRACT

THE RELATIONSHIP BETWEEN VOCABULARY SIZE, LEXICAL DIVERSITY, LEXICAL DENSITY AND EFL WRITING SCORES: A CROSS-SECTIONAL STUDY

Musa TÖMEN

MA in English Language Teaching - Department of Foreign Language Education

Anadolu University Graduate School of Educational Sciences, June, 2016

Supervisor: Prof. Dr. Gül DURMUŞOĞLU KÖSE

Vocabulary is an important aspect of language learning. Many studies were carried on vocabulary, vocabulary knowledge, vocabulary instruction, vocabulary learning and since 1990s the focus of the studies have shifted on the effects of vocabulary on second language (L2) learning. For Nation (2000) knowing a word includes form, meaning and use. The role of vocabulary knowledge in receptive and productive L2 skills cannot be ignored as it is used in predicting the learners' competence. It is obvious that this would affect L2 learners in expressing their thoughts through productive skills. The aim of the study is to find out the vocabulary size, lexical density and lexical diversity in argumentative essays of Turkish ELT students and to compare these with their writing scores. It is assumed that high vocabulary size, lexical density, and lexical diversity may lead to higher scores. The data, including 309 argumentative essays written by 165 first year and 144 fourth year students, were collected from Anadolu University ELT Department. In order to analyze the data and find answers to the research questions the LFP was used to obtain an estimate productive vocabulary size of the students, the vocd-D was used for lexical diversity of the essays, and the lexical density formula was used. According to the results, among other variables, only lexical diversity was found to have a significant correlation with 1st year students' essay scores and it significantly explained 7.8% of the 1st year essay scores. For the 4th year essays, no significant effect of the variables was found. Moreover, the results also demonstrated that while the variables accounted for 8.7% of the vocabulary scores in the 1st year essays, they did not yield a significant explanation for the 4th year essays vocabulary scores. Therefore, considering the

results of the study, it can be concluded that lexical features are not the only factors effecting writing scores.

Key Words: Vocabulary size, Lexical diversity, Lexical density, EFL writing, Turkish ELT students.

ÖZET

SÖZCÜK BOYUTU, KELİME ÇEŞİTLİLİĞİ VE KELİME YOĞUNLUĞU İLE YABANCI DİL OLARAK İNGİLİZCE YAZMA PUANLARI ARASINDAKİ İLİŞKİ: KESİTSEL ÇALIŞMA ÖRNEĞİ

Musa TÖMEN
İngilizce Öğretmenliği Programı
Yabancı Diller Eğitimi Anabilim Dalı
Anadolu Üniversitesi Eğitim Bilimleri Enstitüsü, Haziran, 2016
Danışman: Prof. Dr. Gül DURMUŞOĞLU KÖSE

Sözcük, dil öğreniminin önemli bir boyutudur. Sözcük bilgisi, sözcük öğretimi ve sözcük öğrenimi ile ilgili birçok çalışma yapılmış, 1990'lardan beri de çalışmaların odak noktası sözcük bilgisinin ikinci dil edinimine etkisi üzerine yoğunlaşmıştır. Nation (2000)'a göre bir sözcüğü bilmek sözcüğün formunu, anlamını ve kullanımını bilmeyi gerektirir. Sözcük bilgisi öğrencilerin sözcük yeterliğini yordamak için kullanıldığından sözcük bilgisinin algısal ve üretimsel dil becerilerindeki rolü yadsınamaz. Sözcük bilgisinin ikinci dil öğrenenlerin düşüncelerini üretimsel becerilerle ifade etmesini etkileyebileceği açıktır. Bu çalışmanın amacı anadili Türkçe olan İngilizce Öğretmenliği öğrencilerinin sözcük boyutu, sözcük çeşitliliği ve sözcük yoğunluklarını bulmak ve bu değerleri kompozisyon notları ile karşılaştırmaktır. Kapsamlı sözcük boyutu, sözcük çeşitliliği ve sözcük yoğunluğunun yüksek kompozisyon notlarına sebep olacağı düşünülmektedir. Çalışma için veri 165 birinci sınıf, 144 dördüncü sınıf Anadolu Üniversitesi İngilizce Öğretmenliği Bölümü öğrencisinden, 309 tartışmacı (argumentative) kompozisyon türünde toplanmıştır. Veriyi incelemek ve araştırma sorularına cevap bulmak için sözcük boyutu LFP kullanılarak, sözcük çeşitliliği vocd-D hesaplama yöntemi kullanılarak ve sözcük yoğunluğu ise sözcük yoğunluğu formülü kullanılarak hesaplanmıştır. Sonuçlara göre, diğer değişkenler arasında sadece sözcük çeşitliliğinin birinci sınıf öğrencilerinin kompozisyon notlarıyla anlamlı bir şekilde ilişkili olduğunu göstermiştir. Sözcük

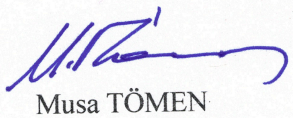
çeşitliliği kompozisyon notlarının %7,8'ini anlamlı bir şekilde açıklamaktadır. 4. Sınıf kompozisyonları için değişkenler arasında kompozisyon notlarını anlamlı bir şekilde etkileyen değişken bulunamamıştır. Sonuçlar, bu değişkenlerin birlikte 1. sınıf kompozisyon notlarının %8,7'sini açıkladığını ancak 4. sınıf notlarını ise anlamlı bir şekilde açıklamadığını göstermiştir. Sonuç olarak, kelimeye bağlı değişkenlerin yazma notlarını etkileyen tek faktör olmadığı söylenebilir.

Key Words: Sözcük boyutu, Sözcük çeşitliliği, Sözcük yoğunluğu, İkinci dilde yazma, Anadili Türkçe olan İngilizce Öğretmenliği öğrencileri.

15/08/2016

ETİK İLKE VE KURALLARA UYGUNLUK BEYANNAMESİ

Bu tezin bana ait, özgün bir çalışma olduğunu; çalışmamın hazırlık, veri toplama, analiz ve bilgilerin sunumu olmak üzere tüm aşamalardan bilimsel etik ilke ve kurallara uygun davrandığımı; bu çalışma kapsamında elde edilemeyen tüm veri ve bilgiler için kaynak gösterdiğimi ve bu kaynaklara kaynakçada yer verdiğimi; bu çalışmanın Anadolu Üniversitesi tarafından kullanılan "bilimsel intihal tespit programı"yla tarandığını ve hiçbir şekilde "intihal içermediğini" beyan ederim. Herhangi bir zamanda, çalışmamla ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara razı olduğumu bildiririm.



Musa TÖMEN

ACKNOWLEDGEMENTS

This thesis could only be written thanks to the guidance, support, and help of many people around me.

First of all, I would like to express my deepest gratitude to my thesis supervisor, Prof. Dr. Gül DURMUŞOĞLU KÖSE, for her continuous support and feedback. Without doubt, this thesis would not have been completed without her caring support, patience and being a research model.

I would like to thank my thesis committee, Prof. Dr. Zülal BALPINAR, Assoc. Prof. Dr. İlknur SAVAŞKAN, Assist. Prof. Dr. Hasan ÇEKİÇ, and Assist. Prof. Dr. Gonca SUBAŞI for their feedback to this thesis.

I also owe all my professors at the ELT department of Anadolu University a debt of gratitude. It was a great chance for me to study and conduct research at this department.

I would also like to thank my colleagues Samet TAŞÇI, Yusuf ÖZTÜRK, Ahmet TANIR and Pınar KARAHAN for their feedback, proofreading and technical support during the process. Buğra KAŞ and Dilek KARAKOÇ also deserve my gratitude for grading the students' essays and accepting to help without any hesitations.

I also feel obliged to thank our students Selen ULUÇAY, Gizem BALCI, Zeynep COŞKUN and Rüya KUTNAL for their support in transferring the essays into Word format.

I cannot forget to thank Prof. Dr. Mustafa TAŞKIN, who guided me and directed me pursue what I want when I as in Elazığ.

My biggest thanks are for my brother, my mother and especially my mother Meryem TÖMEN, who urged me follow an academic career. Thank you all for supporting my every decision.

Musa TÖMEN

Eskişehir, 2016

To my family,

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATONS

L2 : Second of foreign language

SLA : Second language acquisition

VKS : Vocabulary knowledge scale

WAT : Word associates test

L1 : First language

MTLD : Measure of Textual Lexical Diversity

Vocd-D : A lexical diversity measure

LFP : Lexical frequency profile

COCA : Corpus of Contemporary American English

BNC : British National Corpus

GSL : General Service List

AWL : Academic Word List

YDS : Foreign Language Examination

ELT : English Language Teaching

VKT : Vocabulary knowledge test

VLT : Vocabulary Level Tests

PVLT : Productive Vocabulary Level Tests

UWL : University Word List

LD : Lexical density

ESL : English as a Second Language

TTR : Type-Token Ratio

1. INTRODUCTION

Vocabulary is an important aspect of language learning. Many studies were carried on vocabulary, vocabulary knowledge, vocabulary instruction, vocabulary learning and since 1990s the focus of the studies have shifted on the effects of vocabulary on second language (L2) learning. Vocabulary knowledge has been defined variously. Cronbach (1942) introduces five aspects for knowing a word: *generalization, application, breadth, precision,* and *availability*. *Generalization* is to be able to define a word, *application* is to be able to select and recognize situations to use a word appropriately, *breadth* is to be aware of other meanings of a word, *precision* is to be able to use a word precisely, and *availability* is the ability to use the word in discourse. For Nation (2000) knowing a word includes form, meaning and use. He also introduces the terms receptive and productive vocabulary in his detailed definition. Simply, it is known that receptive vocabulary knowledge of a learner consists of words that the learner is able to remember while reading or listening. Productive vocabulary knowledge of a learner, on the other hand, includes the words that the learner is able to use accurately and appropriately while speaking or writing (Nation, 2000).

Vocabulary is considered as the heart of language comprehension and use (Hunt & Beglar, 2005). Productive vocabulary knowledge is widely accepted to be the most encountered language problem that L2 learners encounter (Nation, 1990; Schmitt, 1997; Mokhtar, 2010). Vocabulary use is also an essential indicator of language knowledge both in one's native language and in an L2. While without grammar one can communicate to some extent, without vocabulary it is not possible to convey a message. Nevertheless, the vocabulary studies were not as important as other areas of L2 research such as methodology and grammar teaching until 1990s. Recent research in second language acquisition (SLA) has stated that vocabulary knowledge is the prerequisite condition for the development of other language skills (Gass & Selinker, 2008; Nation, 2006; Roche & Harrington, 2013). Without doubt, vocabulary knowledge has utmost effect on reading and listening comprehension as many studies conducted on the issue state so. Nation (2006); Hu and Nation (2000), for example, state that in order for a learner to comprehend an academic text, s/he should know at least 8,000 word families. 8,000 word families are stated to be enough

to translate approximately 98% text coverage, which shows the effect of vocabulary on reading comprehension. Schmitt (2000); Hu and Nation (2000); Dang and Webb (2013); Silverman et al. (2015) and Zhang et al. (2014) also conducted studies focusing on the correlational relationship between the size of vocabulary the learners have and their reading comprehension scores. Ling (2015); Vandergiff and Baker (2015) and Wang (2015) showed that L2 vocabulary knowledge correlate positively with listening comprehension. It is stated in the studies above that there is a positive correlation between vocabulary knowledge and receptive skills of L2 learners, that is, vocabulary knowledge contributes significantly to reading and listening comprehension. Moreover, L2 learners in the studies regarding listening comprehension stated that they were having difficulties when they encountered unknown vocabulary in the listening texts given (Hamouda, 2013; Solak & Altay, 2014). It can be said that learners who face problems even in receptive skills related to their receptive vocabulary knowledge, are definitely facing problems in productive skills and therefore, in productive vocabulary use.

A certain level of vocabulary knowledge plays an important role in L2 learning and it is necessary for proficiency and fluency. Thus, in terms of productive vocabulary use, the impact of vocabulary knowledge on L2 learners' productive language skills -writing and speaking- has also been studied. Several studies have found that the learners who state that they are experiencing difficulties in writing are directly affected by lack of vocabulary knowledge (Begriche, 2013; Putra, 2014; Rudy, 2013; Yang, 2015). Several other studies have revealed that there is a positive correlation with vocabulary knowledge and speaking; and L2 learners have difficulties such as long silences to choose a word, hesitations while speaking because of their lack of vocabulary knowledge (Fhonna, 2014; Khotimah, 2014; Tahir, 2015; Perez Manzanilla & Diaz Cabrera, 2014). These problems are related to the productive vocabulary knowledge of L2 learners.

The role of vocabulary knowledge in receptive and productive L2 skills cannot be ignored as it is used in predicting the learners' competence. It is obvious that this would affect L2 learners in expressing their thoughts through productive skills. What makes vocabulary difficult for L2 learners can be attributed to many reasons. However, as Mobarg (1997) claims, its being an infinite system (open-ended set) unlike grammar or phonology, which are finite systems, results in difficulties for teachers to organize their teaching.

Vocabulary knowledge has many facets. Its common definition is that if an L2 learner knows a word, s/he should go beyond the ability to produce it again or give a definition of a word close to its dictionary definition (Nation, 2000; Nation & Webb, 2011). There are different hierarchies proposed framing vocabulary knowledge by Henriksen (1999), Read (2004) and Nation (2000). The most comprehensive framework, which includes nine different aspects of word's form, meaning and use (Figure 1), is the one Nation (2000) put forward and accordingly, a learner has to have receptive and productive mastery in nine different aspects to completely know a word. It is known that lack or deficit in any one of these areas of word knowledge results in misinterpretation of the message given in the writing (Folse, 2008).

Form	Spoken	Receptive	What does the word sound like?
		Productive	How is the word pronounced?
	Written	Receptive	What does the word look like?
		Productive	How is the word spelled?
	Word parts	Receptive	What parts can we recognize in this word?
		Productive	What word parts are needed to express meaning?
Meaning	Form and meaning	Receptive	What meaning does this form signal?
		Productive	What word form can be used to express this meaning?
	Concept and referents	Receptive	What is included in this concept?
		Productive	What items does the concept refer to?
	Associations	Receptive	What other words does this make us think of?
		Productive	What other words are possible to use instead of this one?
Use	Grammatical functions	Receptive	In what patterns does this word occur?
		Productive	In what patterns is this word required to use?
	Collocations	Receptive	What other words or types of words occur with this one?
		Productive	What words or types of words must we use with this one?
	Constraints on use (register, frequency, etc.)	Receptive	Where, when, and how often would we expect to encounter this word?
		Productive	Where, when, and how often can we use this word?

Figure 1. *Aspects of Vocabulary Knowledge (Nation, 2000, p. 40-41)*

To define what vocabulary knowledge is of importance and there are several definitions of it. However, this results in a necessity of measurement. The necessity to assess complete vocabulary knowledge led the researchers to create vocabulary tests and researchers have used these tests including matching and/or elicitation tasks or ranking knowledge via ordinal scales (Gonzalez, 2013). Vocabulary Knowledge Scale (VKS) (Paribahkt & Wesche, 1993) and the Word Associates Test (WAT) (Read, 1993) are among the most influential ones. However, both tests have faced criticisms of Nation and Webb (2011) and Schmitt and Ng (2011) in terms of their not being able to measure the vocabulary knowledge accurately and they also claim that in the last stage of VKS, in which the learner is asked to build up a sentence with the word given, the learner should not only know the target word, but also the other words and syntactic structure surrounding it. What Waring (2002) says about the interpretation problems of VKS is that while comparing the pre- and post-test scores of a learner, the average score is taken into consideration. Assume that the average scores of a learner from the pre-test and the post-test, consisting 18 vocabulary items, are close to each other as in the example below. According to Waring (2002) the interpretation of these scores is not possible.

Test 1 1 0 1 0 0 1 2 3 3 2 2 3 4 5 5 4 5 4 = 45/18 = 2.5

Test 2 3 1 4 5 1 0 2 2 1 4 1 3 4 5 1 5 3 4 = 49/18 = 2.7

Table 1. *Vocabulary Knowledge Scale (Paribakht & Wesche, 1993)*

Point value	Self-report categories
1 point	I have never seen this word before.
2 points	I have seen this word before, but I don't know what it means.
3 points	I have seen this word before, and I think it means …… (synonym or translation)[a]
4 points	I know this word. It means …… (synonym or translation)[b]
5 points	I can use this word in a sentence.[c]

a. Learner needs to provide a synonym in English or L1 translation

b. Learner needs to provide a synonym in English or L1 translation

c. Learner needs to complete number 4 and 5.

What Waring (2002) argues is that a total mean score cannot tell anything on which ratings for which word had changed because of the treatment, which shows a lack in validity of the measurement.

The criticism for WAT is its susceptibility to guessing (Schmitt & Ng, 2011, p.107).

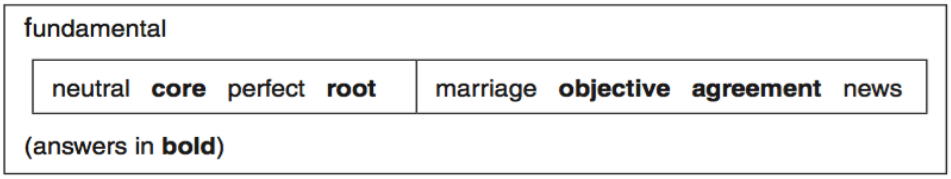


Figure 2. *WAT item*

As seen above, in this test the learners are asked to choose the word or the words with the closest meaning to the target word *fundamental* from the box on the left and choose the common collocation or collocations of the target word from the right box. This is why WAT is problematic because lexical knowledge cannot directly be concluded from the results of this test (Schmitt & Ng, 2011).

Therefore, Laufer and Nation (1995) states an analysis of a learner's written productive vocabulary capabilities can demonstrate the vocabulary knowledge of the learner as the written work consists of accurately used actual vocabulary knowledge in production.

1.1. Concept Definitions

It is important to give briefly the basic definitions of the key concepts that will be used in the thesis in order to make research questions more clear and concise. These concepts are referred in details in methodology section.

Vocabulary size is the measurement of vocabulary items from different frequency levels and the Lexical Frequency Profile is used for this purpose.

Lexical density is the proportion of the content words to the total words in a written text. It gives a percentage according to what information load of the text can be presumed. It also provides some insights into the vocabulary level of the learners.

Lexical diversity is the measure of the variety of vocabulary deployed by a speaker or a writer and provides information about productive vocabulary rather than receptive vocabulary.

What differentiates vocabulary size from the lexical diversity is that although lexical diversity calculation gives a number, which is difficult to interpret, vocabulary size shows the frequency levels of the vocabulary items used in the text and it helps to interpret the vocabulary use of the learners more easily. For example, to what extent they use high frequent vocabulary items and low frequent vocabulary items and so on.

1.2. Statement of the Problem

Writing in a second language is stated to be one of the most important skills in language learning (Jackson, 2004; Choi, 2012; Aliakbari & Boghayeri, 2014). Writing skill is a prerequisite for academic competency as it is a productive skill along with the speaking and also students produce something concrete by using these skills for teachers to evaluate. There are various studies stating that students learning a second language have difficulty and have problems in writing. Besides, as Manchon (2011) argues writing reflects the language development of learners, and that language development is a supporter of writing. Studies conducted to illuminate the issues having effects on writing quality, writing performance of the learners are on syntactic complexity (Beers & Nagy, 2007), L1 influence (Staples & Reppen, 2016), instruction types and duration (Min, 2016), linguistic proficiency including lexical density and variety, vocabulary size, grammar etc. This study will focus on the lexical density, lexical diversity and vocabulary size. While there are studies claiming that these lexical features are more related to the quality of writing performances of the learners and they provide predictive data to evaluate the learners' writing achievement in L2 (Lemmouh, 2008; Douglas, 2010); there are also claims that writing achievement is not only related to these lexical features (Mellor, 2010; Wang, 2014; Lavallee & McDonough, 2015).

According to deBoer (2014), lexical diversity could be the range of displayed vocabulary in a given text. It is seen as an essential element of evaluating a text quantitatively and this assessment has various uses for linguistic and educational research such as language acquisition, linguistic interaction, demographic language performance, and language impairment (Malvern & Richards, 2012). Lexical

diversity has been considered a predictor of learners' general language proficiency (Zareva et al. 2005), essential indicator of the quality of their writing (Laufer & Nation, 1995) and speaking (Jarvis, 2002) performances. It is widely assumed as an important quality indicator of test performance of the learners. Lexical diversity measures can be used to evaluate how a language learner can effectively integrate vocabulary into their language production, which is of greater interest to language teachers and researchers than results on tests measuring passive vocabulary (Nation, 2007). It helps language instructors to understand how language learners use diverse vocabulary items in their productive performances. This understanding provides insights for instructors to plan and guide their teaching in class.

There are different formulas to measure lexical diversity in a text. Researchers tend to agree that two measures seem to be particularly reliable, the Measure of Textual Lexical Diversity (MTLD) and vocd-D (McCarthy & Jarvis, 2010). MTLD and vocd-D analyses are available on a website (http://tool.cohmetrix.com/). The criticisms of lexical diversity measures are mainly based on the text length and MTLD is proven to be less affected by the text length (Koizumi, 2012; McCarthy & Jarvis, 2010).

Vocabulary size calculations, in this case Lexical Frequency Profile (LFP), show how many words a text includes among the vocabulary lists generated from British National Corpus (BNC) and Corpus of Contemporary American English (COCA). The BNC/COCA word family lists consists of 29 word family lists, twenty-five of which contain word families based on frequency and size data and four of which are list of proper names, marginal words (swear words, exclamations), transparent compounds and abbreviations. Vocabulary size calculations can be made on the website (http://www.lextutor.ca/vp/comp/) according to the vocabulary lists as mentioned above BNC/COCA list or General Service List of English (GSL) words (2000 most useful word families of English) and Academic Word List (AWL) (Coxhead, 2000).

The data created from 1st and 4th year ELT students' argumentative essays will be subjected to these calculations to see and highlight the productive lexical use of the students.

1.3. Significance of the Study

The results of the studies conducted on vocabulary knowledge put forward that vocabulary knowledge has two dimensions, receptive and productive (Nation, 2000; Webb, 2005). Therefore, vocabulary knowledge cannot be conceptualized only as how many words learners know (i.e. vocabulary size) but how these words are used (i.e. productive vocabulary) (Meara, 2002; Schmitt et. al., 2010). There are various ways to measure the active and passive vocabulary knowledge of the learners, such as Vocabulary Knowledge Tests, LFP, lexical density and lexical diversity calculations.

With this study, the aim is to examine the current situation of Turkish ELT students by analyzing their essays in terms of vocabulary size, lexical density and lexical diversity. Moreover, this study aims to find out the relationship of these concepts with the students' writing scores.

How L2 learners use productive vocabulary knowledge in their writing is the question. Lexical richness, encompasses lexical density, lexical diversity and vocabulary size, is among under-researched topic (Skehan, 2009). There are few studies conducted on vocabulary knowledge and its impact on L2 skills or academic competence (Karakoç, 2016; Yüksel, 2012). Therefore, the results of this cross sectional study are believed to contribute to the literature on the aforementioned issues together with some implications for writing courses and vocabulary teaching.

1.4. Aims and Research Questions

The aim of the study is to find out the vocabulary size, lexical density and lexical diversity of Turkish EFL students in their argumentative essays and to compare these with their writing scores. It is assumed that large vocabulary size, lexical density, and lexical diversity may lead to higher scores.

The ELT students are accepted to the department according to their university entrance exam scores from a foreign language proficiency examination (YDS) consisting of 80 multiple-choice questions. Thus, students studying at the same university have, more or less, similar proficiency levels because each university has a score range to accept students.

Productive aspect of L2 learning is ignored in Turkey. Turkish high school education does not give importance in productive language skills and YDS consists of reading comprehension and grammar questions. It is expected that a 4-year education in English would increase the proficiency levels of students and hence their productive vocabulary knowledge. Therefore, this thesis aims to compare the 1st and 4th year students' productive vocabulary use and its effect on their writing scores to find out whether a 4-year English medium education has led to an improvement in their productive vocabulary use in their writing in our sample. The following research questions are posed:

1. What are the vocabulary size, lexical density and lexical diversity in 1st and 4th year Turkish ELT student essays?
2. Are there any differences between 1st and 4th year ELT students essays in terms of
    a. vocabulary size?
    b. lexical density?
    c. lexical diversity?
3. What is the correlation between
    a. vocabulary size and
        i. the 1st year students' argumentative essay scores?
        ii. the 4th year students' argumentative essay scores?
    b. lexical density and
        i. the 1st year students' argumentative essay scores?
        ii. the 4th year students' argumentative essay scores?
    c. lexical diversity and
        i. the 1st year students' argumentative essay scores?
        ii. the 4th year students' argumentative essay scores?
4. To what extent do vocabulary size, lexical density and lexical diversity account for the essay scores and vocabulary subsection scores of the 1st and 4th year students?

2. LITERATURE REVIEW

This chapter will, firstly, give the basic definitions of LFP, lexical density and lexical diversity. Since these concepts accounts for productive vocabulary knowledge of the students, receptive and productive vocabulary knowledge will also be covered.

Secondly, the relationship between vocabulary and writing in L2 will be discussed briefly.

Finally, theoretical and empirical studies dealing with the LFP, lexical density and lexical diversity, related to L2 writing will be presented respectively to form the basis of the current study.

2.1. Receptive and Productive Vocabulary Knowledge

Research on second language acquisition has shifted from grammar towards vocabulary after the introduction of the Natural Approach (Krashen, 1989), in which comprehensible and meaningful input were put forward instead of structural accuracy. Later on, the Lexical Approach (Lewis, 1993) was introduced. What Lewis (1993) says is that vocabulary plays an important role in language acquisition and it is the core of language proficiency because proficiency mainly requires understanding and producing lexical phrases or chunks.

Traditionally, knowing a word is defined as to be able to recognize the form and understand the meaning of a word when encountered. However, there are dimensions of knowing a word according to Henriksen (1999), Read (2004) and Nation (2001) as mentioned in the introduction chapter. Richards (1976) also defines the nature of vocabulary knowledge and what knowing a word means with these eight assumptions

1. Vocabulary keeps expanding even in adulthood unlike syntax, which shows little development in adult life.
2. Knowing a word means knowing how probable to encounter that word in written or spoken discourse. For many words we also know the kind of words most likely to be found related to the word.
3. Knowing a word is to be aware of the limitations of its usage according to variations of function and situation.
4. Knowing a word means knowing the syntactic behavior associated with the word.

5. Knowing a word entails knowledge of the underlying form of a word and the derivations that can be made from it.
6. Knowing a word entails knowledge of lexical network of the word and other words in the language.
7. Knowing a word means knowing the semantic value of a word.
8. Knowing a word means knowing many different meanings associated with a word.

It can be seen among the assumptions that Richard (1976) incorporates morphological and syntactic features of a word together with its frequency and register into the definition of vocabulary knowledge. The missing part of the definition is the dimensions of vocabulary knowledge: receptive and productive vocabulary knowledge (Nation, 1990). The validity of this distinction resembles to the distinction between receptive language skills of listening and reading, and productive language skills of writing and speaking (Crow, 1986). Basically, receptive vocabulary knowledge involves perceiving and recognizing a word while listening or reading and understanding its meaning; productive vocabulary knowledge, on the other hand, is defined as to be able to express and convey a message via speaking or writing and to produce an appropriate spoken or written form to do this (Nation, 2000, p.38).

There is not a clear border between receptive and productive vocabulary knowledge as there are objections claiming that while using receptive knowledge, productive knowledge is also required and vice versa (Milton, 2007). However, it is not impossible to classify characteristic aspects of receptive and productive vocabulary knowledge.

Nation (2000, p.41-42) arrays these features for receptive vocabulary knowledge

1. To be able to recognize the word when heard,
2. To be familiar with the written form to recognize while reading,
3. To be able to recognize the affixes and to relate these to its meaning,
4. To know that the word itself signals a particular meaning,
5. To know what the word means in the particular context,
6. To know that there are other related words,
7. To be able to recognize that the word used correctly in the sentence it occurs,
8. To be able to know possible collocations of the word,
9. To be able to know whether it is common or pejorative.

The features of productive vocabulary knowledge are as follows according to Nation (2000, p.42)

1. To be able to say it with correct pronunciation with correct stress,
2. To be able to spell it correctly,
3. To be able to use right word parts in appropriate forms,
4. To be able to produce the word in various contexts,
5. To be able to produce synonyms and antonyms for the word,
6. To be able to use the word correctly in the original context,
7. To be able to produce collocations of the word,
8. To know where, when and how often to use the word.

This distinction resulted in new problems; such as "how many words one must be able to recognize automatically irrespective of context in order to be able to use the higher level processing strategies with success" (Laufer, 1997, p.23), how many words a native speaker knows, how many words there are in the target language (Nation & Waring, 1997), what types of words there are in the target language and most importantly how to measure this knowledge.

Along with the technological innovations, corpus-based studies have become widespread on vocabulary studies, which enables much more accurate and detailed description (Biber, 2006). Corpus-based studies contributed to the literature by providing a great deal of linguistic information from lexical frequency to collocations, chunks and to lexical diversity.

Many studies have been conducted to find out more about vocabulary and its bounds with other language skills; the distinction between receptive and productive knowledge, knowledge and use (Henriksen, 1999); the interconnection between vocabulary knowledge and language proficiency with respect to reading (Hu & Nation, 2000), writing (Begriche, 2013; Putra, 2014; Rudy, 2013; Yang, 2015), listening (Ling, 2015; Vandergiff & Baker, 2015; Wang, 2015) and speaking (Fhonna, 2014; Khotimah, 2014; Tahir, 2015; Perez Manzanilla & Diaz Cabrera, 2014); the effect of word frequency and word lists in vocabulary learning (Coxhead, 2000); the effect of tasks (Laufer & Hulstijn, 2001); the comparisons of vocabulary learning strategies (explicit vs. implicit and incidental vs. intentional) (Ellis, 1994; Ellis & He, 1999); and how to test vocabulary knowledge (size, depth, receptive, productive) (Bogaards, 2000; Laufer & Nation, 1995, 1999; Nation, 2001; Wesche & Paribakht, 1996).

The recent studies on vocabulary in Turkish context are mainly about vocabulary teaching techniques and vocabulary learning strategies. Some of these studies are as follows. Çelik and Toptaş (2010) studied on vocabulary learning

strategy use of Turkish EFL learners; Kök and Canbay (2011) studied on vocabulary consolidation strategies; Başöz (2014) investigated the effectiveness of computer assisted instruction on vocabulary achievement; Aitkuzhinova et. al. (2016) examined the effects of teaching vocabulary to Turkish young learners in semantic clustering way through digital storytelling. Studies focusing on the lexical diversity, lexical density, vocabulary size, dimensions of vocabulary knowledge and their effects on language skills are very limited. Topkaraoğlu (2013) investigated vocabulary size and lexical depth in lexical competence. The experimental group to which Topkaraoğlu (2013) implemented vocabulary teaching activities for 14 weeks outperformed the control in Vocabulary Level Tests (VLT), Productive Vocabulary Level Tests (PVLT). Karakoç (2016) demonstrated the impact of vocabulary knowledge on reading, writing and proficiency scores of B2.2 level Turkish Prep-School students at a state university. She used vocabulary knowledge tests to measure the vocabulary knowledge of the students and looked at the relationship between the receptive vocabulary knowledge and reading performance and the relationship between productive vocabulary knowledge and their writing performance using their scores on VKTs, a reading exam and a writing exam.

There are few doctoral dissertations investigating lexical aspects of the learners' written texts (Yüksel, 2012; Ünaldı, 2011). Ünaldı (2011) conducted a comparative study focusing on the lexical networks Turkish EFL learners. 49 essays written by Turkish EFL learners and 100 essays written by native speakers were compared in terms of lexical networks, lexical cohesion and syntactic features in learners' texts as the first objective of the study. What Yüksel (2012) investigated was the general and academic lexical competence and performance of Turkish ELT students. She conducted the study with 371 students. Through multiple test approach, the receptive general and academic vocabulary size of the students were measured and the general and academic lexical performance of the students were determined by calculating the LFPs and lexical diversity in their argumentative essays. What is striking in this study is that the students who were found to have large vocabulary size and depth could not reflect their receptive vocabulary knowledge in essay writing task.

2.2. Measurement of Productive Vocabulary Knowledge

To be able to measure receptive vocabulary knowledge is not sufficient alone to provide a satisfactory description of the total vocabulary knowledge as vocabulary knowledge has many facets (Zareva, 2005). Therefore, the attention shifted towards to measure the productive vocabulary knowledge of the learners. It was not easy to conduct studies on productive vocabulary measures since the nature of productive vocabulary is context-specific (Lee & Muncie, 2006).

Laufer and Nation (1995) put forward the productive version of Vocabulary Levels Tests, including a sentence with a missing word some letters of which are provided. This productive version of the level tests (PVLT) has been reported valid and can make the comparisons among the learners of different proficiency levels (Laufer, 1998).

Laufer and Nation (1995) proposed another measurement way, LFP, measuring the amount of vocabulary from different frequency levels that learners used in their writings.

Webb (2008) used a translation test to assess the productive vocabulary knowledge by stating that the PVLT actually measures the receptive vocabulary knowledge because given letters may help the learners recognize the word.

Meera and Fritzpatrick (2000) alternatively proposed Lex30, a kind of word association test, in which the learners are given a list of stimulus words and asked to produce responses to the stimuli by claiming that it is easy to apply and it does not require much time.

Despite these alternatives, the LFP is the most commonly adopted measurement of productive vocabulary knowledge used in vocabulary research to analyze the vocabulary use in learners' written works.

2.2.1. Lexical frequency profile (LFP)

Lexical Frequency Profile (LFP) (Laufer & Nation, 1995) measures the amount of vocabulary from different frequency levels, vocabulary size, which learners used in their writings. It can also be defined as a tool measuring the relative

proportion of words from different frequency levels. It was first developed to assess the lexical difficulty of an L2 reading text, but it is widely used to calculate the lexical richness based on word frequency lists (Utku, 2014). LFP calculates the relative proportion of words in the first 1,000 most frequent words, the second 1,000 most frequent words (based on General Service List (GSL)), the 570 most frequent academic words (Academic Word List (AWL), also known as University Word List (UWL)), and words that are not in any of the lists (beyond 2k).

The word lists are important sources for vocabulary research and they are prepared with the idea in mind that some words are more frequent than others. The oldest one and on which the LFP based on is GSL, which was compiled by West in 1953 (as cited in Laufer & Nation, 1995). GSL contains about 2000 base words. 165 word families in the lists are function words and the rest consists of content words. What makes GSL more useful than a simple frequency count is that each word's different parts-of-speech and different meaning senses are listed (Yüksel, 2012). Despite its age, GSL is still valid according to Nation (2004), who questioned the coverage of GSL against the BNC.

Xue and Nation (1984) constructed the UWL and Coxhead (2000) compiled the AWL by adding the words that are not within the scope of 2000 words of GSL.

In this study, the essays will be analyzed in terms of students' percentages of using AWL and off-list vocabulary items as an aspect that shows their productive vocabulary knowledge.

Meara (2005) and Meara and Bell (2001) criticized the LFP for it requires texts over 200 words and for it does not work well with low-level learners. However, they did not propose a new way or method to analyze written texts. Laufer (2005) responded this criticism quite convincingly and that is one of the reasons why it is still preferred to determine the vocabulary size of the texts.

2.2.2 Lexical density

Lexical density (LD) is a measure of lexical richness like the LFP. It measures the proportion of lexical words (content words) to the total number of words in a written text. If the proportion of content words in a text is higher, in other words, proportion of function words is lower, this means that the text includes more information (Johansson, 2008). According to Ure, if an item does not have lexical

properties, it can be described in terms of grammar, which means that these words have a grammatical-syntactic function (as cited in Johansson, 2008).

$$LD = \frac{\text{Number of lexical tokens} \times 100}{\text{Total number of tokens}}$$

Figure 3. *Lexical Density Formula (Laufer & Nation, 1995, p.309)*

Lexical density is calculated by the formula above and Ure ( )concluded that majority of the spoken texts have a lexical density of below 40%, whereas majority of the written texts have a lexical density of 40% or higher (as cited in Johansson, 2008). It must be noted that lexical density is dependent on the syntactic and cohesive properties of the written text. That is, the number of function words influences LD measure, which affects its validity. Therefore, there is little relationship between LD and the quality of writing (Engber, 1995). In this study, this assumption will also be checked by calculating the correlation between LD and the students' essay scores which are given by two raters according to a writing evaluation rubric (ESL Composition Profile). While scoring, the raters did not take LD into consideration. Therefore, the correlation between LD and given scores will be a verification or refutation to Engber (1995).

2.2.3. Lexical diversity

Lexical diversity is considered an end-product of a language and it is the measure of the variety of vocabulary deployed by a speaker or a writer and also it is regarded as to be indicator of higher linguistic skills, speaker competence (Avent & Austermann, 2003, Carrel & Monroe, 2004; Grela, 2002). Lexical diversity is a quantitative measure of a written text, providing the information about productive vocabulary rather than receptive vocabulary.

It is intuitively thought that the vocabulary used by a writer is much more diverse than the vocabulary used by a college student in compositions or lexical diversity of 1st graders is normally less than post-graduate students, which in turn is less diverse than one who has completed a PhD. Using a quantifiable measure can be used to test these opinions and intuitions on the quality of a text because quantifiable

knowledge provides an objective and verifiable approach to evaluate the texts (McCarthy, 2005). Consequently, predictions can be formed and the texts can be judged scientifically with quantifiable measures, among one of which is lexical diversity. Lexical diversity cannot give answers to every question regarding text quality alone, but it offers a useful tool for researchers and teachers.

Although lexical diversity itself is not a disputed concept in linguistics, its measurement has led to questions, because of which various measurement forms, such as Type-Token Ratio (TTR), the MTLD, D (vocd-D), have been put forward (McCarthy, 2005). The dispute on these measurement forms is mainly based on the text length which is said to affect the validity of lexical diversity measure as can be concluded in longer texts the possibility of new words appearing is reduced (Malvern et al., 2004; Jarvis, 2002). The longer the text is, the less reliable the results get.

Therefore, a necessity to be able to measure lexical diversity in a more valid way, new measurement forms have been proposed claiming that the sensitivity one after another. Even though, each form claims to be more valid and less affected by the text length, Vocd-D and MTLD are considered more robust approaches to LD assessment (Malvern et al., 2004; McCarthy & Jarvis, 2007, 2010).

2.2.4. The Vocd-D

One of the approaches to quantify lexical diversity is D parameter, firstly created by Malvern and Richards (1997). Then it was replaced with a newer calculation again by Malvern et al. (2004) with more stable and empirically smoothed via random text samplings (McKee et al., 2000). The operation of D measure is not easy; therefore, a software is needed for the measure. The name Vocd-D comes from the name and a command of the software. In order to obtain a D score the following vocd procedures are followed (McCarthy, 2005):

1. Random samples of text are compiled without alterations.
2. Sample sizes are chosen from 35-50 tokens.
3. 100 samples of every token size is taken to get a mean score.
4. A D score is produced for a TTR of each participant.
5. An average D-score is taken among the prior D-scores.
6. This process is done for 3 times to get the final average D score.

And also these four criteria should be considered and followed in measuring the lexical diversity:

1. Lexical diversity measure should be text length independent.
2. Lexical diversity measure needs to produce scores widely ranging between low diversity scores and high diversity scores, which shows the sensitivity.
3. Lexical diversity measure should be computed without need for lemmatization.
4. Lexical diversity measure should be able to solve a text sequentially. (McCarthy, 2005)

Although there are reliability issues about lexical diversity measures, in terms of text length, deleting the function words before analysis or not, the MTLD and Vocd-D are accepted more rigid and successful than the others (McCarthy & Jarvis, 2007, 2010; Malvern et al., 2004; Crossley & McNamara, 2010). Thus, in this study lexical diversity of the sample texts were calculated by Vocd-D.

2.3. Empirical Studies

Studies on lexical diversity have primarily focused on how to measure lexical diversity; therefore, the relationship between lexical diversity and writing proficiency has not received much attention. Some of the studies using lexical diversity, lexical profile and lexical density calculations to find out their relationship to writing quality are listed below in a table.

As can be seen in the table below, in most of the studies lexical diversity is proven to have impact on students' writing scores and writing quality. While Lemmouh (2008); Douglas (2010); Gonzalez (2013) and Karakoç (2016) reported statistically significant effect of lexical diversity on writing scores; Mellor (2010); Wang (2014) and Lavallee and McDonough (2015) reported that the effect was relatively low and not significant.

Lemmouh (2008) studied with 37 Swedish advanced learners of English. He compared the lexical richness scores of their essays with their essay grades, course grades and vocabulary knowledge as measured by three tests. Moreover, he administered a 14-item questionnaire to the teachers rating the essays to find out whether the teachers give importance to lexical richness while grading. The results showed that there was a relationship between overall course grade and use of

advanced vocabulary in essays; teachers primarily based their grades on grammar and content rather than lexical features.

In his study with non-native university students, Douglas (2010) looked at the effect of lexical richness on these students' academic success. He defined writing skill as representative of academic success of non-native students and therefore he used a writing test to evaluate the academic success of them. In this study the lexical richness was calculated with lexical profiling measurement. He concluded that lexical richness had an important role in writing assessment, and university level writing quality was a predictor of academic success.

Gonzalez (2013), examined the effect of vocabulary size and lexical diversity in advanced non-native speakers' and native speakers' academic compositions on their writing scores. The lexical diversity of 172 essays were measured with MTLD and the vocd-D; and vocabulary size was measured by word frequency means. The essays were rated by three raters according to TOEFL IBT Independent Writing Rubric. The results showed that lexical diversity had more effect on writing scores than vocabulary size. She also found that native speakers' lexical diversity and vocabulary size profiles were higher than non-native speakers' and these lexical profiles had a significant difference among the individual score levels of the rubric.

Karakoç (2016), aimed to demonstrate the impact of vocabulary knowledge on reading, writing, and proficiency scores of Turkish preparatory school students. She used LFP to determine the lexical level of the student essays. She found out there was a significant relation between that lexical level of the student essays and their productive vocabulary knowledge.

Mellor (2010), investigated the writing quality of L2 learners with regard to essay length and lexical diversity. In his study, he used a dataset collected from 34 students. A native speaker rated the essays. He used several measures for lexical diversity and also compared these measures. According to his results, essay length predicted the essay ratings more accurately than lexical diversity.

Wang (2014), directly focused on the relationship between lexical diversity and EFL writing proficiency. He used forty-five texts written by Chinese high school students in a national based English test. Lexical diversity was measured by TTR and vocd-D. His results showed that lexical diversity of higher proficient students did not differ significantly than lower proficient students; and there were not any statistical significant relationship between lexical diversity and students' writing scores.

Lavallee and McDonough (2015) compared the cause-effect essays of 94 English for academic purposes (EAP) students. Three raters graded the essays by using a holistic rubric and five lexical features of the essays (AWL word use, content word frequency, word familiarity, imagability, and lexical diversity). The results did not show a statistically significant correlation between essay ratings and lexical features.

On the other hand, LFP was regarded as reliable and stable measure of vocabulary size of the students by Laufer and Nation (1995) and the results of Lemmouh (2008); Yüksel(2012); Karakoç (2016) and Signes and Arroitia (2015) stated the same.

Essay length is also among the variables that have an impact on writing score as Mellor (2010) claims. Students' not being able to transfer their receptive vocabulary knowledge into productive vocabulary use in their writings is another finding of the studies (Lemmouh, 2008; Yüksel, 2012; Karakoç, 2016).

That native speakers' lexical diversity and vocabulary size differ significantly from non-native speakers' as Gonzalez (2013) discovered in her study forms the base of the idea that students with higher language proficiency levels know and can use more vocabulary items in productive language skills.

Table 2. *Empirical Studies*

	Study	Aim	Tools	Related Findings
Laufer & Nation (1995)	Vocabulary Size and Use: Lexical Richness in L2 Written Production	To find out if there is a correspondence between the vocabulary size of intermediate learners as reflected in their writing and a more direct measure of vocabulary size.	LFP	• It is possible to get a reliable and stable measure of lexical richness in two writings of the same learner. • The LFP can discriminate between learners of different proficiency levels. • The LFP has a correlation with an independent measure of vocabulary size.
Lemmouh (2008)	The Relationship Between Grades and the Lexical Richness of Student Essays	To examine the relationship between Swedish university students' essay grades and lexical richness.	• VLT • PVLT • LFP	• The LFP can be used as a diagnostic tool to identify students with poor vocabulary knowledge. • Students using more academic and low-frequency vocabulary, determined by the LFP, are more successful writers.
Douglas (2010)	Non-Native English Speaking Students at University: Lexical Richness and Academic Success	To measure the lexical richness of non-native and native English speaking students and compare them to academic outcomes.	• TTR • Effective Writing Test (EWT)	• Lower measures of lexical richness seemed to affect the assessment of writing exams. • Students with higher lexical richness performed better in EWT.
Mellor (2010)	Essay Length, Lexical Diversity and Automatic Essay Scoring	To investigate if essay length and lexical diversity together may replace essay ratings. To determine which lexical diversity measure is better.	• TTR • Guiraud's Index • Yule's K • VocD • Hapax	• Essay length was found to be the dominant predictor of essay ratings, while lexical diversity had a relatively little effect. • Advanced Guiraud was the best in clearly identifying the high rated and low rated

Table 2. (Continued) *Empirical Studies*

			• Advanced Guiraud	essays. • Students have large vocabulary size and depth (receptive vocabulary knowledge).
Yüksel (2012)	Cross-sectional Evaluation of Turkish ELT Majors' General and Academic Lexical Competence and Performance	To evaluate the general and academic lexical competence and performance of Turkish ELT students.	• VLT • WAT • Test of Academic Vocabulary • TTR • LFP	• Students cannot use their receptive vocabulary knowledge in production. • LFP is reliable in assessing lexical diversity in students' argumentative essays. • Students' vocabulary knowledge increases across the years but their lexical competence and performance do not increase in the same manner.
Gonzalez (2013)	The Intricate Relationship Between Measures Of Vocabulary Size And Lexical Diversity As Evidenced In Non-Native And Native Speaker Academic Compositions	To find out to what extent vocabulary size and lexical diversity contributes to writing scores on advanced non-native and native speakers' academic compositions.	• MTLD • VocD • CELEX (Word Frequency Means)	• Lexical diversity has more impact on writing score than vocabulary size. • Native speakers' lexical diversity and vocabulary size profiles significantly differ from non-native speakers'. • Vocabulary size has moderate correlation with lexical diversity, which shows that mid-size vocabulary may be more important in writing than using less frequency vocabulary.
Wang (2014)	The Relationship between Lexical Diversity and EFL Writing Proficiency	To explore the relationship between lexical diversity and EFL writing proficiency	• Chinese National Matriculation English Writing Test • TTR	• There is not a significant relationship between the lexical diversity measures and the students' writing scores. • Lexical diversity of high graded students does not differ from the lexical diversity

Table 2. (Continued) *Empirical Studies*

			• VocD • Two argumentative essays written at the beginning and at the end of the programme.	of low graded students.
Mazgutova & Kormos (2015)	Syntactic and Lexical Development in an Intensive English for Academic Purposes Programme	To show the syntactic and lexical development of L2 learners' academic writing after a one-month intensive English for Academic Purposes programme.	• MTLD • CELEX	• The students showed improvement with regard to lexical diversity in their essays. • Students began using more advanced vocabulary, a characteristics of academic context after the programme.
Lavallee & McDonough (2015)	Comparing the Lexical Features of EAP Students' Essays by Prompt and Rating	To examine the relationships among the lexical features (AWL word use, content word frequency, word familiarity, imagability, lexical diversity) of students' essays, essay writings, and writing prompts	• MTLD • Coh-Metrix • AWL • TOEFL Writing Rubric	There is no significant correlation between essay ratings and lexical features.
Signes & Arroitia (2015)	Analysing Lexical Density and Lexical Diversity in Unviersity Students' Written Discourse	• To determine if writing quality assessment based on LFP is valid. • To see if there is development in lower level students' writings in a semester.	• LFP • Textalyser (LD analysis)	LFP provided stable measure of lexical richness in two writings of the same learner.
Karakoç (2016)	The Impact of Vocabulary Knowledge on Reading, Writing and Proficiency Scores of B2.2 Level Turkish Students	To demonstrate the multidimensional nature of vocabulary knowledge development and its relation to the students' reading and writing performance together with the general English ability.	• LFP • VKT • Reading and Writing exam	• The students' receptive vocabulary knowledge is larger than the productive vocabulary knowledge. • Vocabulary knowledge contributes significantly to reading and writing performances of the students. • There is a correlation between the lexical level of the student essays and students' productive vocabulary knowledge.

3. METHODOLOGY

A cross-sectional, quantitative, correlational and descriptive research method was used in this study. The first aim of this study is to present the vocabulary size, lexical density and lexical diversity of 1st and 4th year Turkish ELT students' argumentative essays. The second aim is to look at the difference of 1st year students' essays and 4th year students' essays in terms of these lexical features. The third aim is to find out if there is a correlational relationship between these lexical features and students' argumentative essay scores for the 1st year essays and the 4th year essays separately. And the final aim is to find out to what extent these lexical features explain the writing scores.

3.1. The Setting and the Data

The data were collected from Anadolu University ELT Department. The students have to get sufficient scores on English language examination, a part of a standardized university entrance test carried out by Student Selection and Placement Center affiliated by the Council of Higher Education in Turkey, to enroll in the program. In this examination, there are 80 multiple-choice questions mainly devoted to reading comprehension and grammar (vocabulary knowledge, sentence completion, translation, reading passages, paraphrasing, paragraph completion, irrelevant sentence in a passage).

In ELT department, the students have two compulsory writing courses (Written Communication, Academic Writing and Report Writing) in the first year of their four-year education. In these courses the students are taught how to write a paragraph, how to write an essay, essay types (Opinion, Cause and Effect Analysis, Summary-Analysis, Problem-Solution, Argumentative), and APA style.

To be able to enroll the ELT departments of universities, students are required to get pre-determined scores. For example, in 2015 the base point for the ELT department of Anadolu University was 418,598. The language proficiency of the students are not the only variable in these scores, however, by looking at their scores calculated mainly according to their YDS exam performance it can be assumed that the proficiency levels of the students are more or less close to each other.

Moreover, in Anadolu University, students have to pass the preparatory school, which has an exit criterion according to a certain proficiency level. The preparatory school uses the Global Scale of English (GSE). Students with A level have to pass a proficiency examination to enter their department. That proficiency exam consists of a multiple-choice exam to assess reading and language use, a listening exam, a speaking exam and a writing exam. Students have to get 60 or more out of 100 as the average of all these exam scores to pass the preparatory school. This also provides evidence for the proficiency levels of the students. Consequently, it can be concluded that they all enter the faculty with similar proficiency levels.

The data include 309 argumentative essays written by 165 first year and 144 fourth year students. The students were asked to write an argumentative essay with the prompt below. The topic was selected from the Louvain Corpus of Native English Essays (LOCNESS). It is a corpus of native English essays made up of 324,304 words in total. As LOCNESS comprised of argumentative essays in general, argumentative essay and the topic below were chosen instead of other essay genres.

Write a well-developed argumentative essay on the topic below:

Technology and Imagination

Some people say that in our modern world, dominated by science, technology and industrialization, there is no longer a place for dreaming and imagination. Discuss your opinion about this statement. (See Appendix I)

3.2. Instruments

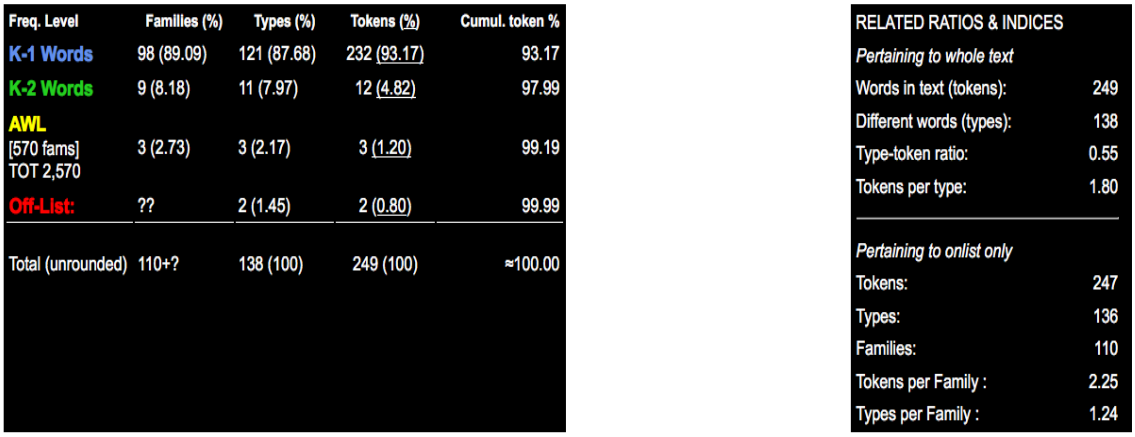
In order to analyze the data and find answers to the research questions the LFP was used to obtain an estimate productive vocabulary size of the students, the vocd-D was used for lexical diversity of the essays, and lexical density formula was used.

In order to evaluate the students' essays ESL Composition Profile was used. It is an analytical writing rubric consisting of the following subsections: content, organization, discourse markers, vocabulary, sentence construction, and mechanics (see Appendix II).

3.2.1. Lexical Frequency Profile

In LFP the text was assessed according to Nation's (1986) word lists on a website (http://www.lextutor.ca/vp/comp/). The output shows the number and the percentage of word types and word tokens from the text (Meara, 2005). Laufer and Nation (1995) stated that the LFP is reliable and valid as it correlates well with an independent measure of vocabulary size. LFP gives the proportion of the words according to the first 1,000 most frequent words, the second 1,000 most frequent words (based on GSL), the 570 most frequent academic words AWL, and words that are not in any of the lists (beyond 2k).

In this study, in order to determine the vocabulary size of the students the LFP analysis were made online on the website given above (see Appendix figures below for a sample output) and the percentages stating the AWL use and beyond 2k use are taken into consideration because the students are regarded as having the first 2000-word knowledge because of their assumed proficiency.



Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. token %
K-1 Words	98 (89.09)	121 (87.68)	232 (93.17)	93.17
K-2 Words	9 (8.18)	11 (7.97)	12 (4.82)	97.99
AWL [570 fams] TOT 2,570	3 (2.73)	3 (2.17)	3 (1.20)	99.19
Off-List:	??	2 (1.45)	2 (0.80)	99.99
Total (unrounded)	110+?	138 (100)	249 (100)	≈100.00

RELATED RATIOS & INDICES	
Pertaining to whole text	
Words in text (tokens):	249
Different words (types):	138
Type-token ratio:	0.55
Tokens per type:	1.80
Pertaining to onlist only	
Tokens:	247
Types:	136
Families:	110
Tokens per Family :	2.25
Types per Family :	1.24

Figure 4. *The LFP analysis sample output*

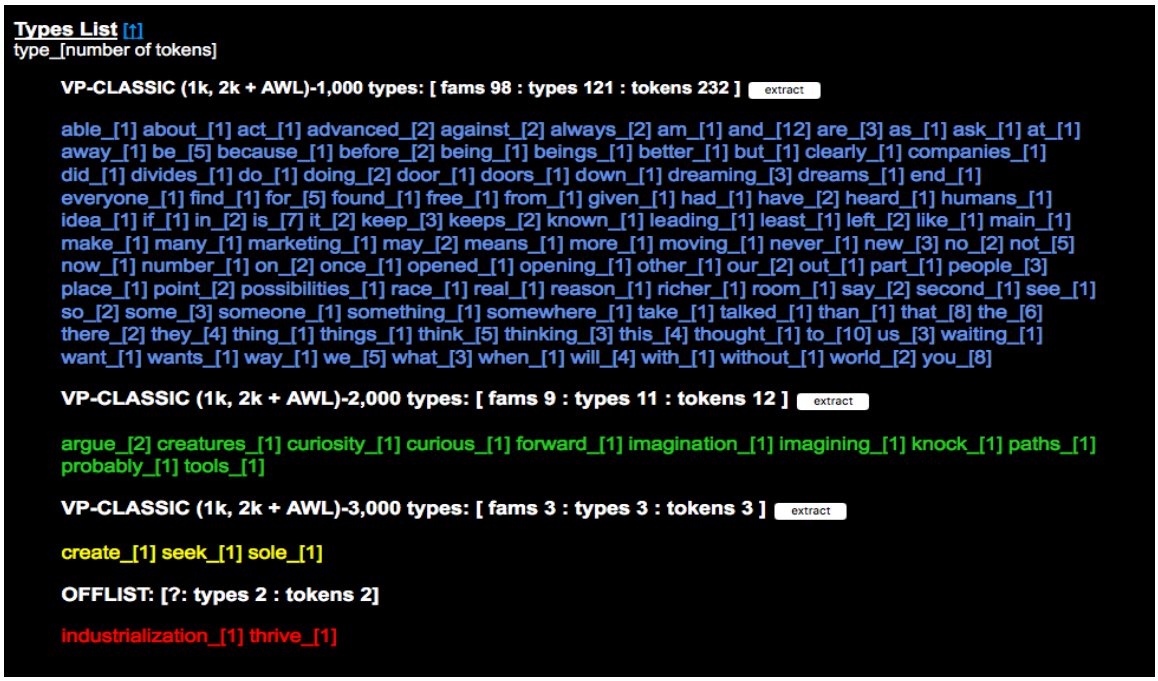


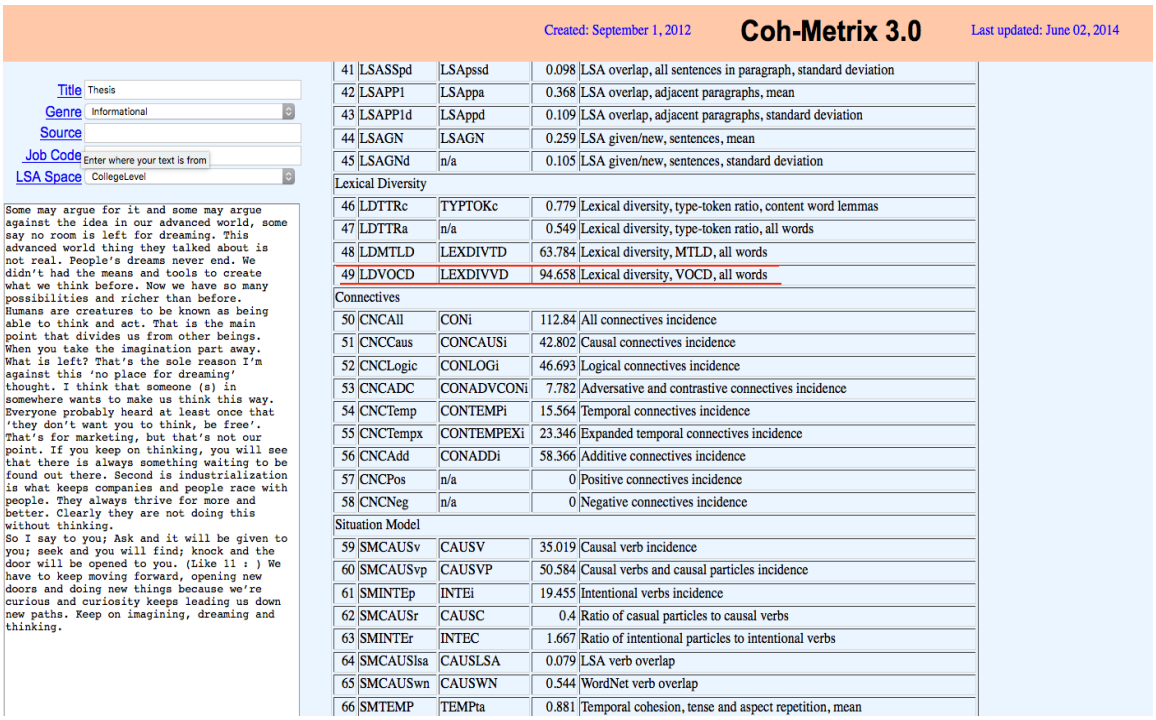
Figure 5. *The LFP analysis sample output*

3.2.2. The Vocd-D

The Vocd-D is an index evaluating the texts' lexical diversity (see 2.2.3.).

The Vocd-D does not calculate the lexical diversity in the texts having less than 100 words. The calculation was done online on a website providing lexical analyses of a text (http://tool.cohmetrix.com/).

Lower Vocd-D score shows less diversity and higher Vocd-D score shows high diversity in the texts the students wrote.



41	LSASSpd	LSApssd	0.098	LSA overlap, all sentences in paragraph, standard deviation
42	LSAPP1	LSAppa	0.368	LSA overlap, adjacent paragraphs, mean
43	LSAPP1d	LSAppd	0.109	LSA overlap, adjacent paragraphs, standard deviation
44	LSAGN	LSAGN	0.259	LSA given/new, sentences, mean
45	LSAGNd	n/a	0.105	LSA given/new, sentences, standard deviation
Lexical Diversity				
46	LDTTRc	TYPTOKc	0.779	Lexical diversity, type-token ratio, content word lemmas
47	LDTTRa	n/a	0.549	Lexical diversity, type-token ratio, all words
48	LDMTLD	LEXDIVTD	63.784	Lexical diversity, MTLD, all words
49	LDVOCD	LEXDIVVD	94.658	Lexical diversity, VOCD, all words
Connectives				
50	CNCAll	CONi	112.84	All connectives incidence
51	CNCCaus	CONCAUSi	42.802	Causal connectives incidence
52	CNCLogic	CONLOGi	46.693	Logical connectives incidence
53	CNCADC	CONADVCONi	7.782	Adversative and contrastive connectives incidence
54	CNCTemp	CONTEMPi	15.564	Temporal connectives incidence
55	CNCTempx	CONTEMPEXi	23.346	Expanded temporal connectives incidence
56	CNCAdd	CONADDi	58.366	Additive connectives incidence
57	CNCPos	n/a	0	Positive connectives incidence
58	CNCNeg	n/a	0	Negative connectives incidence
Situation Model				
59	SMCAUSv	CAUSV	35.019	Causal verb incidence
60	SMCAUSvp	CAUSVP	50.584	Causal verbs and causal particles incidence
61	SMINTEp	INTEi	19.455	Intentional verbs incidence
62	SMCAUSr	CAUSC	0.4	Ratio of casual particles to causal verbs
63	SMINTEr	INTEC	1.667	Ratio of intentional particles to intentional verbs
64	SMCAUSlsa	CAUSLSA	0.079	LSA verb overlap
65	SMCAUSwn	CAUSWN	0.544	WordNet verb overlap
66	SMTEMP	TEMPta	0.881	Temporal cohesion, tense and aspect repetition, mean

Figure 6. *Lexical diversity (Vocd-D analysis sample output)*

3.2.3. Lexical Density

Lexical density of the essays was calculated online. High density shows there are more content words in the essay than the function words and vice versa. LD analyses of the essays were done and sample output is given below.

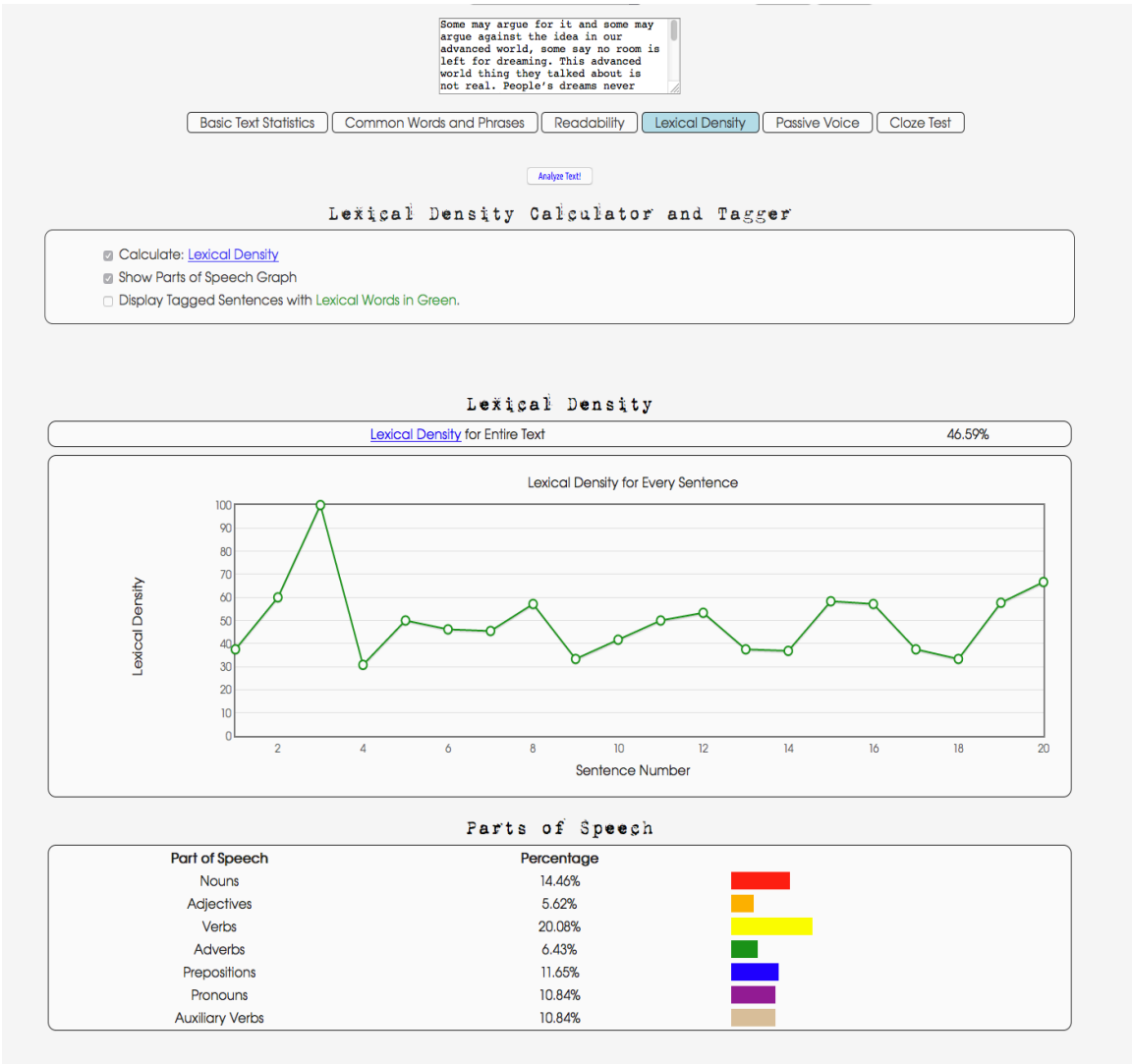


Figure 7. *Lexical density analysis sample output*

3.2.4. ESL Composition Profile

In order to evaluate the students' essays an analytical writing rubric, adapted from ESL Composition Profile, was used (see Appendix II). We chose this rubric as the writing papers of the students at Anadolu University, School of Foreign Languages are still evaluated with this rubric. It has six subsections: content, organization, discourse markers, vocabulary, sentence construction, and mechanics, ranging from *excellent to very good* to *very poor.*

The maximum and minimum points in the sections are as follows:

- Content: 25-5
- Organization: 15-2
- Discourse Markers: 10-1
- Vocabulary: 15-2
- Sentence Construction: 30-6
- Mechanics: 5-2

3.3. The Data Collection Procedure

The study was conducted in the second semester of 2015-2016 academic year. The essays were collected from the students during their class hours with the permission of their teachers. 1st year students wrote the essays in their writing courses and 4th year students wrote the essays in their translation courses. It was voluntary for students to participate to the study and they signed a consent form (see Appendix IV) before starting to write the essays. They were informed clearly about the procedure and they were said that the data collected from them were going to be used in creating a data set. They were given 60 minutes to write the essays and no dictionary was allowed during the process.

The essays were then typed into digital files and were sent to the raters for the evaluation process. The raters were two experienced English instructors who both work at state universities and have been teaching English for six years. They both have taught writing and they are also familiar to the ESL Composition Profile. They rated all the papers according to the subsections of the rubric score bands. In order to get the overall scores of the students, the average score of the two raters were taken. The inter-rater reliability of the raters was found to be *.941* for overall scores, which shows high consistency between the raters.

Table 3. *Inter-rater reliability for overall scores*

Reliability Statistics	
Cronbach's Alpha	N of Raters
.941	2

Another inter-rater reliability analysis was run for vocabulary subsection scoring of the raters. The Cronbach's alpha was found to be *.98*, which again shows very high consistency between the raters.

Table 4. *Inter-rater reliability for vocabulary scores*

Reliability Statistics	
Cronbach's Alpha	N of Raters
.980	2

Lexical density and diversity calculations were done on a website and all the results were compiled on an Excel sheet for SPSS analysis. Lexical Density scores show percentage of content word use in the essay analyzed; lexical diversity scores, calculated according to Vocd-D formula, show the diversity in the essay analyzed and higher score shows higher diversity.

4. RESULTS AND DISCUSSION

In this chapter, the results of the data analysis procedure will be reported and discussed in response to the research questions respectively.

For the first research question, vocabulary size, lexical density and the lexical diversity in the essays will be calculated and only descriptive statistics will be provided.

For the second research question, the difference between 1st and 4th year ELT student essays in terms of vocabulary size, lexical density, and lexical diversity will be compared and the results of t-test analyses showing the difference between 1st year and 4th year students' vocabulary size, lexical density, and lexical diversity will be provided.

For the third research question, a series of correlation analyses will be given demonstrating the relationships between 1st and 4th year students' essay scores and vocabulary size, lexical density, and lexical diversity.

For the fourth and final research question, multiple regression analyses will be given demonstrating the effect size of these variables on students' overall essay scores and vocabulary subsection scores.

4.1. Vocabulary Size, Lexical Density and Lexical Diversity of the Students' Essays

Before analyzing the dataset of 1st and 4th year students' essays, looking at the students' essay scores may be useful in interpreting the results of the research questions in detail. As it is seen in the table below, the students in this sample do not differ from one another with regard to this particular writing task. The 1st year students' mean score is, *M*= 52.703, and the 4th year students' mean score is, *M*= 53.625 (see Appendix for essay scores).

Table 5. *Descriptive Statistics of the Essay Scores*

	N	Minimum	Maximum	Mean	Std. Deviation
Writing Scores_1	165	17.0	89.5	52.703	13.4577
Writing Scores_4	144	21.5	87.0	53.625	15.1258

When the mean scores were compared it seems that there is not a difference between the 1st and the 4th year students with regard to this writing task or essay type. There may be several reasons for this. One of the reasons may be that the 1st year students wrote the essays in their writing classes as an assignment and the 4th year students wrote the essays in their translation classes. Therefore, it can be concluded that the 1st year students probably did the task more attentively while the 4th year students regarded it as a forced labor.

Another reason may lie under the curriculum of the 1st and 4th year of ELT department. In the first year the students take the basic language classes such as reading, writing, listening and grammar and the curriculum starts to deal more with methodology for the next years. They are more into the methodological courses. Thus, the 4th year students probably do not remember much about essay types, and organization.

It is also necessary to look at students' vocabulary subsection scores because the lexical features are expected to directly affect the raters' scoring in this subsection. As it is clearly seen in the table below, the students' vocabulary subsection scores are not also that different from each other, as well (see Appendix for vocabulary scores).

Table 6. *Descriptive Statistics of Vocabulary Subsection Scores*

	N	Minimum	Maximum	Mean	Std. Deviation
Vocab_1	165	2	14	7.68	2.080
Vocab_4	144	4	13	7.84	2.094

1st year students' essays consist of 20510 tokens in total. 15511 of these belong to K-1 Words list (see Appendix VI), 2113 belong to K-2 Words list (see Appendix VII), 1204 belong to AWL (see Appendix VIII), and 1682 tokens are off-list. 1st year students used only 2994 different words/types in their essays. Their percentage of using AWL and off-list words together was found to be 10.17%.

Table 7. *1st year students' essays (165 essays)*

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. token %
K-1 Words	665 (56.40)	1294 (43.22)	15511 (75.63)	75.63
K-2 Words	296 (25.11)	443 (14.80)	2113 (10.30)	85.93
AWL [570 fams] TOT 2,570	218 (18.49)	330 (11.02)	1204 (5.87)	91.80
Off-List:	??	929 (31.03)	1682 (8.20)	100.00
Total (unrounded)	1179+?	2994 (100)	20510 (100)	≈100.00

4th year students' essays include 30157 tokens in total. 24626 of these belong to K-1 Words list, 2263 belong to K-2 Words list, 1520 belong to AWL, and 1748 tokens are off-list. 4th year students used 3205 types in their essays and their percentage of AWL and off-list words together was found to be 10.84%.

Table 8. *4th year students' essays (144 essays)*

Freq. Level	Families (%)	Types (%)	Tokens (%)	Cumul. token %
K-1 Words	708 (55.23)	1425 (44.46)	24626 (81.66)	81.66
K-2 Words	332 (25.90)	500 (15.60)	2263 (7.50)	89.16
AWL [570 fams] TOT 2,570	242 (18.88)	410 (12.79)	1520 (5.04)	94.20
Off-List:	??	870 (27.15)	1748 (5.80)	100.00
Total (unrounded)	1282+?	3205 (100)	30157 (100)	≈100.00

The 4th year students used more words in their essays when compared to the 1st year students but when the proportion of beyond 2k and academic word use examined, the proportions seem similar to each other.

The table below shows how much beyond 2k vocabulary was used by 1st year students. Only one section of the data was presented (see Appendix IX for all students). As it is seen on the table, students' essays vary in terms of total word number, AWL and off-list word use. The mean score of AWL and Off-list word use is 10.17%.

Table 9. *1st year students' descriptive statistics (see Appendix X for all students)*

Student	AWL	Off-list	AWL+Off-list	Total Words
1.	1,18	0,78	1,96	256
2.	6,25	1,25	7,5	240
3.	5,88	5,51	11,39	272
4.	5,14	7,01	12,15	214
5.	2,99	2,99	5,98	302
6.	5,18	3,63	8,81	193
7.	9,44	6,01	15,45	234
8.	5,13	1,83	6,96	276
9.	5,9	5,31	11,21	341
10.	5,44	3,77	9,21	244
11.	4	3,56	7,56	232
12.	6,61	4,41	11,02	233
13.	3,48	4,42	7,9	635
14.	5,48	3,2	8,68	226
15.	4,53	1,65	6,18	250
16.	7,69	3,42	11,11	240
17.	3,24	1,8	5,04	283
18.	2,7	2,9	5,6	485
19.	4,73	6,22	10,95	406
20.	3,77	3,77	7,54	339

When 4th year students' essays are examined, it is seen that they do not differ much from 1st year students in terms of beyond 2k vocabulary use but total words in their essays. It can be concluded that although the number of the words used in the essays increases, the percentage of different word use does not change much (*M*=10,77).

Table 10. *4th year students' descriptive statistics (see Appendix XI for all students)*

Student	AWL	Off-list	AWL+Off-list	Word
1.	6,88	5,07	11,95	273
2.	7,84	5,6	13,44	278
3.	1,21	1,62	2,83	249
4.	6,8	1,46	8,26	214
5.	4,1	2,46	6,56	125
6.	3,75	6,67	10,42	244
7.	4,2	0,7	4,9	149
8.	5,41	5,41	10,82	153
9.	6,22	6,7	12,92	210
10.	4,88	6,1	10,98	167
11.	7,87	1,57	9,44	255
12.	5,93	3,7	9,63	136
13.	4,86	3,47	8,33	149
14.	6,94	1,63	8,57	252
15.	2,81	2,11	4,92	286
16.	10,38	3,46	13,84	268
17.	1	3,01	4,01	408
18.	8,29	3,41	11,7	416
19.	7,66	5,86	13,52	222
20.	6,51	3,07	9,58	269

When the lexical density of the data sets calculated, it was found that they more or less had the same density percentage. Lexical density for the entire data set of 1st year students' essays is 50.8% and for the 4th year data set, it is 50.76%. The distribution of content words in the 4th year data set as follows: nouns 25.56%, adjectives 6.76%, verbs 13.14%, and adverbs 5.29%. This distribution is again nearly the same in the 1st year data set: nouns 25.38%, adjectives 6.55%, verbs 13.2%, and adverbs 5.67% (see Appendix XII for all students' density percentages). The density scores being over 40% confirms what Engberg (1995) stated about the written and spoken discourse lexical density.

In lexical diversity calculations Vocd-D formula was used and the results for the corpora are in the table below. Diversity score is not calculated in the essays consisting of less than 100 words (see Appendix XIII for all students' diversity scores).

Table 11. *Lexical diversity of the data set*

	N	Minimum	Maximum	Mean
Vocd-D	309	0,000	143,492	83.378
Valid N	309			

4.2. The Comparison of Vocabulary Size, Lexical Density and Lexical Diversity between 1st and 4th Year Data Sets

In order to find out whether 1st and 4th year students' essays differ from each other in terms of vocabulary size, lexical density, and lexical diversity a series of independent samples t-tests were utilized and mean differences of the students were analyzed.

As it is seen in the tables below, the mean difference between the vocabulary sizes of the essays is quite low (*M= .5982)* and to compare the mean scores of 1st year essays (M=10.1787) and 4th year essays (M=10.7769) an independent sample t-test was conducted. The yielded results showed that there was not a significant difference between the essays in terms of vocabulary use beyond 2k lists, which is defined as vocabulary size in this study; t(307)=2.180, *p*=.108. It means that in our sample, 1st and 4th year students' essays can be considered the same in this regard.

Table 12. *Descriptive statistics of vocabulary size*

		N	Mean	Std. Deviation	Std. Error Mean
AWL_Off	1stYear	165	10.178	3.432	.267
	4thYear	144	10.776	3.040	.253

Table 13. *Independent samples t-test of vocabulary size*

AWL_Off	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
Equal variances assumed	2.180	.141	-1.611	307	.108	-.598
Equal variances not assumed			-1.624	306.929	.105	-.598

Another t-test was utilized in order to find the difference between the lexical density mean scores of 1st and 4th year essays. Below are the tables showing the results of the t-test. The mean difference between groups was found to be .45896 and according to the t-test results, this difference was not statistically significant; $t(307)=.461$, $p=.869$. Therefore, it can be said that in terms of lexical density, two data sets are close to each other. The percentage's being over 40% confirms that the lexical density of written discourse is above 40% (Engberg, 1995). Whether the LD has a relation with writing quality is within the scope of the next research question but it is worth mentioning here what Laufer and Nation (1995) states about the LD. The fewer grammatical words in the text does not always mean a richer or denser text, it may reflect the cohesion and the word order of the text.

Table 14. *Descriptive statistics of LD*

		N	Mean	Std. Deviation	Std. Error Mean
LD	1stYear	165	51.057	4.180	.325
	4thYear	144	50.981	3.838	.319

Table 15. *Independent samples t-test of LD*

LD	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
Equal variances assumed	.461	.497	.165	307	.869	.458
Equal variances not assumed			.166	306.194	.868	.456

The last t-test was run in order to see whether the 1st and 4th year data sets differ in terms of lexical diversity, calculated by Vocd-D formula. The mean score of 1st year data was found to be M=79.182, and the mean score of 4th year data was found to be M=88.187. The mean difference between two groups (*M*=-9.005) was found to be statistically significant according to the t-test results ($t(307)=1.929$, $p< .01$). It means that 4th year students' essays are more diverse than the 1st year students' essays, and this can be generalized to the population. Having the higher diversity mean scores, 4th year students used more diverse words in their essays when compared to the 1st year students. It is normally considered that if lexical diversity is high in the learner's output, it indicates much higher proficiency than lower lexical diversity (Malvern et al., 2004).

Table 16. *Descriptive statistics of lexical diversity (Vocd-D scores)*

		N	Mean	Std. Deviation	Std. Error Mean
LD	1stYear	165	79.182	21.410	1.666
	4thYear	144	88.187	22.642	1.886

Table 17. *Independent samples t-test of lexical diversity (Vocd-D scores)*

LD	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference
Equal variances assumed	.1.929	.166	-3.590	307	.000	-9.005
Equal variances not assumed			-3.577	296.059	.000	-9.005

4.3. The Correlational Relationship between Vocabulary Size, Lexical Density, and Lexical Diversity with Essay Scores

In order to answer the third research question of the study two sets of correlation analyses were run. In the first analysis, the relationship of vocabulary size, lexical density, and lexical diversity scores of 1st year students with their argumentative essay scores was found.

Table 18. *Correlation Analysis of 1st year essays*

		Writing Scores_1	LD_1	AWL_Off_1	Vocd-D_1
Writing Scores_1	Pearson Correlation	1	-.057	-.033	.260**
	Sig. (2-tailed)		.465	.670	.001
	N	165	165	165	165
LD_1	Pearson Correlation	-.057	1	.402**	.149
	Sig. (2-tailed)	.465		.000	.056
	N	165	165	165	165
AWL_Off_1	Pearson Correlation	-.033	.402**	1	.120
	Sig. (2-tailed)	.670	.000		.124
	N	165	165	165	165
Vocd-D_1	Pearson Correlation	.260**	.149	.120	1
	Sig. (2-tailed)	.001	.056	.124	
	N	165	165	165	165

**. Correlation is significant at the 0.01 level (2-tailed).

It is seen that there is not a significant correlation between lexical density and essay scores (r= -.057, p > .01), there is not a significant correlation between vocabulary size and essay scores (r= -.033, p > .01), but there is a significant weak positive correlation between lexical diversity and argumentative essay scores of 1st year students (r= .260, p < .01). It means that students with high lexical diversity also get higher scores in their argumentative essays. The effect size of lexical diversity on essay scores seems low and it will be analyzed with multiple regression analysis for the fourth research question.

It is not within the scope of our research questions but it is worth mentioning that based on the results of the study, lexical density and vocabulary size of the students are moderately correlated (r=.402, p < .01). That is, students using more words from AWL and beyond 2k are also found to have more lexical density in their essays.

Table 19. *Correlation Analysis of 4th year students*

		Writing Scores_4	LD_4	AWL_Off_4	Vocd-D_4
Writing Scores_4	Pearson Correlation	1	-.017	.069	.033
	Sig. (2-tailed)		.840	.411	.695
	N	144	144	144	144
LD_4	Pearson Correlation	-.017	1	.309**	.343**
	Sig. (2-tailed)	.840		.000	.000
	N	144	144	144	144
AWL_Off_4	Pearson Correlation	.069	.309**	1	.240**
	Sig. (2-tailed)	.411	.000		.004
	N	144	144	144	144
Vocd-D_4	Pearson Correlation	.033	.343**	240**	1
	Sig. (2-tailed)	.695	.000	.004	
	N	144	144	144	144

**. Correlation is significant at the 0.01 level (2-tailed).

When the same correlation analysis was done for the 4th year students' essay scores and lexical features, unlike 1st year students', it was seen that there was not a significant correlation between lexical density and essay scores (r= -.017, p > .01), there was not a significant correlation between vocabulary size and essay scores (r= .069, p > .01), and there was not a significant correlation between lexical diversity and essay scores

(r= .033, p > .01). This means that the abovementioned lexical features do not explain the essay scores of 4th year students in a significant way. Lexical diversity was found to be not correlating with writing scores of the students for the 4th year students.

For the 4th year data, it was found that lexical density had a significant weak positive correlation with vocabulary size and lexical diversity respectively (r= -.309, p < .01), (r= .343, p < .01). It was also found that there was a significant weak positive correlation between vocabulary size and lexical diversity (r= .240, p < .01).

4.4. The Effect of Lexical Density, Vocabulary Size, and Lexical Diversity on the Essay Scores Vocabulary Scores

The correlation analyses showed that only lexical diversity had a significant effect on students' essay scores for the 1st year students. In order to see to what extent lexical diversity explains the essay scores of the students a multiple regression analysis was run for 1st year students and 4th year students separately. The results of the regression analysis for the 1st year students' essays indicated that 7.8% of the variance (essay scores) was explained by the independent variables (lexical density, vocabulary size and lexical diversity) (R^2= .078, F(3,161)=4.550, p < .01). However, while lexical diversity significantly predicted the 7.5% of the essay scores alone (β = .075, p < .01), other variables (lexical density and vocabulary size) only explained .3% of the essay scores non-significantly.

Table 20. *Multiple Regression Analysis for the 1st Year Data (Essay Scores)*

Model	R	R Square	Adjusted R Square	Change Statistics				
				R Square Change	F Change	df1	df2	Sig. F Change
1	.057[a]	.003	-.003	.003	.537	1	163	.465
2	.058[b]	.003	-.009	.000	.021	1	162	.886
3	.280[c]	.078	.061	.075	13.055	1	161	.000

a. Predictors: (Constant), Lexical Density_1

b. Predictors: (Constant), Lexical Density_1, Vocabulary Size_1

c. Predictors: (Constant), Lexical Density_1, Vocabulary Size_1, Lexical Diversity_1

The results of the regression analysis for the 4th year students showed that none of the predictors explained the 4th year students' essay scores significantly ($F(3,140)=.343$, $p > .01$). The table below states that it is not necessary to interpret the results of the multiple regression analysis for the 4th year students because there is no significant value found in the analysis.

Table 21. *ANOVA for the 4th Year Data Regression Analysis (Essay Scores)*

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	9.415	1	9.415	.041	.840[b]
	Residual	32707.335	142	230.333		
	Total	32716.750	143			
2	Regression	209.347	2	104.673	.454	.636[c]
	Residual	32507.403	141	230.549		
	Total	32716.750	143			
3	Regression	238.583	3	79.528	.343	.794[d]
	Residual	32478.167	140	231.987		
	Total	32716.750	143			

a. Dependent Variable: Writing Scores_4
b. Predictors: (Constant), Lexical Density_4
c. Predictors: (Constant), Lexical Density_4, Vocabulary Size_4
d. Predictors: (Constant), Lexical Density_4, Vocabulary Size_4, Lexical Diversity_4

So as to see whether these lexical features affect vocabulary scores of the students, two other multiple regression analyses were run for the 1st and the 4th year data successively. The results yielded from the regression analysis for the 1st year students' essays showed that 8.7% of the variance (vocabulary scores) was explained by the independent variables (lexical density, vocabulary size and lexical diversity) ($R^2= .087$, $F(3,161)= 14.676$, $p < .01$). Only lexical diversity significantly accounted for 8.3% of the vocabulary scores ($\beta = .083$, $p < .01$), other variables (lexical density and vocabulary size) only explained .4% of the vocabulary scores non-significantly.

Table 22. *Multiple Regression Analysis for the 1st Year Data (Vocabulary Scores)*

Model	R	R Square	Adjusted Square	R Change Statistics				
				R Square Change	F Change	df1	df2	Sig. F Change
1	.060[a]	.004	-.003	.004	.589	1	163	.444
2	.060[b]	.004	-.009	.000	.000	1	162	.992
3	.295[c]	.087	.070	.083	14.676	1	161	.000

a. Predictors: (Constant), Lexical Density_1
b. Predictors: (Constant), Lexical Density_1, Vocabulary Size_1
c. Predictors: (Constant), Lexical Density_1, Vocabulary Size_1, Lexical Diversity_1

On the other hand, the results of the regression analysis for the 4th year students showed that none of the predictors explained the 4th year students' vocabulary scores significantly ($F(3,140)=.436$, $p > .01$). The case was the same for the 4th year overall essay scores. The table below states that it is not necessary to interpret the results of the multiple regression analysis for the 4th year students because there is no significant value found in the analysis affecting vocabulary scores.

Table 23. *ANOVA for the 4th Year Data Regression Analysis (Vocabulary Scores)*

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.018	1	.018	.004	.950[b]
	Residual	627.309	142	4.418		
	Total	627.326	143			
2	Regression	4.873	2	2.437	.552	.577[c]
	Residual	622.453	141	4.415		
	Total	627.326	143			
3	Regression	5.808	3	1.936	.436	.728[d]
	Residual	621.518	140	4.439		
	Total	627.326	143			

a. Dependent Variable: Vocabulary Scores_4
b. Predictors: (Constant), Lexical Density_4
c. Predictors: (Constant), Lexical Density_4, Vocabulary Size_4
d. Predictors: (Constant), Lexical Density_4, Vocabulary Size_4, Lexical Diversity_4

The results showed that lexical features in the essays of the students mainly show similarities but only lexical diversity in the essays of the fourth year students is significantly higher than the essays of the first year students. That is, their essays are more diverse than the first year students' essays, which can be interpreted as they are able to use more diverse words in their essays.

5. CONCLUSION

This study aimed to find out the relationship between vocabulary size, lexical diversity, lexical density and EFL writing scores. In this chapter, the conclusions that are drawn from the results of the analyses will be presented. The research questions will be addressed one by one with the summary of the results.

5.1. Vocabulary Size, Lexical Density and Lexical Diversity of the Students' Essays

The first research question was posed to obtain the descriptive statistics of the data set. As also stated in the results chapter, first the essay scores and vocabulary scores of the students were given to give the overall picture of the essays. The students' essay scores and vocabulary scores seem very close, which shows that the groups do not differ in terms of these scores (essay scores mean M_1= 52.703 and M_4= 53.625; vocabulary scores mean M_1= 7.68 and M_4= 7.84). The possible reasons of this situation were discussed in the results section (see 4.1.1).

The results of the vocabulary size analysis are different from Laufer and Nation (1995). They compared the vocabulary size of different proficiency level students and found that the groups differ from each other significantly. The results of Signes and Arroitia (2015) also looked at the different proficiency level students vocabulary size and their results also showed that students differ in terms of vocabulary use beyond 2k. In our study, the 1st and the 4th year students' essays seem very close to each other with this regard. The 1st year data consist of 10.17%, the 4th year data consist of 10.77% AWL and off-list word use. It can intuitively be concluded that 4-year education in the department did not contribute much to the students in terms of productive vocabulary use in this particular writing task.

The lexical density of the data sets was also found to be similar. Lexical density for the entire data set of 1st year students' essays is 50.8% and for the 4th year data set, it is 50.76%. The lexical density of written discourse is expected to be over 40% (Engberg, 1995). Therefore, this result supported the claim that lexical density of written discourse is expected to be over 40%.

The lexical diversity of the data sets range from 0 to 143.492 (M= 83.378). The minimum score is 0, because vocd-D does not give a diversity score for the texts less than 100 words.

5.2. The Comparison of Vocabulary Size, Lexical Density and Lexical Diversity between 1st and 4th Year Data Sets

The second research question sought the answer whether the differences in lexical features of 1st and 4th year data sets are significant or not. According to the t-tests results, the two data sets are not significantly different in terms of vocabulary size (t(307)=2.180, p=.108) and lexical density (t(307)=.461, p=.869). However, the lexical diversity scores seem to differ significantly t(307)=1.929, p= .000).

What can be concluded from this result is that if lexical diversity is high in the learner's output, it may be an indicator of higher proficiency than lower lexical diversity (Malvern et al., 2014). Thus, in our data sets the 4th year students can be regarded as more proficient by looking at this particular result (M_1= 79.182, M_4= 88.187). Nonetheless, in order to come up with a general conclusion, the effect of lexical diversity on writing scores should be examined.

There are also contradicting results with regard to lexical diversity of different proficiency level students' essays. In Gonzalez (2013) the lexical diversity of high proficient students' essays was higher than lower proficient students. However in Wang (2014), the diversity scores of the different proficiency level students did not differ significantly. This study supports the findings of Gonzalez (2013) in this regard.

5.3. The Correlational Relationship between Vocabulary Size, Lexical Density, and Lexical Diversity with Essay Scores

The third research question was asked to see the correlational relationship between vocabulary size, lexical density, and lexical diversity with essay scores.

According to the results of 1st year data, the lexical diversity significantly correlate with essay scores (r= .260, p < .01). This result for the 1st year students

confirms the studies that have found the similar relationship between lexical diversity and writing scores of the students (Douglas, 2010; Mellor, 2010; Gonzalez, 2013). However, Lavallee and McDonough (2015) and Wang (2014) could not find a significant relationship between the lexical diversity and the writing scores of the students. It is clear that there is a controversy on lexical diversity and writing quality and much research is needed to come up with more rigid conclusion on the issue. The conflicting results may be because of different lexical diversity measure formula use, the sample texts, the participants, the essay type etc.

The correlation analysis for the 4th year data did not result in the same manner. That is, there is not a significant correlation between lexical density and essay scores (r= -.017, p > .01), there is not a significant correlation between vocabulary size and essay scores (r= .069, p > .01), and there is not a significant correlation between lexical diversity and essay scores (r= .033, p > .01). What can be concluded from this result is that these lexical features are not related with the essay scores of the 4th year students. This discrepancy between two groups of the study also shows that lexical diversity measure is still problematic because writing is a multi-dimensional process, a part of which is vocabulary. And as it is seen from the results of the study, the vocabulary itself is not enough to explain the scores the students get from their essays. Other aspects of writing quality seem to be more effective than the lexical features.

5.4. The Effect of Lexical Density, Vocabulary Size, and Lexical Diversity on the Essay Scores Vocabulary Scores

To answer the last research question multiple regression analyses were utilized for the 1st and the 4th year data separately. The results showed that, for the 1st year data, the lexical features account for 7.8% of the essay scores significantly (R^2= .078, F(3,161)=4.550, p < .01). It shows that the essay scores do not mostly depend on the lexical features of the essays, there are other factors affecting the rating of the essays.

According to the results of the regression analysis for the 4th year students, none of the predictors explained the 4th year students' essay scores significantly

($F(3,140)=.343$, $p > .01$). That is the essay scores of the 4th year students do not depend on the lexical features.

Among these lexical features only lexical diversity was found to have a significant impact on the essay scores as in the studies of Mellor (2010), Douglas (2010) and Gonzalez (2013). There are also contradicting results, claiming that the lexical diversity does not have an impact on writing scores such as Wang (2014) and Lavallee & McDonough (2015).

Vocabulary size was also found to have an impact on writing scores in studies of Lemmouh (2008), Gonzalez (2013) and Karakoç (2016) but in our data set vocabulary set can only explain a very small percentage of essay scores.

To what extent these lexical features explain the vocabulary scores of the students was the second part of the last research question. The results for the 1st year students' essays showed that 8.7% of the variance (vocabulary scores) can be explained by the independent variables (lexical density, vocabulary size and lexical diversity) ($R^2= .087$, $F(3,161)= 14.676$, $p < .01$). Only lexical diversity significantly accounted for 8.3% of the vocabulary scores ($\beta = .083$, $p < .01$), other variables (lexical density and vocabulary size) only explained .4% of the vocabulary scores non-significantly.

The results for the 4th year students' essays did not yield a significant result as in the case of their essay scores ($F(3,140)=.436$, $p > .01$). Only very small percentage of vocabulary scores can be explained by the lexical features in this data set.

The results of the study showed that Turkish ELT students have difficulty in writing, because the average of both first and fourth year students' essay scores are considerably low. Moreover, both fourth and first year students' essays consisted of 200 words on average. It may show that the students do not want to write or cannot write as much as the expected essay length. The vocabulary size and lexical diversity scores also show that the students have difficulty in expressing themselves in writing because they almost did not use words out of 1k and 2k word lists. This study showed once more that our students do really have problems or drawbacks in writing, as a productive skill. In order to be able to say that essay genre did have an effect on the results, the same research design can be duplicated with other essay genres. However, it is clear that in this sample, the students had difficulty in writing an argumentative essay on the given topic

and this may give some insights to the writing lecturers. For example, more essays should be assigned to the students with at least 350 words limit to make the students write more.

It was also seen that the students did not really know the basic outline of the essay genre. They lost points from the content section of the rubric and they also lost points from the discourse markers section because they could not use the relevant discourse markers suitable to essay genre. This shows that writing courses should be more challenging and essay writing should be implemented more into the four-year curriculum via assignments or examinations. Two compulsory writing courses in the first year seem to be not efficient and adequate. Writing lecturers should raise the students' awareness on the importance of expressing the ideas via writing.

Another finding of the study was that despite being low, lexical diversity had an effect on essay scores. A vocabulary course, in which the awareness of the students on vocabulary profiles, vocabulary lists, frequency of the words are raised, can be implemented to the curriculum. They are not only students but also future teachers, so they should also be aware of the importance of lexical diversity and productive vocabulary knowledge in productive aspects of language learning and teaching.

The vocabulary level of the students should be enhanced via vocabulary activities implemented in writing courses such as affixation exercises, vocabulary level tests throughout the semester, paragraph completion exercises etc. That is, vocabulary should be considered as a separate skill like grammar, writing or speaking. Rich vocabulary use should be encouraged and students using rich vocabulary should be rewarded to raise the awareness.

For further studies it can be suggested that the receptive vocabulary aspect can also be included in research design to support the productive vocabulary measures of the written texts of the students. The developmental process of the students can also be studied by making them write three or four essays in an academic year. It can also be useful in interpreting their actual productive vocabulary use.

The exam papers of the students can be used as data sets to eliminate the question whether they have written attentively and whether the essays reflect their real production.

To conclude, the conflicting results of the lexical features of the students' essays and whether they have an impact on writing scores and vocabulary scores seem to be the

case in our study as well. What is clear in our study is that the raters do not consider the lexical features much while rating the students' essays because these lexical features did not seem to have a direct impact neither on essay scores nor on vocabulary scores.

REFERENCES

Aitkuzhinova-Arslan, A., Gün, S., & Üstünel, E. (2016). Teaching vocabulary to Turkish young learners in semantically related and semantically unrelated sets by using digital storytelling. *Journal of Language and Linguistic Studies*, *12*(1), 42-54.

Aliakbari, M., & Boghayeri, M. (2014). A needs analysis approach to ESP design in Iranian context. *Procedia-Social and Behavioral Sciences*, *98*, 175-181.

Avent, J.R, & Austermann, S (2003). Reciprocal scaffolding: a context for communication treatment in aphasia. *Aphasiology, 17*, 397-404

Basoz, T., & Cubukcu, F. (2014). The effectiveness of computer assisted instruction on vocabulary achievement. *Mevlana International Journal of Education*, *4*(1), 44-52.

Beers, S. F., & Nagy, W. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre?. Reading and Writing, 22(2), 185–200.

Begriche, F. (2014). The role of teaching vocabulary to enhance foreign language learners' writing skill (Master's Thesis). Retrieved from http://dspace.univ-biskra.dz:8080/jspui/handle/123456789/4741

Biber, D. 2006. University language: A corpus-based study of spoken and written registers. Amsterdam: John Benjamins.

Bogaards, P. (2000) Testing L2 Vocabulary Knowledge at a High Level: the case of the Euralex French Tests. Applied Linguistics, 21 (4), 490-516.

Carrell, P.L., & Monroe, L.B. (1993). Learning styles and composition, *The modern Language Journal*, 77, 148-162.

Choi, J. (2012). Self-Access English Learning Needs: Student and Teacher Perspectives. Listening, 42(49), 15.

Coxhead, A. (2000). A New Academic Word List. TESOL QUARTERLY, 34 (2), 213- 238

Cronbach, L. J. (1942). An analysis of techniques for diagnostic vocabulary testing. *The journal of educational research*, *36*(3), 206-217.

Crossley, S. A., & McNamara, D. S. (2010). Cohesion coherence, and expert evaluations of writing proficiency. In R. Catrambone, & S. Ohlsson (Eds.), Proceedings of the 32nd annual conference of the cognitive science society (pp. 984–989). Austin, TX: Cognitive Science Society.

Çelik, S., & Toptaş, V. (2010). Vocabulary learning strategy use of Turkish EFL learners., Procedia Social and Behavioral Sciences 3, 62-71, http://www.sciencedirect.com/science/article/pii/S187704281001387X

Dang, T. N. Y., & Webb, S. (2013). The lexical profile of academic spoken English.

English for Specific Purposes. http://dx.doi.org/10.1016/j.esp.2013.08.001

deBoer, F. (2014). Evaluating the comparability of two measures of lexical diversity. *System*, *47*, 139-145.

Douglas, S. R. (2010). *Non-native English speaking students at university: Lexical richness and academic success* (Doctoral dissertation, University of Calgary).

Ellis, R. (1994) The Study of Second Language Acquisition. Oxford: Oxford University Press.

Ellis, R. & He, X. (1999). The role of modified input and output in the incidental acquisition of word meanings. Studies in Second Language Acquisition, 21, 285- 301.

Engber, C.A. (1995) The relationship of lexical proficiency to the quality of ESL compositions. Journal of Second Language Writing 4, 2: 139-155.

Fhonna, R. (2014). The correlation between mastering vocabulary and speaking ability (case study at SMA 10 Fajar Harapan Banda Aceh). Visipena, 5(1), 90-99.

Folse, K. (2008). Myth 1: Teaching vocabulary is not the writing teacher's job. In J. Reid (Ed.) Writing myths: Applying second language research to classroom teaching (pp.1-17). Ann Arbor: University of Michigan Press.

Gass, S. M. & Selinker, L. (2008). *Second language acquisition: An introductory course*.

González, M. C. (2013). *The intricate relationship between measures of vocabulary size and lexical diversity as evidenced in non-native and native speaker academic compositions* (Doctoral dissertation, University of Central Florida Orlando, Florida).

Gregori-Signes, C., & Clavel-Arroitia, B. (2015). Analysing Lexical Density and Lexical Diversity in University Students□ Written Discourse. *Procedia-Social and Behavioral Sciences*, *198*, 546-556.

Grela, Bernard G. (2002). Lexical verb diversity in children with Down syndrome.

Clinical Linguistics & Phonetics, 14, 251-263

Hamouda, A. (2013). An investigation of listening comprehension problems encountered by Saudi students in the EL listening classroom. International Journal of Academic Research in Progressive Education and Development, 2(2), 113-155.

Henriksen, B. (1999) Three dimensions of vocabulary development. Studies in Second Language Acquisition 21: 303-317

Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. Reading in a Foreign Language, 13(1), 403–430.

Hunt, A., & Beglar, D. (2005). A framework for developing EFL reading vocabulary. *Reading in a Foreign language*, *17*(1), 23.

Hulstijn, J. and B. Laufer. 2001. Some empirical evidence for the Involvement Load Hypothesis in vocabulary acquisition. Language Learning 51: 539-558

Jackson, M. E., (2004). Will electronic journals eliminate the need for ILL?, Interlending & Document Supply, 32(3), 192-193.

Jarvis, S. (2002). Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing 19*, 1: 57–84.

Johansson, V. (2008). Lexical Diversity and Lexical Density in Speech and Writing: A Developmental Perspective. Working Papers, Lund University, Dept. of Linguistics and Phonetics. 53. pp. 61-79.

Karakoç, D. (2016). *The Impact Of Vocabulary Knowledge On Reading, Writing And Proficiency Scores Of B2.2 Level Turkish Students: A study with Anadolu University English Prep-School Students*, (Unpublished MA Thesis), Graduate School of Educational Sciences, Anadolu University, Eskişehir.

Khotimah, S. (2014). The use of problem based learning to improve students' speaking ability. ELT Forum: Journal of English Language Teaching, 3(1), 50-56.

Koizumi, R. (2012). Relationships between text length and lexical diversity measures: can we use short texts of less than 100 tokens. *Vocabulary Learning and Instruction*, *1*(1), 60-69.

Kök, İ. & Canbay, O. (2011). An experimental study on the vocabulary level and vocabulary consolidation strategies, *Procedia Social and Behavioral Sciences* 15, p. 891-894.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Addition evidence for theinput hypothesis. Modern Language Journals, Vol. 73, 440-464.

Laufer , B. (1997) The Lexical Plight in Second Language Reading in J. Coady and T. Huckin (Eds) Second Language Vocabulary Acquisition and Pedagogy Cambridge CUP, 140-55

Laufer, B. (1998). The development of passive and active vocabulary in a second language: same or different? Applied Linguistics 12: 255-271.

Laufer, B. (2005). Lexical frequency profiles: From Monte Carlo to the real world: A response to Meara (2005). *Applied Linguistics*, *26*(4), 582–588.

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

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied linguistics*, *16*(3), 307–322.

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language testing*, *16*(1), 33–51.

Lavallée, M., & McDonough, K. (2015). Comparing the Lexical Features of EAP Students' Essays by Prompt and Rating. *TESL Canada Journal*, *32*(2), 30-44.

Lee, S. H., & Muncie, J. (2006). From receptive to productive: Improving ESL learners' use of vocabulary in a postreading composition task. *TESOL Quarterly*, *40*(2), 295–320.

Lemmouh, Z. (2008). The relationship between grades and the lexical richness of student essays. *Nordic Journal of English Studies*, *7*(3), 163-180.

Lenko-Szymanska, A. (2002). How to trace the growth in learners‟ active vocabulary: A Corpus-based study, in B. Ketteman and G. Marko (eds.) Teaching and Learning by Doing Corpus Analysis. Amsterdam: Rodopi. pp. 217-230.

Lewis, M. (1993). *The lexical approach* (Vol. 1, p. 993). Hove: Language Teaching Publications.

Ling, G. U. I. (2015). Predictability of vocabulary size on learners' EFL proficiency: Taking VST, CET4 and CET6 as instruments. Studies in Literature and Language, 10(3), 18-23. http://dx.doi.org/10.3968/6679

Malvern, D.D. and Richards, B.J. 1997: A new measure of lexical diversity. In Ryan, A. and Wray, A., editors, *Evolving models of language*. Clevedon: Multilingual Matters, 58–71.

Malvern, D., Richards, B. J., Chipere, N., & Durán, P. (2004). Lexical diversity and language development: Quantification and assessment: New York: Palgrave Macmillan.

Malvern, D., & Richards, B. (2012). Measures of lexical richness. *The Encyclopedia of Applied Linguistics*.

Manchón, R. M. (2011). Learning-to-write and writing-to-learn in an additional language. Amsterdam: John Benjamins.

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, *29*, 3-15.

McCarthy, P. M. (2005). An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity (MTLD). *Dissertation Abstracts International*, *66*, 12.

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior research methods*, *42*(2), 381-392.

Meara, P. (2002) The rediscovery of vocabulary. Second Language Research 18, 4: 393-407. [1.4

Meara, P. (2005). Lexical frequency profiles: A Monte Carlo analysis. *Applied Linguistics 26*(1), 32-47.

Meara, P. & Fitzpatrick, T. (2000). Lex30: an improved method of assessing productive vocabulary in an L2. System 28: 19–30.

Meara, P., & Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect, 16*, 5-19.

Mellor, A. (2011). Essay length, lexical diversity and automatic essay scoring. *Memoirs of the Osaka Institute of Technology*, *55*(2), 1-14.

Milton, J. (2007) `Lexical profiles, learning styles and the construct validity of lexical size tests', in Daller, H., Milton, J., and Treffers-Daller J. (eds.) Modelling and assessing vocabulary knowledge.Cambridge:CambridgeUniversity Press,pp.47-58.

Min, H. T. (2016). Effect of teacher modeling and feedback on EFL students' peer review skills in peer review training. *Journal of Second Language Writing*, *31*, 43-57.

Mobarg, M. (1997) Acquiring, teaching and testing vocabulary. International Journal of Applied Linguistics 7, 2: 201-222. [1.3

Mokhtar, A. A. (2010). Achieving Native-like English Lexical Knowledge: The Non- native Story. Journal of Language Teaching and Research, Vol. 1, No. 4, pp. 343-352

Nation, P. (1990). Teaching and Learning Vocabulary: Boston: Heinle & Heinle

Nation, P. (1995). The Word on Words: An Interview with Paul Nation. Interviewed by N. Schmitt. The Language Teacher 19 (4), 5-7

Nation, I. S. P. (2000) Learning Vocabulary in Another Language. Cambridge: Cambridge University Press.

Nation, P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes*, *63*(1), 59–82.

Nation, I. S. P. (2007). The four strands. Language Learning and Teaching, 1(1), 1-12.

Nation, I.S.P. & Webb, S. (2011). Researching and analyzing vocabulary. Boston: Heinle-Cengage.

Paribakht, T. S., & Wesche, M. B. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada journal*, *11*(1), 09-29.

Pérez Manzanilla, I. S., & Díaz Cabrera, K. M. (2014). Factors that may have an impact on advanced EFL students' speaking ability (Master's Thesis). Retrieved from http://cdigital.uv.mx/handle/123456789/35250

Putra, A. R. (2014). Using picture series to improve the writing skill on recount of 8th graders SMP Muhammadiyah 3 Depok in the 2013–2014 Academic Year (Doctoral dissertation). Retrieved from http://eprints.uny.ac.id/id/eprint/18475

Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. Language Testing, 10, 355-371.

Read, J. (2004). Research in Teaching Vocabulary. Annual Review of Applied Linguistics. pp. 146-161

Richards, J. C. (1976). The role of vocabulary teaching. TESOL Quarterly, 10, 77-89.

Roche, T., & Harrington, M. (2013). Recognition vocabulary knowledge as a predictor of academic performance in an English as a foreign language setting. *Language Testing in Asia*, *3*(1), 12.

Rudy, M. (2013). EFL Writing strategies of the second year students of SMPIT Daarul 'Ilmi Kemiling Bandar Lampung. Paper presented at International Conference on Education and Language (ICEL), Bandar Lampung University.

Schmitt, N. (1997). „Vocabulary learning strategies." In N. Schmitt and M. McCarthy, (Eds.), Vocabulary: Description, Acquisition and Pedagogy 199-227. Cambridge University Press, Cambridge.

Schmitt, N. (2000). Vocabulary in Language Teaching. Cambridge: CUP

Schmitt, N. (2010). Researching Vocabulary: A vocabulary research manual. New York: Palgrave MacMillan.

Schmitt, N., Ching, Ng, J. W & Garras, J. (2010). The Word Association Format: Validation Evidence. Language Testing, online. 1-22.

Silverman, R. D., Proctor, C. P., Harring, J. R., Hartranft, A. M., Doyle, B., & Zelinke, S. B. (2015). Language skills and reading comprehension in English monolingual and Spanish–English bilingual children in grades 2–5. Reading and Writing, 28(9), 1381-1405. http://dx.doi.org/10.1007/s11145-015-9575

Skehan, P. (2009). Lexical performance by native and non-native speakers on language-learning tasks. In *Vocabulary Studies in First and Second Language Acquisition* (pp. 107-124). Palgrave Macmillan UK.

Solak, E., & Altay, F. (2014). Prospective EFL teachers' perceptions of listening comprehension problems in Turkey. Journal of International Social Research, 7(30).

Staples, S., & Reppen, R. (2016). Understanding first-year L2 writing: A lexico-grammatical analysis across L1s, genres, and language ratings. *Journal of Second Language Writing,* 32, p.17-35.

Tahir, S. Z. (2015). Improving students' speaking skill through voice chat at University of Iqra Buru. Journal of Modern Education Review, 5(3), 296-306. http://dx.doi.org/10.15341/jmer(2155-7993)/03.05.2015/009

Topkaraoğlu, M., & Dilman, H. (2013). Effects of Studying Vocabulary Enhancement Activities on Students' general Language Proficiency Levels. *Batı Anadolu Eğitim Bilimleri Dergisi*, *2013*(8).

Unaldi, I. (2011). A comparative investigation of lexical networks of Turkish learners of English as a foreign language: A corpus based study. *Unpublished doctoral dissertation, Cukurova University, Adana, Turkey*.

Utku, R. (2014). *Meta-analysis of the Lexical Frequency Profile* (Unpublished MA Thesis), The University of Queensland School of Languages and Comparative Cultural Studies Brisbane, Australia

Vandergrift, L., & Baker, S. (2015). Learner variables in second language listening comprehension: An exploratory path analysis. Language Learning, 65(2), 390- 416. http://dx.doi.org/10.1111/lang.12105

Wang, S. (2015). An empirical study on the role of vocabulary knowledge in EFL listening comprehension. Theory and Practice in Language Studies, 5(5), 989- 995. http://dx.doi.org/10.17507/tpls.0505.14

Wang, X. (2014). The relationship between lexical diversity and EFL writing proficiency. *University of Sydney Papers in TESOL*, *9*.

Waring, R. (2002). Basic principles and practice in vocabulary instruction. The Language Teacher. Retrieved in June, 2016, from http://jalt-publications.org/old_tlt/articles/2002/07/waring

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. Studies in Second Language Acquisition, 27: 33-52.

Webb, S. (2008). Receptive and productive vocabulary size. Studies in Second Language

Acquisition. Vol. 30. pp. 79-95

Webb, S., & Nation, P. (2008). Evaluating the vocabulary load of written text. *TESOLANZ Journal*, *16*, 1–10.

Wesche, M. & Paribakht, T.S. (1996). Assessing second language vocabulary knowledge: Depth versus breadth. Canadian Modern Language Review, 53: 13- 40.

Xue, G., & Nation, I. S. P. (1984). A university word list. Language Learning and Communication, Vol. 3, p. 215–299.

Yang, Y. I. (2015). An investigation of Chinese junior high school teachers' and students' attitudes towards EFL writing. International Journal of Research Studies in Education, 5(2). http://dx.doi.org/10.5861/ijrse.2015.1209

Yüksel, İ (2012) *Cross-sectional evaluation of general and academic lexical competence and performance.* (Doctoral Dissertation). Anadolu University, Eskişehir, Turkey

Zareva, A. (2005). Models of L2 learners‟ vocabulary knowledge assessment. System, 33 (4), 547–562.

Zareva, A., Schwanenflugel, P., Nikolova, Y. (2005). Relationship between lexical competence and language proficiency: Variable sensitivity. Studies in Second Language Acquisition, 27 (4), 567-595.

Zhang, J., McBride-Chang, C., Wong, A. M. Y., Tardif, T., Shu, H., & Zhang, Y. (2014). Longitudinal correlates of reading comprehension difficulties in Chinese children. Reading and Writing, 27(3), 481-501. http://dx.doi.org/10.1007/s11145-013-9453-4

APPENDICES

Appendix-I. Writing Prompt

29/02/2016

Name: Elif Nur
Surname: Doğu
Group: 4-A

4A15

Write a well-developed argumentative essay on the topic below:

Technology and Imagination
Some people say that in our modern world, dominated by science, technology and industrialization, there is no longer a place for dreaming and imagination. Discuss your opinion about this statement.

Allocated time: 60 minutes Thinking or Imagining

Are imagination and dreaming crucial parts of humanbeing? Can we keep up without thinking or trying to produce something new or something which has never been done / thought before? There are, of course, some opposite ideas about this topic. Some of you say yes, and some say no.

In this modern world where technology is frequently used, we see different invention everyday. People find something totally new or develop their ideas in a completely unusual way. We possibly cannot bump into to the new things just like that. To improve ourselves or add something useful to the life - not only our lives but also to all the humanbeings' experience - we have to contemplate on the related subject / case. But, is thinking about the related topic, a way of dreaming? Or should we try to be more creative? Only thinking will be enough to find a different solution? Will that improve us and the world? Surely, not. If we want to enrich ourselves and the society in which we live, we are supposed to make use of the former experience we had, in this way we can see the deficiency and what we need, and then, to correlate these experiences to our lives according to the current conditions of the time.

With the help of this way of thinking - we call it imagination - we can surely develop new ideas and science or technology shouldn't prevent us from doing this. Everything starts in human mind where the imagination comes from. If we cannot imagine - which will be new - we cannot transfer those images to the reality. And the societies who don't like to think, dream, imagine are always obliged to repeat themselves and fall behind.

Appendix-II. ESL Composition Profile

ESL COMPOSITION PROFILE		
	RANGE	**CRITERIA**
Content	**25-22**	**EXCELLENT TO VERY GOOD:** knowledgeable—substantive—thorough development of thesis/genre—relevant to assigned topic
	21-17	**GOOD TO AVERAGE:** some knowledge of subject—adequate range—limited development of thesis/genre—lacks detail
	16-11	**FAIR TO POOR:** limited knowledge of subject—little substance—inadequate development of thesis/genre
	10-5	**VERY POOR:** does not show knowledge of subject—non-substantive—not pertinent—OR not enough to evaluate
Organization	**15-13**	**EXCELLENT TO VERY GOOD**: organization clearly stated and supported—well organized and very through development of introduction, body and conclusion, well-organized and very through development of supporting details
	12-9	**GOOD TO AVERAGE:** somewhat choppy—main ideas stand out, but organization unclear—limited development of introduction, body and conclusion—and/or limited development of supporting details
	8-5	**FAIR TO POOR:** ideas confused or disconnected—lacks logical sequencing and development of introduction, body and conclusion, and/or limited development of supporting details
	4-2	**VERY POOR:** does not communicate—no organization—OR not enough to evaluate
Discourse Markers	**10-9**	**EXCELLENT TO VERY GOOD:** Appropriate use and wide range of cohesive devices (signal words, pronouns, key words, demonstrative adjectives)
	8-6	**GOOD TO AVERAGE:** Mostly appropriate use and range of cohesive devices (signal words, pronouns, key words, demonstrative adjectives)
	5-3	**FAIR TO POOR:** Limited use and range of cohesive devices (signal words, pronouns, key words, demonstrative adjectives)
	2-1	**VERY POOR**: little or no linkage between sentences
Vocabulary	**15-13**	**EXCELLENT TO VERY GOOD**: sophisticated range—effective word/idiom choice and usage—word form mastery—appropriate register
	12-9	**GOOD TO AVERAGE:** adequate range—occasional errors of word/idiom form, choice, usage but meaning not obscured
	8-5	**FAIR TO POOR:** limited range—frequent errors of word/idiom form, choice, usage and/or meaning confused or obscured
	4-2	**VERY POOR**: essentially translation—little knowledge of English vocabulary, idioms, word form OR not enough to evaluate
Sentence Construction	**30-26**	**EXCELLENT TO VERY GOOD:** effective use of simple, compound, and complex sentences—effective use of coordinators, subordinators, and transitions—few errors of S-V agreement, verb tense, number, word order/function, articles, pronouns, prepositions
	25-20	**GOOD TO AVERAGE:** inconsistent control of simple, compound and/or complex sentences—minor problems in the use of coordinators, subordinators, and transitions—several errors of S-V agreement, verb tense, number, word order/function, articles, pronouns, prepositions
	19-14	**FAIR TO POOR:** major problems in simple, compound and/or complex sentences—frequent errors of negation, agreement, tense, number, word order/function, articles, pronouns, prepositions and/or fragments, run-ons, comma splice—meaning confused or obscured
	13-6	**VERY POOR:** virtually no mastery of sentence construction rules—dominated by errors—does not communicate—OR not enough to evaluate
Mechanics	**5**	**EXCELLENT TO VERY GOOD:** demonstrates mastery of conventions—few errors of spelling, punctuation, capitalization—includes clearly defined paragraphs and title—legible handwriting
	4	**GOOD TO AVERAGE:** occasional errors of spelling, punctuation, capitalization—unclear paragraphing—but meaning not obscured
	3	**FAIR TO POOR:** frequent errors of spelling, punctuation, capitalization, paragraphing—poor handwriting—meaning confused or obscured
	2	**VERY POOR:** no mastery of conventions—dominated by errors of spelling, punctuation, capitalization, paragraphing—handwriting illegible—OR not enough to evaluate
OFF	**0**	**Off-topic and /or off-genre Failure to understand and /or complete the task**

· Adapted from Holly L. Jacobs, V. Faye Hartflel, Jane B. Hughey, and Deanna R. Wormut (1981). Newbury House Publisher.

Appendix-III. Consent Form

CONSENT FORM

Your signature below indicates that you have decided to volunteer as a research participant for this study, and that you have given your consent that your essay will be used in building a corpus of Turkish students' written English.

Student's Name	Signature
1. Murat AKKAYA	
2. Yusuf Can AKARSU	
3. Özge GÜNAY	
4. Ümmü Gülsüm TUFAN	
5. Gizem KURNALI	
6. Elifnur DAĞCI	
7. Ümit BIYIK	
8. Mehmet KOCATÜRK	
9. Korcan BAŞARAN	
10. Hande Direnç AÇIKGÖZ	
11. Simge ÇAL	
12. Özge OCAK	
13. Esra Nur ÖZ	
14. Muhammet Furkan GİRGİN	
15. Mehmet YILMAZ	
16. Tuğba DİRLİKLİ	
17. Didem DEMİR	
18. Ozan Fırat BARAN	
19. Büşra COŞKUN	
20. Mücahit BAYRAM	
21. Emel BAYHAN	
22. Burak KESİM	
23. Mehmet ACAR	
24. Nurşah KARACA	
25. Furkan ÜN	
26. Oguljahan GELDİGURBANOVA	
27. Alina SOVDEKOVA	

Appendix-IV. Essay and Vocabulary Scoring of the two raters (1st Year Essays)

	Rater 1		Rater 2	
Essays	Vocab	Overall	Vocab	Total
1	8	61	10	51
2	8	66	8	76
3	6	55	7	64
4	6	60	7	53
5	8	61	6	68
6	5	49	6	59
7	7	53	7	61
8	9	65	7	57
9	9	68	9	61
10	8	59	8	67
11	4	33	4	42
12	6	45	6	40
13	7	57	7	67
14	5	22	3	18
15	12	87	12	92
16	6	53	7	63
17	7	53	7	61
18	6	38	6	43
19	8	57	8	63
20	13	88	11	82
21	11	76	11	82
22	10	67	10	75
23	9	53	9	61
24	8	55	8	50
25	9	52	8	47
26	11	81	11	73
27	10	55	11	63
28	9	56	10	63
29	11	76	11	70
30	8	48	8	55
31	10	67	10	75
32	9	57	9	66
33	8	49	7	45
34	7	47	7	54

35	7	43	7	52
36	9	64	9	71
37	8	55	8	63
38	6	45	6	53
39	7	52	7	61
40	9	58	7	50
41	12	69	9	61
42	5	40	5	49
43	7	43	7	51
44	12	75	10	70
45	11	66	11	72
46	12	68	10	60
47	10	70	9	63
48	9	54	9	61
49	8	52	8	61
50	6	38	6	46
51	11	73	11	80
52	12	73	12	64
53	10	65	10	73
54	7	49	7	56
55	7	48	7	55
56	8	50	8	57
57	7	39	7	45
58	7	44	7	51
59	8	54	7	45
60	6	40	6	47
61	5	36	5	44
62	8	51	8	58
63	6	35	6	43
64	7	48	7	55
65	6	39	6	48
66	7	47	7	40
67	6	41	6	47
68	5	38	5	33
69	2	14	2	20
70	5	29	5	35
71	5	34	5	41

72	7	48	7	40
73	7	51	7	42
74	6	52	6	60
75	7	53	7	45
76	4	21	4	28
77	5	38	5	44
78	8	56	8	47
79	5	27	5	35
80	6	33	6	40
81	7	51	7	61
82	8	52	8	60
83	7	47	7	55
84	5	20	5	28
85	7	41	7	35
86	8	41	8	49
87	4	22	4	31
88	3	17	3	24
89	3	14	3	20
90	9	59	9	66
91	8	49	8	54
92	10	68	9	61
93	9	55	9	64
94	6	36	6	44
95	7	51	7	60
96	7	48	7	56
97	5	32	5	40
98	8	55	8	62
99	7	43	7	52
100	11	72	11	64
101	7	44	7	51
102	8	48	7	40
103	7	48	7	40
104	8	47	8	41
105	5	26	5	33
106	6	37	6	44
107	6	42	6	51
108	8	53	8	43

109	6	41	6	48
110	7	52	7	59
111	14	91	13	84
112	11	71	11	79
113	7	50	7	58
114	9	61	9	67
115	8	56	8	63
116	9	61	9	68
117	8	55	8	64
118	8	61	8	54
119	9	61	9	67
120	10	56	10	66
121	9	60	9	68
122	8	58	8	64
123	12	83	12	75
124	9	61	9	67
125	8	52	8	61
126	9	67	9	60
127	6	34	6	41
128	7	34	7	41
129	8	52	8	61
130	9	49	9	57
131	7	48	7	40
132	11	72	11	81
133	6	47	6	55
134	8	56	8	64
135	7	41	7	47
136	6	45	6	51
137	7	47	7	53
138	6	46	6	40
139	6	40	6	47
140	5	35	5	41
141	8	46	8	40
142	10	59	10	64
143	9	60	9	52
144	7	41	7	48
145	6	35	6	41

146	8	50	8	58
147	10	56	10	62
148	12	82	12	74
149	9	52	9	59
150	7	46	7	51
151	6	43	6	50
152	8	56	8	61
153	9	55	9	62
154	11	64	11	70
155	9	53	9	63
156	8	59	8	65
157	7	46	7	52
158	8	52	8	43
159	6	36	6	44
160	8	54	8	48
161	3	17	3	22
162	8	56	8	61
163	6	39	6	45
164	8	52	8	59
165	7	41	7	51

Appendix-V. Essay and Vocabulary Scoring of the two raters (4th Year Essays)

	Rater 1		Rater 2	
Essays	Vocab	Total	Vocab	Total
1	8	53	8	57
2	9	57	9	60
3	7	45	7	47
4	7	48	7	50
5	4	27	4	32
6	5	35	5	37
7	5	28	5	33
8	5	26	5	30
9	5	28	5	30
10	6	33	6	36
11	7	40	7	44
12	5	28	5	29
13	6	34	6	39
14	11	70	11	74
15	7	48	7	50
16	8	50	8	54
17	7	46	7	47
18	9	59	9	62
19	6	44	6	46
20	7	44	7	45
21	9	67	9	74
22	12	83	12	85
23	8	56	8	57
24	10	69	10	71
25	9	59	9	63
26	9	53	9	53
27	11	73	11	75
28	10	69	10	73
29	9	63	9	64
30	10	69	10	75
31	13	85	13	89

32	12	80	12	82
33	9	60	9	62
34	5	32	5	34
35	8	56	8	59
36	7	50	7	53
37	8	63	8	67
38	9	63	9	66
39	11	75	11	78
40	6	42	6	44
41	6	44	6	46
42	5	25	5	30
43	10	57	10	61
44	9	61	9	67
45	8	49	8	49
46	7	46	7	50
47	8	48	8	53
48	8	53	8	57
49	10	64	10	70
50	6	38	6	41
51	7	46	7	52
52	8	53	8	57
53	10	71	10	72
54	8	57	8	59
55	6	45	6	49
56	11	77	11	81
57	7	50	7	54
58	8	57	8	60
59	7	53	7	54
60	4	28	4	28
61	6	47	6	49
62	6	39	6	40
63	7	44	7	46
64	6	40	6	43
65	9	63	9	66
66	7	46	7	46
67	6	40	6	40
68	11	78	11	80

69	8	58	8	61
70	8	54	8	54
71	12	76	12	78
72	8	56	8	58
73	7	40	7	44
74	10	67	10	72
75	5	32	5	32
76	8	54	8	56
77	6	46	6	51
78	12	76	12	77
79	13	84	13	87
80	6	41	6	43
81	8	50	8	54
82	7	44	7	47
83	7	46	7	50
84	8	56	8	59
85	6	35	6	35
86	8	51	8	53
87	6	36	6	37
88	6	32	6	33
89	4	19	4	24
90	5	36	5	40
91	6	38	6	39
92	8	49	8	54
93	7	46	7	48
94	6	45	6	47
95	10	71	10	77
96	6	47	6	47
97	8	60	8	64
98	11	76	11	78
99	12	73	12	74
100	10	71	10	72
101	8	53	8	56
102	10	71	10	73
103	5	40	5	40
104	11	79	11	83
105	6	46	6	46

106	9	65	9	68
107	7	56	7	58
108	7	48	7	48
109	11	81	11	82
110	12	82	12	86
111	10	67	10	67
112	8	51	8	54
113	6	39	6	39
114	9	68	9	71
115	10	70	10	73
116	7	51	7	53
117	8	55	8	55
118	11	71	11	76
119	12	76	12	77
120	8	58	8	62
121	10	70	10	72
122	7	51	7	53
123	7	52	7	53
124	6	39	6	39
125	7	47	7	49
126	10	63	10	63
127	11	69	11	69
128	9	60	9	64
129	8	51	8	52
130	6	35	6	35
131	9	62	9	67
132	4	25	4	26
133	5	26	5	30
134	7	42	7	42
135	7	43	7	44
136	8	50	8	50
137	5	33	5	35
138	5	30	5	31
139	6	37	6	39
140	8	52	8	52
141	7	51	7	54
142	7	46	7	49

143	6	41	6	41
144	7	49	7	52

Appendix-VI. Overall Data Set (1st Year Essays)

				Vocabulary Range		
Essays	Word	Lexical Density	VocdD	AWL	Off-list	AWL+Off-list
1	256	47.98%	101.642	1,18	0,78	1,96
2	240	52.32%	77.465	6,25	1,25	7,5
3	272	50.57%	63.363	5,88	5,51	11,39
4	214	52.13%	84.257	5,14	7,01	12,15
5	302	49.83%	110.339	2,99	2,99	5,98
6	193	51.6%	76.887	5,18	3,63	8,81
7	165	53.75%	77.083	9,2	4,91	14,11
8	234	50,00%	91.265	9,44	6,01	15,45
9	276	46.79%	68.306	5,13	1,83	6,96
10	341	49.11%	68.128	5,9	5,31	11,21
11	123	49.59%	62.702	7,32	4,07	11,39
12	178	52.33%	70.862	8,62	4,6	13,22
13	232	52.51%	86.629	4	3,56	7,56
14	83	50.6%	0	9,64	4,82	14,46
15	635	49.61%	97.893	3,48	4,42	7,9
16	226	56.54%	100.11	5,48	3,2	8,68
17	250	47.74%	63.548	4,53	1,65	6,18
18	109	53.7%	115.958	5,5	13,76	19,26
19	283	45.85%	87.778	3,24	1,8	5,04
20	485	47.88%	96.21	2,7	2,9	5,6
21	406	53.5%	95.123	4,73	6,22	10,95
22	339	48.8%	94.087	3,85	4,44	8,29
23	168	55.41%	90.091	6,21	6,83	13,04
24	193	57.75%	61.348	5,18	3,11	8,29
25	272	54.85%	132.734	2,57	6,99	9,56
26	438	48.72%	62.963	2,74	3,65	6,39
27	196	50,00%	42.962	4,21	3,68	7,89
28	244	48.95%	66.882	5,44	3,77	9,21
29	287	51.26%	90.82	1,09	4	5,09
30	233	48.25%	58.097	6,61	4,41	11,02
31	377	47.14%	93.817	3,8	4,08	7,88
32	319	52.66%	84.743	3,77	3,77	7,54
33	216	49.53%	94.531	6,13	2,36	8,49

34	145	58.62%	85.17	8,22	7,53	15,75
35	155	52.94%	77.238	9,15	1,31	10,46
36	301	52.88%	74.719	8,81	3,73	12,54
37	240	46.35%	76.34	7,69	3,42	11,11
38	173	53.53%	99.341	2,94	4,71	7,65
39	219	46.54%	56.421	7,83	2,3	10,13
40	206	53.96%	85.546	6,93	4,95	11,88
41	287	54.29%	71.019	10,71	1,07	11,78
42	175	58.72%	111.263	10,92	2,3	13,22
43	193	43.92%	73.861	2,65	6,88	9,53
44	323	49.06%	91.772	4,38	3,12	7,5
45	307	55.08%	90.592	2,65	10,93	13,58
46	319	54.17%	131.708	6,39	7,67	14,06
47	374	54.69%	55.793	3,75	2,95	6,7
48	240	48.1%	75.749	4,26	3,83	8,09
49	323	54.6%	143.492	5,7	8,54	14,24
50	233	51.08%	94.073	3,48	3,48	6,96
51	366	48.91%	51.256	13,9	3	16,9
52	352	51.86%	92.09	3,45	4,02	7,47
53	269	53.96%	69.699	7,6	6,08	13,68
54	235	43.97%	64.458	3,02	5,6	8,62
55	213	52.15%	104.284	6,22	6,7	12,92
56	212	50.72%	41.678	5,31	3,86	9,17
57	285	50.18%	111.886	8,57	5,71	14,28
58	231	50.22%	78.148	9,13	4,78	13,91
59	272	49.45%	83.785	6,27	2,21	8,48
60	199	51.26%	67.994	12,56	3,52	16,08
61	220	47.03%	65.076	1,83	10,5	12,33
62	274	49.63%	82.105	3,3	8,79	12,09
63	171	55.63%	71.156	5,56	10,49	16,05
64	217	53.46%	56.106	1,38	1,38	2,76
65	134	56.39%	60.492	9,02	4,51	13,53
66	223	59.19%	70.194	12,61	6,31	18,92
67	149	55.1%	86.488	14,97	8,84	23,81
68	161	49.36%	80.356	7,01	5,1	12,11
69	139	51.8%	84.24	10,71	3,57	14,28
70	193	56.99%	71.042	4,57	4,06	8,63

71	131	68.46%	56.068	13,08	1,54	14,62
72	165	51.52%	93.002	4,85	4,85	9,7
73	161	56.25%	71.367	10,62	4,38	15
74	191	52.38%	56.826	4,76	4,23	8,99
75	177	49.13%	78.06	9,25	3,47	12,72
76	99	44.44%	0	5,05	0	5,05
77	181	50.83%	75.343	4,4	4,4	8,8
78	221	51.16%	55.249	6,51	3,26	9,77
79	144	51.39%	84.613	5,59	5,59	11,18
80	172	48.82%	105.929	7,69	2,37	10,06
81	199	50.51%	80.197	4,52	3,52	8,04
82	158	55.06%	58.03	10,76	6,33	17,09
83	137	57.46%	57.807	8,89	4,44	13,33
84	214	52.45%	105.879	1,95	4,88	6,83
85	181	50,00%	80.944	5,11	7,39	12,5
86	182	59.78%	70.409	6,21	1,69	7,9
87	293	44.91%	83.266	2,11	3,51	5,62
88	138	55.15%	71.021	3,68	0,74	4,42
89	191	43.68%	87.934	3,68	3,16	6,84
90	213	56.25%	98.342	6,7	7,18	13,88
91	174	55.81%	87.995	2,91	4,65	7,56
92	295	49.31%	121.968	4,48	9,66	14,14
93	199	50,00%	88.735	6,22	7,77	13,99
94	139	52.9%	48.355	5,15	3,68	8,83
95	237	55.13%	74.042	6,84	4,27	11,11
96	164	49.03%	62.42	7,1	4,52	11,62
97	125	47.2%	61.852	3,94	0,79	4,73
98	219	49.77%	71.645	2,82	4,69	7,51
99	158	53.64%	56.467	7,24	3,95	11,19
100	292	50.7%	77.783	6,34	3,87	10,21
101	256	45.85%	89.64	5,95	6,35	12,3
102	213	50.47%	60.791	6,13	4,72	10,85
103	226	46.82%	87.395	7,27	3,18	10,45
104	176	44.83%	53.79	4,02	2,87	6,89
105	101	55.45%	83.445	6,93	4,95	11,88
106	216	53.02%	71.663	5,14	3,27	8,41
107	165	50,00%	93.25	9,38	3,75	13,13

108	234	54.59%	101.647	7,42	3,49	10,91
109	160	44.94%	57.742	3,16	1,9	5,06
110	208	47.78%	89.105	5,94	3,96	9,9
111	592	49.74%	107.012	3,44	2,58	6,02
112	297	51.56%	80.948	7,99	4,17	12,16
113	202	53.3%	97.739	3,55	8,12	11,67
114	241	44.35%	71.307	6,25	3,33	9,58
115	267	41.22%	68.777	3,45	5,75	9,2
116	243	48.54%	67.976	7,11	2,09	9,2
117	218	45.5%	58.103	4,74	4,74	9,48
118	221	44.5%	93.82	6,88	3,21	10,09
119	257	53.17%	114.939	6,75	6,35	13,1
120	257	50.98%	104.302	5,08	2,34	7,42
121	206	53.69%	56.596	11,33	2,46	13,79
122	261	56.98%	99.683	3,83	4,21	8,04
123	342	50.58%	65.344	6,14	4,68	10,82
124	188	56.52%	96.219	7,03	5,95	12,98
125	200	52.85%	74.414	4,66	2,07	8,45
126	232	52.89%	80.151	5,78	2,67	13,47
127	196	49.22%	58.507	6,22	7,25	13,47
128	364	48.47%	109.31	5,57	2,51	8,08
129	286	46.91%	83.879	5,42	3,97	9,39
130	275	49.45%	97.845	3,28	5,11	8,39
131	193	47.59%	65.555	5,35	1,6	6,95
132	309	50.83%	87.275	6,6	3,96	10,56
133	144	51.43%	66.634	7,14	2,86	10
134	200	52.31%	83.284	4,62	2,05	6,67
135	217	47.39%	100.172	2,36	3,3	5,66
136	151	55.03%	126.921	5,37	5,37	10,74
137	174	55.75%	91.245	5,17	1,72	6,89
138	189	50.27%	72.099	2,67	6,42	9,09
139	161	53.8%	84.726	6,33	6,33	12,66
140	165	51.88%	67.013	8,12	3,12	11,24
141	182	52.81%	80.313	5,06	5,06	10,12
142	434	43.69%	78.837	2,34	3,97	6,31
143	254	52.46%	88.918	9,43	3,28	12,71
144	193	43.39%	73.581	3,72	4,26	7,98

145	107	58.49%	59.852	3,77	3,77	7,54
146	272	52.43%	105.39	4,12	3,75	7,87
147	227	51.35%	71.297	4,07	6,33	10,4
148	497	51.23%	117.958	3,27	10,82	14,09
149	211	43.27%	69.645	6,28	2,9	9,18
150	204	44.78%	59.629	2,99	0,5	3,49
151	170	53.25%	62.755	4,17	2,38	6,55
152	233	59.48%	89.655	6,47	4,31	10,78
153	255	47.01%	79.229	4,78	1,2	5,98
154	275	58.8%	110.925	6,06	10,23	16,29
155	195	54.45%	74.097	7,85	8,9	16,75
156	291	42.81%	53.743	3,85	3,85	7,7
157	189	49.73%	59.453	3,26	4,89	8,15
158	245	54.92%	91.075	3,25	7,72	10,97
159	148	49.66%	88.709	4,14	8,97	13,11
160	224	42.4%	64.121	5,56	4,63	10,19
161	65	55.56%	0	8,06	0	8,06
162	297	43.84%	71.329	3,78	2,06	5,84
163	231	47.16%	75.316	1,75	3,95	5,7
164	222	47.47%	75.392	4,17	1,85	6,02
165	281	47.48%	66.806	2,9	6,88	9,78

Appendix-VII. Overall Data Set (4th Year Essays)

					Vocabulary Range	
Essays	Word	LD	Vocd-D	AWL	Off-list	AWL+Off-list
1	273	56.09%	98.739	6,88	5,07	11,95
2	278	56.93%	88.11	7,84	5,6	13,44
3	249	45.34%	66.301	1,21	1,62	2,83
4	214	48.56%	77.729	6,8	1,46	8,26
5	125	52.94%	75.032	4,1	2,46	6,56
6	244	50.83%	103.56	3,75	6,67	10,42
7	149	49.65%	52.293	4,2	0,7	4,9
8	153	52.7%	89.7	5,41	5,41	10,82
9	210	54.07%	94.337	6,22	6,7	12,92
10	167	44.24%	61.026	4,88	6,1	10,98
11	255	53.94%	82.93	7,87	1,57	9,44
12	136	49.63%	52.717	5,93	3,7	9,63
13	149	56.55%	100.765	4,86	3,47	8,33
14	252	47.35%	62.271	6,94	1,63	8,57
15	286	46.1%	78.037	2,81	2,11	4,92
16	268	52.69%	97.815	10,38	3,46	13,84
17	408	47.37%	88.66	1	3,01	4,01
18	416	47.56%	94.866	8,29	3,41	11,7
19	222	51.35%	105.416	7,66	5,86	13,52
20	269	48.28%	100.074	6,51	3,07	9,58
21	260	54.05%	123.641	10,81	4,63	15,44
22	375	51.07%	80.959	4,02	6,7	10,72
23	245	46.67%	81.516	6,22	5,81	12,03
24	340	50.75%	55.856	8,77	4,39	13,16
25	386	58.59%	104.543	7,22	3,87	11,09
26	443	53.46%	63.727	6,24	2,54	8,78
27	380	54.28%	70.779	5,63	1,07	6,7
28	305	46.96%	67.401	6,73	1,68	8,41
29	314	50.49%	77	11,26	2,98	14,24
30	388	52.76%	117.319	3,15	5,51	8,66
31	374	52.01%	92.125	9,09	3,74	12,83
32	363	46.93%	114.383	5,59	7,26	12,85
33	284	46.76%	68.596	5,78	2,89	8,67

34	197	46.94%	93.311	6,67	3,59	10,26
35	269	52.26%	72.465	8,58	2,24	10,82
36	236	51.74%	66.104	10,04	2,62	12,66
37	335	51.67%	113.107	5,52	7,98	13,5
38	336	49.1%	85.418	3,31	2,11	5,42
39	353	54.62%	104.817	9,2	6,61	15,81
40	200	53.3%	96.748	3,54	4,04	7,58
41	215	50.47%	77.817	4,25	3,77	8,02
42	312	44.04%	81.839	4,28	3,62	7,9
43	268	49.62%	109.328	4,91	5,66	10,57
44	310	45.21%	83.066	6,6	4,62	11,22
45	240	54.08%	96.708	2,95	2,95	5,9
46	189	49.72%	118.317	4,97	3,87	8,84
47	269	52.27%	68.807	7,55	2,26	9,81
48	268	49.43%	76.897	7,58	4,55	12,13
49	351	50.72%	91.872	5,48	3,17	8,65
50	165	53.37%	128.334	8,43	6,02	14,45
51	243	49.58%	101.053	3,35	5,86	9,21
52	293	43.01%	93.614	3,5	5,24	8,74
53	267	53.85%	114.54	5,43	5,43	10,86
54	262	50.57%	84.507	5,73	4,58	10,31
55	230	54.5%	74.625	4,5	2,7	7,2
56	522	47.47%	119.487	5,08	5,66	10,74
57	276	49.82%	78.814	4,71	3,99	8,7
58	279	41.07%	69.707	8,3	3,61	11,91
59	269	49.07%	65.853	5,28	5,28	10,56
60	132	41.35%	54.045	3,05	2,29	5,34
61	214	48.34%	50.335	7,11	5,21	12,32
62	212	51.94%	96.1	11	3,5	14,5
63	236	59.48%	141.566	6,03	12,07	18,1
64	211	54.37%	86.471	6,8	3,88	10,68
65	365	47.62%	78.038	5,03	3,07	8,1
66	193	45.65%	69.94	7,07	2,72	9,79
67	261	48.45%	82.633	6,59	4,26	10,85
68	543	52.68%	95.496	5,74	4,63	10,37
69	333	50.61%	79.848	4,6	5,21	9,81
70	317	56.13%	119.53	10,65	6,13	16,78

71	355	53.56%	58.943	9,97	7,12	17,09
72	251	53.82%	102.025	6,85	8,47	15,32
73	247	49.59%	97.862	6,15	6,56	12,71
74	444	51.82%	114.414	3,41	4,77	8,18
75	227	52.47%	70.578	5,36	4,91	10,27
76	302	46.67%	108.504	3	4,67	7,67
77	224	52.73%	91.93	5,45	4,55	10
78	460	49.56%	68.951	4,79	5,01	9,8
79	503	51.41%	102.307	5,65	4,84	10,49
80	334	55.76%	99.012	5,15	3,94	9,09
81	222	52.73%	53.491	4,09	3,64	7,73
82	293	54.48%	139.04	5,38	5,38	10,76
83	262	50,00%	73.318	9,68	3,23	12,91
84	221	49.31%	116.537	6,45	5,07	11,52
85	348	49.13%	85.499	5,29	3,53	8,82
86	173	51.76%	98.298	4,71	5,88	10,59
87	201	49.24%	115.695	8,08	6,57	14,65
88	145	52.41%	61.629	10,42	4,17	14,59
89	148	51.37%	65.08	10,27	4,79	15,06
90	79	52.7%	0	9,33	10,67	20
91	196	58.55%	134.091	5,76	8,38	14,14
92	245	54.29%	131.629	6,22	7,88	14,1
93	318	54.26%	108.434	3,46	4,4	7,86
94	263	52.73%	97.383	4,69	4,69	9,38
95	254	46.37%	92.811	9,27	5,65	14,92
96	274	54.41%	74.006	2,95	2,21	5,16
97	191	61.78%	96.584	8,38	7,85	16,23
98	341	50.29%	102.739	4,68	5,56	10,24
99	308	47.84%	69.173	9,33	3,33	12,66
100	341	46.73%	58.254	4,39	4,68	9,07
101	248	51.03%	80.525	8,54	2,44	10,98
102	200	53.77%	89.622	10	3	13
103	363	46.41%	70.963	6,91	1,38	8,29
104	197	49.73%	53.63	3,08	2,05	5,13
105	566	47.86%	92.684	6,19	7,08	13,27
106	216	53.05%	72.128	3,24	3,24	6,48
107	362	54.37%	95.338	8,91	2,51	11,42

108	226	46.46%	57.342	9,73	3,54	13,27
109	253	50,00%	68.154	4,8	4,4	9,2
110	449	55.36%	102.124	8,28	4,25	12,53
111	212	48.67%	45.86	5,34	2,67	8,01
112	507	54.45%	89.25	8,93	3,57	12,5
113	285	45.41%	111.658	6,8	2,43	9,23
114	150	52.74%	99.972	6,8	4,76	11,56
115	350	44.13%	67.12	3,15	3,15	6,3
116	407	46.15%	65.835	3,59	2,82	6,41
117	200	54.77%	52.34	13,13	0,51	13,64
118	246	49.17%	98.221	7,08	5,42	12,5
119	409	51.62%	89.327	4,49	4,49	8,98
120	422	53.57%	126.65	9,95	5,45	15,4
121	326	47.69%	60.808	4,31	8	12,31
122	355	50.86%	70.173	4,01	4,3	8,31
123	334	44.85%	87.721	5,78	3,34	9,12
124	300	56.66%	72.284	8,19	3,75	11,94
125	283	56.09%	123.621	4,83	4,83	12,54
126	312	51.16%	81.368	3,63	8,91	15,36
127	486	53.38%	103.532	9,89	5,47	15,36
128	345	54.81%	100.646	6,73	6,14	12,87
129	372	53.17%	107.938	7,16	3,58	10,74
130	325	51.42%	103.856	5,4	6,35	11,75
131	193	51.04%	97.91	6,81	6,81	13,62
132	207	48.77%	90.609	2,49	8,46	10,95
133	151	53.33%	92.536	5,33	2,67	8
134	266	49.81%	124.011	5,66	7,92	13,58
135	188	52.72%	97.129	6,52	3,26	9,78
136	253	53.82%	116.995	4,02	4,82	8,84
137	218	46.26%	117.015	2,34	8,88	11,22
138	137	59.4%	54.862	6,02	2,26	8,28
139	166	58.54%	139.327	5,49	9,76	15,25
140	261	58.62%	106.836	5,75	2,68	8,43
141	315	47.76%	81.914	5,1	1,59	6,69
142	306	54.13%	102.368	6,93	5,94	12,87
143	237	43.48%	67.033	6,52	6,96	13,48
144	282	46.07%	72.178	8,93	3,93	12,86