T.C.

YEDİTEPE UNIVERSITY

GRADUATE SCHOOL OF EDUCATIONAL SCIENCES

THE EFFECTS OF EXPLICIT TEACHING OF FORMULAIC

LANGUAGE ON ACADEMIC WRITING

Armağan BAKILI AKKOÇ

İstanbul, 2017

THE EFFECTS OF EXPLICIT TEACHING OF FORMULAIC

LANGUAGE ON ACADEMIC WRITING

by

ARMAĞAN BAKILI AKKOÇ

Dissertation submitted in partial fulfilment of the requirements for the

degree of Master of Arts

in

Department of English Language Teaching

YEDITEPE UNIVERSITY

March, 2017

YEDİTEPE ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ

TEZ TESLİM ve ONAY TUTANAĞI

KONU: The Effects of Explicit Teaching of Formulaic Language on Academic Writing

ONAY:

Dr. Erkan Karabacak

(Danışman)   (İmza)

(varsa Eş-Danışman)   (İmza)

Yard. Doç. Dr. Evrim Eveyik Aydın

(Üye)   (İmza)

Sibel Tatar

(Üye)   (İmza)

TESLİM EDEN   : Armağan BAKILI AKKOÇ
TEZ SAVUNMA TARİHİ   : 27.03.2017
TEZ ONAY TARİHİ   : 27.03.2017

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Armağan BAKILI AKKOÇ

ABSTRACT

The present study investigated the effects of the explicit teaching of formulaic language on the overall quality of academic writing, specifically argumentative writing of university students. Participants are 44 freshman and 27 sophomore students at the ELT Department of Yeditepe University which were grouped as the experimental and the control. Before the intervention, a pre-test was conducted to determine the use of formulaic language and the overall quality of the essays. Then, the experimental group was explicitly instructed on 40 target formulaic language items, an organisational template constructed for argumentative writing and three thesis statement structures for two weeks. After the intervention a post-test, a final test and a delayed post-test were given to all students in order to see whether the explicit instruction had any effect on the use of formulaic language that was taught and the overall quality of writing.

The descriptive statistics indicated that the experimental group increased the overall quality scores of their writing after the intervention and this increase was found to be statistically significant through Wilcoxon signed rank test. Although a decrease was observed in the delayed post-test of the experimental group students, the mean scores of them in the overall quality of the essays were higher than those of the control group in each test, which was also found to be statistically significant by the inferential statistics. On the other hand, the overall quality scores of the control group obtained from the essays did not change immediately but increased later in the final test; however, the Friedman test revealed that this increase was not statistically significant. Mann-Whitney U test results also indicated that there was a statistically significant difference between the experimental and the control groups regarding both the type and the tokens of target formulaic language items used. After the intervention, experimental

group students used greater number and more variety of formulaic language items in each test than their counterparts in the control group. Moreover, a positive correlation was found between the frequency of the formulaic language items used and the overall quality scores of the essays. The explicit teaching of organisational template constructed for argumentative writing had some positive effects on the organisation of the essays as well since the percentage of the students following the recommended templates gradually increased for the experimental group while it decreased for the control group. On the other hand, the results indicated that the students of both groups did not use some of the elements of the recommended template in their essays and mostly missed elements were mainly related to stating the opposing views. Lastly, the thesis statement writing was also affected by explicit teaching, as the percentage of the students at the pre-test who wrote a thesis statement not in compliance with any recommended structures gradually decreased at each test after intervention.

In light of these findings, it could be stated that the explicit instruction of the target formulaic language items raised the awareness of the students about the use of formulaic language and improved the overall quality of the argumentative writing, but it did not lead to any consistent improvement on students' later production in the long term. However with the help of pedagogical approaches mentioned in this study explicit teaching might be helpful for the students to learn formulaic language and improve their writing and it should be the focus of further research considering the possible limitations.

ÖZET

Bu çalışma, kalıplaşmış dil ifadelerini doğrudan açıkça derste öğretmenin, üniversite öğrencilerinin bu ifadeleri kullanımlarına ve akademik yazılarının genel kalitesine, özellikle tartışma yazılarına olan etkisini incelemiştir. Katılımcılar, kontrol ve deney olarak gruplandırılan Yeditepe Üniversitesi İngilizce Öğretmenliği Bölümü 44 birinci sınıf 27 ikinci sınıf öğrencileridir. Öğretimden önce, öğrencilerin kalıplaşmış dil ifadelerini kullanımlarımı ve akademik yazma becerilerinin genel kalitesini belirleyebilmek için bir ön test uygulanmıştır. Daha sonra, 40 adet hedef kalıplaşmış dil ifadesi, tartışma yazıları için oluşturulan bir organizasyon şeması ve üç değişik tez cümlesi yapısı deney grubuna, iki hafta boyunca doğrudan açıkça derste öğretilmiştir. Öğretimden sonra ise kalıplaşmış dil ifadeleri kullanımında ve akademik yazma becerilerinin genel kalitesi üzerinde doğrudan açıkça derste öğretmenin bir etkisi olup olmadığını belirlemek için tüm öğrencilere bir son test, bir final testi ve bir gecikmeli son test verilmiştir.

Betimleyici istatistikler, deney grubunun yazılarının genel kalite puanlarının öğretimden sonra arttığını göstermiş ve bu artış Wilcoxon Signed Rank test ile de istatistiksel olarak anlamlı bulunmuştur. Deney grubu öğrencilerinin gecikmeli son test sonuçlarında bir düşüş gözlenmiş olmasına rağmen her testteki yazılarının genel kalitesine ilişkin ortalama puanları kontrol grubu öğrencilerininkinden daha yüksek çıkmış ve bu çıkarsamalı istatistik ile anlamlı bulunmuştur. Diğer bir yandan, kontrol grubu öğrencileri tarafından yazılan tartışma yazılarının genel kalitesine ilişkin puanlarda hemen bir değişiklik görülmemiş fakat daha sonra final testinde bir artış görülmüştür, ama yine de Friedman testi bu artışın istatistiksel olarak anlamlı olmadığı ortaya çıkarmıştır. Mann-Whitney U test sonuçları ise deney ve kontrol grubu arasında

hem kalıplaşmış dil ifadeleri sayısı hem de çeşidi açısından anlamlı fark olduğunu göstermiştir. Öğretimden sonra deney grubu öğrencileri kontrol grubu öğrencilerine kıyasla sayıca daha fazla ve daha çeşitli kalıplaşmış dil ifadeleri kullanmıştır. Buna ek olarak, kalıplaşmış dil ifadelerinin kullanım sıklığı ile yazıların genel kalite puanları arasında da pozitif bir korelasyon bulunmuştur. Tartışma yazıları için oluşturulan organizasyon şemasının doğrudan açıkça derste öğretilmesinin de yazıların organizasyonu üzerinde olumlu etkileri olmuştur. Tavsiye edilen şemayı kullanan öğrenci yüzdesi deney grubu için giderek artarken bu yüzde kontrol grubu için azalmıştır. Bunun yanı sıra, bulgular, her iki grup öğrencilerinin, tavsiye edilen şemanın bazı öğelerini kullanmadıklarını ve bu eksik öğelerin genellikle karşı görüş bildirme ile ilgili olduğunu da ortaya çıkarmıştır. Son olarak, tez cümlesi yazımı da doğrudan açıkça derste öğretimden etkilenmiştir. Ön testte, öğretilen tez cümleleri ile hiçbir benzerlik göstermeyen tez cümleleri yazan deney grubu öğrencilerinin oranı, öğretimden sonra, giderek azalmıştır.

Bu bulgular ışığında, kalıplaşmış dil ifadelerini, doğrudan açıkça derste öğretimin öğrencilerin kalıplaşmış dil ifadelerini kullanmada farkındalıklarını artırdığı ve tartışmacı yazının genel kalitesini de geliştirdiği söylenebilir, fakat bu farkındalık, uzun vadede, öğrencilerin daha sonraki ürünlerinde kalıcı iyileşmeyi beraberinde getirmemiştir. Buna rağmen, bu çalışmada bahsedilen pedagojik yaklaşımların da yardımı ile doğrudan açıkça derste öğretim, öğrencilerin kalıplaşmış dil ifadelerini öğrenmelerinde ve yazılarını geliştirmelerinde faydalı olabilir ve muhtemel kısıtlamalar da dikkate alınarak gelecekteki çalışmaların odağı olmalıdır.

Anahtar kelimeler: kalıplaşmış dil ifadeleri, doğrudan açıkça derste öğretim, akademik yazma becerileri, tartışmacı yazma

To my beloved mother, Nermin Bakılı

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER I

1. INTRODUCTION

The issue of formulaic language has been the subject of growing interest to researchers recently, as it is thought to be a key component of language and essential to the way a second language is used, processed and acquired (Wood, 2002; Schmitt & Carter, 2004; Durrant, 2008; Millar, 2011) since it has many functions and also provides valuable data to understand language development (Wood, 2002; Boers et al., 2006; Wood, 2006; Jiang & Nekrasova, 2007; Ellis et al., 2008; Tremblay et al., 2011; Ellis, 2012; Meunier, 2012; Graham & Osment, 2014; Wood & Appel, 2014).

Formulaic language which can be seen in many forms such as collocations, lexical bundles, idioms and so forth, is found to have facilitative processing advantages not only for speakers but also hearers (Biber et al., 2004; Biber & Barbieri, 2007; Hyland, 2008a; Conklin & Schmitt, 2008; Bardovi-Harlig, 2009; Nekrosova, 2009; Liu & Huo, 2011; Boers & Lindstromberg, 2012; Ortaçtepe, 2013; Wray & Perkins, 2000), by reducing the cognitive load on the brain to generate or comprehend an utterance. This is because the lexical items embedded in formulaic language are stored and retrieved holistically (Jiang & Nekrasova, 2007; Kecskes, 2007; Conklin & Schmitt, 2012) as if they were a single lexical item (Schmitt & Carter, 2004; Wood, 2006; Wood, 2009; Alipour & Zarea, 2013). Thus, the speaker, without spending much time to find the required linguistic knowledge like grammatical rules, can directly and quickly say the intended utterance and, the hearer, in a similar way, can quickly understand the speaker's point without spending much time to decode it (Wray & Perkins, 2000; Hyland, 2008a).

Formulaic language also works as the building blocks of a discourse by helping to shape a speech or writing (Cortes, 2002; Schmitt & Carter, 2004; Biber et al., 2004; Biber & Barbieri, 2007; Wray, 2000, Hyland, 2008a; Hyland, 2008b; Nekrasova, 2009; Ahmadi et al., 2013; Alhassan & Wood, 2015) with the prefabricated sequences to introduce a topic, to elaborate and conclude the topic. Using formulaic language not only provides a technical appropriateness but also helps to sound natural and idiomatic. That is accepted as an indication of the proficiency and a key to admissibility to the discourse community in which formulaic language items are regarded as default expressions (Erman & Warren, 2000, Foster, 2001; Kuiper, 2004; Biber et al., 2004; Schmitt & Carter, 2004; Wood, 2006; Biber & Barbieri, 2007; Conklin & Schmitt, 2008; Ellis et al., 2008; Hyland, 2008a; Hyland, 2008b; Durrant & Schmitt, 2010; Millar, 2011; Liu, 2011; Meunier, 2012; Conklin & Schmitt, 2012).

In order to achieve idiomatic competence, it seems that language learners, are required to learn formulaic language items which exist ubiquitously in the language (Erman & Warren, 2000, Foster, 2001; Kuiper, 2004; Biber et al., 2004; Schmitt & Carter, 2004; Wood, 2006; Biber & Barbieri, 2007; Conklin & Schmitt, 2008; Ellis et al., 2008; Hyland, 2008a; Hyland, 2008b; Durrant & Schmitt, 2010; Millar, 2011; Liu, 2011; Meunier, 2012; Conklin & Schmitt, 2012). Otherwise, they, even the advanced ones, may have some challenges such as misunderstandings, communication breakdowns (Kecskes, 2007), sounding odd and unnatural (Ellis et al., 2008; Li & Schmitt, 2009, Ortaçtepe, 2013; Peters & Pauwels, 2015), being incompetent and correspondingly linguistically inappropriate to the related professional community (Ellis et al., 2008; Li & Schmitt, 2009, Ortaçtepe, 2013; Peters & Pauwels, 2015).

There may be various reasons underlying these challenges such as a slow learner, inept instructor, inadequate material or formulaic language itself (Jones & Haywood, 2004; Bardovi-Harlig, 2009; Liu, 2011; Laufer & Waldman, 2011; Martinez & Schmitt, 2012). For instance, a learner may not be aware of the formulaic language or ignore it, and also required input exposure and supportive exercises may not be provided by instructors or materials (Foster, 2001; Li & Schmitt, 2009; Byrd & Coxhead, 2010; Durrant & Schmitt, 2010; Szudarski & Carter, 2014). Moreover, formulaic language itself is confusing enough for learners since it is not always easy to separate a formulaic sequence from a generated sequence (Wray, 2000; Bishop, 2004a; Bishop, 2004b, Durrant, 2008; Laufer & Waldman, 2011).

To sum up, formulaic language can be said to be a benchmark in language learning. Despite its importance, it is surprising to expect it to be learned through exposure in the classrooms without adequate time. As a natural consequence of some deficiency and ignorance, the students may have difficulty in learning the formulaic language which is reckoned as vital for them. In order to overcome the challenges the students experience, explicit teaching may be a remedy in the classrooms with limited time. Thus, in the present study, the effects of the explicit teaching of formulaic language on the academic writing of the university students will be explored, with the expectation that the results will be helpful to instructors and researchers for further studies.

1.1. Purpose of the Study

Considering the importance of formulaic language, because of its plethora, functions and advantages, it is thought to be worth researching, especially in academic

writing which is believed to be the primary form of assessment and the mostly demanded skill for the university students regardless of their majors. Based on the researcher's own professional experience and her observations with colleagues, it is possible to assume that at university, students are exposed to such instruction methodologies (e.g. autonomous learning, research, paper writing, presentation and so on) and academic genres that they did not experience during their high school years. Most of the learners are used to studying for the university entrance exam which is a multiple choice test, by focussing on test techniques, rather than on critical thinking and producing the verbal or written work until they are enrolled at a university. It is, therefore, thought that students still have some shortcomings in academic writing no matter how good their grammar and vocabulary knowledge is. Although their writing seems technically competent, a gist of proficiency often seems missing. One of the reasons of this might be the deficiency of the learners in the knowledge of the formulaic language which may exist in so many complicated forms. This deduction is in line with the studies of Ohlrogge (2009), Kennedy and Thorp (2007), Read and Nation (2006) and Bonk (2001) who examined either written or oral productions of the learners taking part it high-stakes proficiency exams such as IELTS, ECCE and TOEFL. They found extensive use of formulaic language in the productions of the learners achieving higher scores compared to the ones achieving lower scores. Although the studies indicated that there was strong and significant associations between the second language proficiency and the use of formulaic language, the formulaic language is generally expected to be taught through exposure to the materials in and out of the classroom. However, the challenges that the learners experience and the lack of current methods and materials in acquiring formulaic language at native-like proficiency, mentioned in the introduction part, clearly indicate that there is a need for explicit instruction of formulaic language

supported by useful techniques and activities. This should be based on a well framed methodology with pedagogical concerns (Wray, 2000; Nesselhauf, 2003; Jones & Haywood, 2004; Cortes, 2004; Cortes, 2006; Li & Schmitt, 2009; Wood, 2009; Pang, 2010; Meunier, 2012) because simple exposure to formulaic language in written or spoken materials does not result in automatic acquisition (Cortes, 2002; Cortes, 2004; Jones & Haywood, 2004; Meunier, 2012). Thus, this dissertation will concentrate mainly on the effects of the explicit teaching of formulaic language in academic writing, with an expectation of minimizing current and further problems, and, accordingly, result in learners' improved overall performances.

1.2. Research Questions

The present study attempts to explore specifically the following research questions;

1. Is there any difference between the overall argumentative quality of the essays written by the students who are explicitly taught target formulaic language items and that of those written by the students who are not explicitly taught?

2. Is there any difference between the overall argumentative quality of the essays before and after the treatment?

3. Is there any difference in the use of target formulaic items between the pre-test essays and post-test essays by the experimental group, and also the experimental and the control group?

4. Is there a relationship between the number of target formulaic language items used and the overall argumentative quality of the essays, as well as between the

type of target formulaic language items used and the overall argumentative quality of the essays?

5. To what extent does the explicit teaching of organisational template and the structures of thesis statement affect the organisation of the essays?

1.3. Significance of the study

The present study aims to investigate whether the explicit instruction of formulaic language causes any gain in the overall argumentative quality of the academic writing of university students, as it is believed that they have limited exposure in and out of the classrooms to learn the formulaic language (Cortes, 2002; Cortes, 2004; Jones & Haywood, 2004; Meunier, 2012). That's why something practical, easier and effective, namely the explicit teaching, may be involved in the process of learning the formulaic language. In this context, the study is expected to be important in different dimensions: learner, instructor and curriculum, accordingly material.

First, this study is significant for learners as they are instructed to be more aware of formulaic language and its functions in their writing which is believed to be one of the most demanding challenges university students encounter. Moreover, formulaic language is not only important for their writing but also for other skills like being prepared for speaking, being able to summarise reading, and to take notes while listening. Moreover, in many majors, students are asked to write argumentative papers as part of their assignments and exams, in which they are supposed to be both correct and competent in language use. Being consciously aware of the use of formulaic language, they may sound like a more competent member of that community and be a more proficient language user.

For instructors, this study is of capital importance as well, as it could be impetus to evaluate their teaching methods and evolve them. The instructors are also alerted to be aware of formulaic language, with the purpose of preventing it being ignored, and to include it in their teaching methods and the materials in order to be more helpful to their students.

Lastly, the present study can be accepted to be important from the perspective of the curriculum and the materials used for the related course, by revealing some pedagogical implications which may be crucial to review the current curriculum and the materials. Accordingly, the curriculum can be re-designed and the materials can be edited. Moreover, the instruments to be used in this study can be converted into new materials for the relevant courses.

To sum up, the present study is expected to be important in almost all parties of a learning procedure, namely, learner, instructor, curriculum and materials, by presenting the current literature, the results of the present experimental study, pedagogical implications and materials. All in all, this study is expected to contribute to the field in which there seems to be a shortage of studies on the explicit teaching of the formulaic language in the field of academic writing, which will be illustrated in detail, in the literature review part, and inspire further research.

1.4. Research Assumptions

The research has its own assumptions made in advance by the researchers. First, it is assumed that in the present study, the participants in both control and experimental groups are homogenous in terms of background EFL knowledge and prior language experience they gained either within the university boundaries or outside it, and along

with other factors having an effect on final achievement scores of the participants. This is thought to be achieved by working with regular classes without making changes on the student profiles of the groups. The researcher also assumes that each of the subjects involved in the lectures will sincerely and actively exert 100 % effort in writing the essays by remembering the lectures they were given and respond to the instructor honestly and accurately in the discussion parts since the experiment is embedded in the course and the essays were used for the assessment of the course. Another assumption, is that the instructors involved in the study are qualified enough and identically, as they have already been assigned to the related courses by the related department of the university based on their qualification and experience.

1.5. Organisation of the Study

The present study is organized into five chapters. Chapter I provides an introduction to the importance of the knowledge of formulaic language, the problem of the deficiency in learners' use of formulaic language and the results of this deficiency. The reasons for the difficulty in learning it are discussed while referring to the significance of formulaic language. It also highlights the necessity for teachers' explicit teaching and presents the questions that lead the research. Chapter 2 offers a review of relevant literature beginning with a discussion of the terminology, the functions and benefits of formulaic language. Then it proceeds to the most recent research and their results related to formulaic language investigations which are categorized as four headings; participants, register, genre and teaching. Chapter 3 presents the methodology for the study beginning with the description of the participants proceeding to the procedure which includes implementation design and data collection and concluding with data analysis. While Chapter 4 reveals the results of the investigation,

Chapter 5 presents a discussion of the results, limitations of the research, conclusion with implications from the present study and offers inspirations for further research.

CHAPTER II

2. LITERATURE REVIEW

Formulaic language has been of wide interest recently to many researchers, who study its basic theory in nature (Durrant, 2008; Ellis et al., 2008; Alali & Schmitt, 2012), define and compare its various aspects and approaches (Sun & Wang, 2003; Biber et al., 2004; Cortes, 2004; Biber & Barbieri,2007; Conklin & Schmitt, 2008; Hyland, 2008b; Durrant & Schmitt, 2009; Simpson-Vlach & Ellis, 2010; Martinez & Schmitt, 2012) and introduce new frameworks for analysis (Conrad & Biber, 2005; Boers et al., 2006; Herbel-Eisenmann et al., 2010; Ädel & Erman, 2012). Practical studies but not as many as theoretical ones, are also found in the literature, which compare different participants, such as native versus non-native speakers (Wray, 1999; Cortes, 2002; Cortes, 2004; Conklin & Schmitt, 2008; Ellis et al., 2008; Durrant & Schmitt, 2009; Chen & Baker, 2010; Siyanova-Chanturia et al., 2011a; Ädel & Erman, 2012) students versus expert writers (Cortes, 2002; Cortes, 2004; Wei & Lei, 2011; Ahmadi et al., 2013), or the characteristics of different registers such as written versus spoken discourse (Biber et al., 2004; Biber & Barbieri, 2007), or else different genres such as student essays versus dissertations and published articles (Hyland, 2008b, Jalali & Ghayoomi, 2010; Wei & Lei, 2011). Beside these, little attention has been paid to the teaching of formulaic language, with the studies proposing lists of formulaic language items in general or specific to a discipline (Simpson-Vlach & Ellis, 2010; Byrd & Coxhead, 2010; Liu, 2011; Martinez & Schmitt, 2012), focusing on the use and

functions of them (Wray, 1999; Wray & Perkins, 2000; Biber et al., 2002; Cortes, 2004; Wood, 2006; Cortes, 2006; Hyland, 2008b; Jablonkai, 2009; Li & Schmitt, 2009; Herbel-Eisenmann et al., 2010; Byrd & Coxhead, 2010; Ädel & Erman, 2012; Staples et al., 2013) and suggesting practical exercises (Cortes, 2002; Wood, 2002; Cortes, 2004; Jones & Haywood, 2004; Cortes, 2006; Wray, 2008; Pang, 2010; Alali & Schmitt, 2012; Peters, 2012; Eriksson, 2012;) so as to define them in a better way and provide a well-framed methodology to make it easier to teach or learn formulaic language items.

2.1. Formulaic Language: Terminology

Recently, research has increasingly focussed on formulaic language which is believed to play a crucial role in language learning by providing processing advantages (Wray & Perkins, 2000; Wray 2000; Jones &, Haywood, 2004; Underwood et al., 2004; Schmitt & Carter, 2004; Jiang & Nekrasova, 2007; Wray, 2008; Bannard & Matthews, 2008; Conklin & Schmitt, 2008; Ellis et al., 2008; Nekrasova, 2009; Arnon & Snider, 2010; Millar, 2011; Siyanova-Chanturia et al., 2011a; Tremblay et al., 2011; Siyanova-Chanturia et al., 2011b; Alali & Schmitt,2012; Valsecchi et al., 2013; Graham & Osment, 2013; Szudarski & Carter, 2014), discourse functions (Cortes, 2002; Schmitt & Carter, 2004; Biber et al., 2004; Biber & Barbieri, 2007; Wray, 2000, Hyland, 2008; Hyland, 2008b; Nekrasova, 2009; Ahmadi et al., 2013; Alhassan & Wood, 2015), fluency in comprehension and production of language (Underwood et al., 2004; Kecskes, 2007; Jiang & Nekrasova, 2007; Bannard & Matthews, 2008; Conklin & Schmitt, 2008; Arnon & Snider, 2010; Tremblay et al., 2011; Siyanova-Chanturia et al., 2011b) as well as idiomaticity (Wray & Perkins 2000; Wray, 2000; Wood, 2002; Wood, 2006; Rott, 2009; Durrant & Schmitt, 2010; Byrd & Coxhead, 2010; Millar,

2011; Laufer & Waldman, 2011; Dickinson, 2012; Henriksen, 2013). Many researchers, who seem to have a general consensus on the features, benefits and accordingly, the importance of formulaic language, have different point of views about its defining characteristics, the methodologies to identify it, or even what to call it (Biber, et al., 2004; Read & Nation, 2004; Conrad & Biber, 2005) since formulaic sequences seem to exist in so many forms (Schmitt & Carter, 2004). Wray and Perkins in their study gathered more than 40 terms (2000, p. 3) which refer to various types or subtypes of formulaic language such as multiword sequences/units, lexical phrases/ bundles/chunks, formulas, clusters, routines, fixed expressions, prefabricated chunks (prefabs) among other terms stated in Table 2.1 below.

Table 2.1
Terms used in the literature to describe formulaic language

1. Amalgams	17. Holistic	32. Ready-made expressions
2. Automatic	18. Holophrases	
3. Chunks	19. Idiomatic	33. Ready-made utterances
4. Clichés	20. Idioms	
5. Co-ordinate constructions	21. Irregular	34. Rote
6. Composites	22. Lexical(ized) phrases	35. Routine formulae
7. Conventionalized forms	23. Lexicalized sentence stems	36. Schemata
8. Fixed expressions including idioms	24. Multiword units	37. Semi-preconstructed phrases that constitute single choices
9. Fixed expressions	25. Non-compositional	
10. Formulaic language	26. Non-computational	
11. Formulaic speech	27. Non-productive	38. Sentence builders
12. Formulas/Formulae	28. Petrification	39. Stable and familiar expressions with specialized subsenses
13. Fossilized forms	29. Praxons	
14. Frozen phrases	30. Preassembled speech	
15. Gambits	31. Prefabricated routines and patterns	40. Synthetic
16. Gestalt		41. Unanalysed chunks of speech

For the sake of convenience, this paper shall adopt an overarching term "formulaic language" encompassing multiword units, chunks, collocations, fixed expressions, prefabricated chunks, idioms, proverbs, and the others stated in Table 2.

1., which may be defined as follows: "a sequence, continuous or discontinuous, of words or other meaning elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar" (Wray, 1999, p. 214). In other words, as Ellis (1996, p. 111) stated that the words in a formulaic sequence are "glued together" and stored as a single "big word" with an associated holistic meaning (as cited in Wray, 2000, p. 465). At the same time, Wray (2009) alleges that even single words and morphemes can function as formulaic sequences.

2.2. Formulaic Language: Functions and Benefits

The researchers agree on many features and benefits of formulaic language (Conrad & Biber, 2005; Dickinson, 2012) although they use different terms. Formulaic language items are so pervasive in a natural language (Schmitt & Carter, 2004; Wood, 2006; Ellis et al., 2008; Durrant & Schmitt, 2009; Durrant & Schmitt, 2010; Valsecchi et al., 2013) that at least one-third to one-half of the language is assumed to be composed of formulaic elements (Erman & Warren, 2000; Foster, 2001; Van Lancker-Sidtis & Rallon, 2004), even Altenberg (1998), suggested that possibly as much as 80% of adult native language may be thought to be formulaic (as cited in Wray, 2000) and these items are found to follow each other more frequently than a random probability in a given register (Erman & Warren, 2000; Wray & Perkins, 2000; Biber & Barbieri, 2007; Hyland, 2008a; Hyland, 2008b; Wei & Lei, 2011; Alipour & Zarea, 2013). This frequency is widely supposed to be a key facilitator for fluent comprehension and production by many researchers (Wray, 1999; Wood, 2002; Wray & Perkins, 2000; Schmitt & Carter, 2004; Conklin & Schmitt, 2008; Durrant, 2008; Wood, 2009; Chen & Schmitt, 2008; Ellis et al., 2008; Hyland, 2008b; Li & Schmitt, 2009; Nekrasova,

2009; Pang, 2010; Wood, 2009; Jalali & Ghayoomi, 2010; Liu & Huo, 2011; Laufer & Waldman, 2011; Tremblay et al., 2011; Siyanova-Chanturia et al. 2011; Conklin & Schmitt, 2012; Boers & Lindstromberg, 2012; Martinez & Schmitt, 2012; Ahmadi et al., 2013; Valsecchi et al., 2013; Ortaçtepe, 2013; Graham & Osment, 2014; Szudarski & Conklin, 2014). The processing of formulaic language, not only by native speakers but also by non-native speakers, are studied by many researchers through eye-tracking method (Underwood et al., 2004; Siyanova-Chanturia et al., 2011a; Valsecchi et al., 2013), self-paced reading tasks (Conklin & Schmitt, 2008; Tremblay et al., 2011; Millar, 2011) and reaction exercises (Jiang & Nekrasova, 2007; Siyanova & Schmitt, 2008; Arnon & Snider, 2010), the analysis of the measurements of speech rate (Bannard & Matthews, 2008) and length of runs (Wood, 2006). These studies proved that the formulaic language items are read more quickly than the same words embedded in non-formulaic sequences since they are stored in and retrieved from long-term memory as if they were single lexical units (Schmitt & Carter, 2004; Wood, 2006; Wood, 2009; Alipour & Zarea, 2013). To illustrate, formulaic language items are available when required for both the speaker and the hearer to reach directly and quickly to the targeted utterance or perception by skipping the generative system partially or entirely (Wray 2000; Wray & Perkins 2000), that is, "once the brain is familiar with a linguistic task, it by-passes the processing route required to generate it" (McCrone, 1999; Raichle, 1998, as cited in Wray, 1999, p. 215) This means less load on working memory during the initial processing and following recall (Jiang & Nekrasova, 2007; Kecskes, 2007; Conklin & Schmitt, 2012). Similar to the process of the speaker, the hearer can understand a message more quickly and easily without having recourse to full analytic decoding as the utterance had been heard before (Wray & Perkins, 2000; Wray, 2000). Appertaining to these processing advantages, both the speaker and the hearer can use

formulaic language items for socio-interactional and pragmatic purposes in a quicker and clearer way (Wray & Perkins, 2000; Biber et al., 2004; Biber & Barbieri, 2007; Hyland, 2008a; Conklin & Schmitt, 2008; Bardovi-Harlig, 2009; Nekrosova, 2009; Liu & Huo, 2011; Boers & Lindstromberg, 2012; Ortaçtepe, 2013). For example, in order to get attention in a crowded and noisy classroom, a teacher could say "Excuse me!" instead of "Could you please give me your attention and be quiet?", because the former is easier and quicker for the speaker to say as it is prefabricated and easier and quicker for the hearer to decode as it is more predictable (Wray & Perkins, 2000; Hyland, 2008a). Although the latter is a grammatically correct sentence and has the same intention as the former, it is not likely to achieve the interactional purpose as efficiently.

Another striking feature of formulaic language items is the hypothesis that the mastery of formulaic language is the signal of competent performance (Cortes, 2004; Van Lancker-Sidtis & Rallon, 2004; Jones & Haywood, 2004; Peters & Pauwels, 2015), key to successful language learning and native-like production (Wray, 1999; Wray, 2000; Durrant, 2008; Hyland, 2008a; Hyland, 2008b; Durrant & Schmitt, 2009). Many researchers suggest that formulaic language helps to shape text meanings (Wray & Perkins 2000; Jones & Haywood, 2004; Hyland, 2008a), develop communicative competence (Wray, 2000; Schmitt & Carter, 2004; Alali & Schmitt, 2012), accomplish native-like idiomaticity (Wray & Perkins 2000; Wray, 2000; Wood, 2002; Wood, 2006; Rott, 2009; Durrant & Schmitt, 2010; Byrd & Coxhead, 2010; Millar, 2011; Laufer & Waldman, 2011; Dickinson, 2012; Henriksen, 2013) and also contributes to our sense of distinctiveness in a register Thus, it helps the participants of a community especially novices or newcomers, to sound like others and to be accepted in the relevant community since many formulaic language items are accepted as a preferred way of

saying something like default expressions (Erman & Warren, 2000, Foster, 2001; Kuiper, 2004; Biber et al., 2004; Schmitt & Carter, 2004; Wood, 2006; Biber & Barbieri, 2007; Conklin & Schmitt, 2008; Ellis et al., 2008; Hyland, 2008a; Hyland, 2008b; Durrant & Schmitt, 2010; Millar, 2011; Liu, 2011; Meunier, 2012; Conklin & Schmitt, 2012). Accordingly, the lack of knowledge in formulaic language may cause misunderstanding or communication breakdown as the example given in the study of Kecskes (2007, p. 1) shows:

"Chinese student:	*– I think Peter drank a bit too much at the party yesterday.*
Turkish student:	*– Eh, **tell me about it**. He always drinks much.*
Chinese student:	*– When we arrived he drank beer. Then Mary brought him some vodka. Later he drank some wine. Oh, too much.*
Turkish student:	*– Why are you telling me this? I was there.*
Chinese student:	*– Yes, but you told me to tell you about it.*"

Moreover, as building blocks of a discourse, formulaic language plays an essential role for discourse organisation and conversation management (Cortes, 2002; Schmitt & Carter, 2004; Biber et al., 2004; Biber & Barbieri, 2007; Wray, 2000, Hyland, 2008a; Hyland, 2008b; Nekrasova, 2009; Ahmadi et al., 2013; Alhassan & Wood, 2015) since it highlights the relationship between "what follows and what has been said previously as well as it signals to the hearer when it is appropriate to speak" (Wray, 1999, p. 216). Although language learners are capable of generating grammatically correct sentences, they may seem odd or unnatural not only in written but also in spoken productions if they do not use formulaic language items to frame their writing and speech studied the functions and the use of formulaic language in a

particular discourse(Wray, 1999; Li & Schmitt, 2009). Thus many researchers (Wray, 2000; Wray & Perkins, 2000; Wood, 2002; Biber et al., 2004; Cortes 2004; Wood, 2006; Biber & Barbieri, 2007; Wood, 2009; Hyland, 2008a; Wood, 2009; Jablonkai, 2009; Nekrasova, 2009; Pang, 2010; Liu & Huo, 2011; Tremblay et al., 2011; Staples et al., 2013; Wood & Appel, 2014). Biber et al. (2004) and Biber and Barbieri (2007) provide insights into how formulaic language is functionally divided, by proposing a comprehensive taxonomy which is widely accepted as stance expressions, discourse organizers, and referential expressions. However, it is likely to see that several sub-categories are constructed for each main category as well in the literature (Cortes, 2002; Biber et al., 2004; Hyland, 2008b; Chen & Baker, 2010). A category of special conversational (Cortes, 2002; Conrad & Biber, 2005) or interactional expressions is sometimes also used if spoken data is considered (Biber et al., 1999, as cited in Ädel & Erman, 2012, p. 89). In broad terms, referential expressions are specific attributes to temporal, spatial, physical context (Wood & Appel, 2014) such as "at the beginning of", "the end of", or "at that point", which helps the language users form their ideas. Discourse organisers provide interpretive frames to develop discourse (Wood & Appel, 2014) and mark the relationships among parts of discourse like introducing a topic, detailing an idea or ending it with expressions such as "to begin with", "on the contrary" or "in conclusion". Expressions of stance and interactional expressions are the ones used for interpersonal functions (Cortes, 2004). While the interactional or special conversational expressions are colloquial word combinations used for expressing inquiry like "what are you doing?" and politeness like "thank you", and also reporting clauses like "He said to her" (Conrad & Biber, 2005), expressions of stance convey attitudes, assessments of certainty or modality that frame a proposition (Biber &

Barbieri, 2007; Wood & Appel, 2014; Conrad & Biber, 2005) like "I don't know what …" or "less likely to" (Cortes, 2004).

Lastly, some other features of formulaic language items include that they can be semantically transparent (*e.g. as a result of*) although they do not need to be as in idioms (Wray, 2000; Cortes, 2002; Cortes, 2004; Schmitt & Carter, 2004; Biber & Barbieri, 2007; Hyland, 2008a; Hyland, 2008b; Wei & Lei, 2011; Peters, 2012; Ahmadi et al., 2013), in most cases their meanings can be understood from their components. Besides, they can be syntactically regular (Wray, 2000) which means that they can be fully generated by the language grammar as *it in the example of it was lovely to see you*, and also usually incomplete in structure such as *in the case of* (Cortes, 2004; Biber & Barbieri, 2007; Hyland, 2008a; Nekrasova, 2009; Pang, 2010; Wei & Lei, 2011; Ahmadi et al., 2013). Thus, they are often relatively lacking in salience in the input (Biber & Barbieri, 2007; Durrant & Schmitt, 2010), that is to say that those formulaic language items may not be noticeable or prominent in the input.

2.3. Formulaic Language: Challenges and Reasons

Taking account of the importance of formulaic language explained above and the fact that formulaic language elements such as idioms, collocations, and lexical bundles are presumed to form a high percentage of natural language (Erman & Warren, 2000; Wray & Perkins, 2000; Foster, 2001; Van Lancker-Sidtis & Rallon, 2004; Kuiper, 2004; Schmitt & Carter, 2004; Conrad & Biber , 2005; Liu, 2011; Valsecchi et al., 2013) mastering them seems to be a prerequisite for native-like mastery of a foreign language (Peters & Pauwels, 2015). However, learning formulaic language items can be hampered due to some reasons originated by the learner, the teacher, the material or

the formulaic language itself (Jones & Haywood, 2004; Bardovi-Harlig, 2009; Liu, 2011; Laufer & Waldman, 2011; Martinez & Schmitt, 2012). For instance, learners may not have enough experience with the language in classroom due to sparse input and inadequate instruction (Foster, 2001; Li & Schmitt, 2009; Byrd & Coxhead, 2010; Durrant & Schmitt, 2010; Szudarski & Carter, 2014;) and also in real world due to little exposure to socialization with native speakers (Ortaçtepe, 2013). Thus, they are likely to have a lack of awareness of the input that surrounds them (Wray, 2000; Bishop, 2004a; Bishop, 2004b). As a result of this, they may ignore new formulaic language items (Bardovi-Harlig, 2009) and overuse some expressions previously learned while worrying about being understood by the interlocutors (DeCock, 2000; Kecskes, 2007).

Moreover, unlike native speakers, who process formulaic language as holistic units, Wray (2002) argued that learners may tend to process such language elements word by word making a conscious effort to learn formulaic sequences (as cited in Valsecchi et al., 2013) and they are likely to result in first language transfer (DeCock, 2000; Altenberg & Granger, 2001; Nesselhauf, 2003; Ellis, 2012). As for the teacher, thinking that students are already exposed to the formulaic language through materials, s/he may not provide enough and efficient input in a meaningful context or relate well to the language elements with real world situations (Wray, 2000; Lin, 2013; Wood & Appel, 2014). However, the materials appear to have a lack of presenting formulaic language in a pedagogical way. Moreover, these materials which consist of just lists constituted without pedagogical approach and provide decontextualized sentences with a poor or no methodology (Jones & Haywood, 2004; Wood & Appel, 2014) may fall short of helping students to learn formulaic language. The linguistic structure of formulaic language itself may also create confusion for learners, and even for teachers,

since grammatically generated word strings can appear similar to formulaic sequences (Wray, 2000; Bishop, 2004a; Bishop, 2004b, Durrant, 2008; Laufer & Waldman, 2011), and that the formulaic sequences are ubiquitous and may have transparency mostly may cause learners simply to pay less or no attention to formulaic language elements (Erman & Warren, 2000; Bishop, 2004a; Bishop, 2004b; Wood, 2006; Biber & Barbieri, 2007; Wood, 2009; Laufer & Waldman, 2011).

Unfortunately, due to reasons mentioned above, even advanced learners, who have extensive knowledge of grammar and lexicon, may experience challenges (Chan & Liou, 2005; Wray & Fitzpatrick, 2008; Nekrasova, 2009; Byrd & Coxhead, 2010; Čolović-Marković, 2012) such as little/under use (Li & Schmitt, 2009; Bardovi-Harlig, 2009; Laufer & Waldman 2011; Lin, 2013) or overuse of formulaic language (Cortes, 2002; Cortes, 2004; Durrant & Schmitt, 2009; Qi & Ding, 2011; Ädel & Erman, 2012; Valsecchi et al., 2013), collocational errors (Altenberg & Granger 2001; Nesselhauf, 2003; Rott, 2009), inefficient comprehension and production of language (Henriksen, 2013), inarticulate speech (Hyland, 2008a), incompatible use of language in a discourse (Jones & Haywood, 2004; Wood, 2009; Pang, 2010) and so forth if their utterances lack in formulaicity which is considered a touchstone of second language acquisition (Jones & Haywood, 2004). Such challenges negatively influence the impression of learners' language performance (Szudarski & Carter, 2014) and make the learners sound non-native, unnatural or odd and not belong to related discourse community (Ellis et al., 2008; Li & Schmitt, 2009, Ortaçtepe, 2013; Peters & Pauwels, 2015).

2.4. Experimental Studies on Formulaic Language

Many researchers have conducted experimental studies on formulaic language due to its great importance in language learning. These studies can be divided into four categories regarding their purpose; participants, register, genre and teaching.

2.4.1. Studies comparing the use of formulaic language of different participants

Much of the research on formulaic language depends on the comparison of the participants involved in the research such as native versus non-natives or students versus expert writers. For example, Ping (2009), DeCock (2000) and Durrant and Schmitt (2009) investigated student writing by native speakers and non-native speakers. The results indicated that non-native speakers used as many formulaic language items as natives did, sometimes more, with fewer varied functions compared to native speakers and misuse, underuse or overuse certain types of formulaic language items, e.g. verb phrases (Ping, 2009), two to four word items (DeCock, 2000) and certain favoured formulaic language items generally having strong collocations (Durrant & Schmitt, 2009). In a similar way, some researchers like Alipour and Zarea (2013) and Ahmadi et al. (2013) focused on published articles by both native authors and non-native authors and the results revealed that non-native authors generally made a greater use of formulaic language items in terms of the frequencies, but overused certain type of items like four word items or clausal elements compared to their native counterparts. The reason suggested by DeCock (2000) and Durrant and Schmitt (2009) for non-native speakers' more frequent use of formulaic language items is that they prefer to use high frequent formulaic language items which have L1 equivalents or it may be because they

have already been exposed to those formulaic language items several times in their prior readings according to Ahmadi et al. (2013). Although many researchers seem to have a consensus on that non-native speakers use more but fewer types of formulaic language items, different results exist in the literature as well. For instance, Römer (2009) found very few differences between the use of four word formulaic language items by native speakers and advanced non-native speakers while Ädel and Erman (2012), Ortaçtepe (2013) and DeCock (2004) who removed the repeats and hesitation items from the spoken data, found that native speakers produced more formulaic language items than non-native speakers. As expected, the native speakers in a study by Ortaçtepe (2013) received higher native-like ratings, on the other hand, the ratings of non-natives were observed to improve through exposure to a target language community. Moreover, Chen and Baker (2010) investigated student writing by native speakers and non-native speakers as well as journal articles written by native speakers. Differences regarding structure and function of formulaic language items were found between the journal articles and both groups of student writing, but not between the native and non-native student writing. For instance, journal articles written by professional writers contained more NP-based formulaic language items and referential expressions while both native and non-native student writing had more VP-based formulaic language items and discourse organisers. Learners were also observed to often use a limited variety of formulaic language items, and at times overuse certain expressions rarely seen in the corpus of professional academic writers.

Other studies basically focused on comparisons between expert and non-expert writing. For instance, Cortes (2002; 2004) compared different level university students to professional writers regarding their use of formulaic language in published articles

and student writing in the disciplines of history and biology. The results of this study showed that students seldom used formulaic language items in their writings compared to the ones in professional writers' published articles although they are all native speakers of English. Another study by Wei and Lei (2011) comparing the doctoral dissertations by advanced Chinese EFL writers and published research articles by professional writers, showed that advanced learners generally used formulaic language items more frequently and with greater range. However, the items used by the students in the studies (Cortes, 2002; 2004; Wei &Lei, 2011) did not correspond to the ones selected by the professional writers.

On the other hand, some researchers compared not only the production of formulaic language by native and non-native speakers but also the processing of it by eye-tracking method (Underwood et al., 2004; Siyanova-Chanturia, et al., 2011a; Siyanova-Chanturia et al., 2011b; Valsecchi et al., 2013), self-paced reading tasks (Conklin & Schmitt, 2008; Ellis et al., 2008; Millar, 2011; Tremblay et al., 2011) or analysing the reaction times (Jiang & Nekrasova, 2007; Siyanova & Schmitt, 2008; Arnon & Snider, 2009; Szudarski & Conklin, 2014). The results of these studies proved that formulaic language is stored holistically and this provides processing advantages since both native and non-native speakers responded to formulaic sequences faster and with fewer errors than control sequences which are non-formulaic while native speakers had a greater processing advantage over non-native speakers who were found to have poorer intuitions as well.

2.4.2. Studies comparing formulaic language in various registers

In the literature relating to formulaic language, there is a tendency to examine various registers which are the instantiations of a conventionalised, functional configuration of language tied to certain broad societal situations (Lee, 2001, p. 46) such as legal register, formal register, written register and spoken register since formulaic language often has discernible functions within a particular register as basic building blocks of discourse (Biber et al., 2004; Conrad & Biber, 2005; Wei & Lei, 2011; Alipour & Zarea, 2013). For example, Biber et al. (2004), comparing the use of formulaic language items in spoken register (classroom teaching and conversations) to those in written register (textbooks and academic prose), found that there are differences among these registers in the use of formulaic language items. The results showed that the formulaic language items used in classroom teaching were surprisingly different in both frequency and functions from those in conversation and academic prose, by having more stance and discourse organizing expressions. While spoken registers (classroom teaching and conversation) used a much greater range of different formulaic language items than written registers (textbook and academic prose), classroom teaching used the most formulaic language items in each functional category.

Another noteworthy study by Biber and Barbieri (2007) who extended the previous study of Biber et al. (2004) by investigating the use of formulaic language items in a wide range of spoken university register (classroom teaching, class management, office hours, study groups, service encounters) and written university registers (textbooks, academic prose, course management, institutional writing). The findings of this study indicated that each category of both registers had distinctive formulaic language items regarding both the frequency and the function. For instance,

classroom management and service encounters had the widest variety of bundles within the spoken register and institutional writing within the written register used as many formulaic language items as any spoken register while written course management had more formulaic items than any other sub-categories in both registers. As for the functions of the formulaic items, while spoken register employed mostly stance expressions, written register had mostly referential expressions.

Additional to aforementioned studies, Chen (2008) and Wood and Appel (2014), who compared the EAP/ESP textbooks to discipline-specific textbooks, found that EAP/ESP textbooks were insufficient to present the formulaic language items that were used abundantly in discipline specific texts, also the functions of the formulaic language items were different in these registers. Similarly, Cortes (2004), Hyland (2008b) and Alipour and Zarea (2013) who compared the texts written in different disciplines such as biology and history (Cortes, 2002; 2004), electrical engineering and microbiology from the applied and natural sciences, and business studies and applied linguistics from the social sciences (Hyland, 2008b), computer engineering, physics and applied linguistics (Alipour & Zarea, 2013), found that the formulaic language items differed in forms, frequency and functions in these fields. For instance, while research articles in biology employed much more formulaic language items than those in history (Cortes, 2002; 2004), electrical engineering texts contained the greatest range of high-frequency formulaic language items (Hyland, 2008b) and computer engineering was found to include formulaic language items more frequently than physics and applied linguistics (Alipour & Zarea, 2013). According to the researchers, one of the reasons of these differences might be that the unique distinctiveness of each discipline caused considerable deviations in the frequency, structure and functions of formulaic

language items used. For instance, the fields in natural sciences such as computer engineering and physics which depend on more on numerical and visual data requires utilising more formulaic language items to interpret their data while applied linguistics depend on text-oriented formulaic language items to discursively elaborate and frame the research. In addition, Cortes (2004) and Hyland (2008b) noted that student and published writings had specific discrepancies by discipline as well.

2.4.3. Studies comparing formulaic languages across genres

The feature of the variability across different genres of the formulaic language is another popular topic among the researchers as it is thought that each genre is characterized by peculiar formulaic language items (Hyland, 2008a; Ellis et al., 2008). For example, Hyland (2008a) investigated the use of formulaic language items in three academic genres, published journal articles, PhD, and MA dissertations of EFL students from four different disciplines, electrical engineering, business studies, applied linguistics and microbiology. He believes formulaic language items have distinctive characteristics across academic genres and to know those items which typify their disciplines and those which are respected in the particular genres makes the learners more acceptable in that particular discipline and community. The findings of the study in which the frequencies and patterns were compared across the different corpora, supported the theory that the use of formulaic language varies considerably in frequency, structure and function across different genres. For instance, the journal articles employed far fewer formulaic language items with a narrower range of different items overall compared to the Master and PhD dissertations. Hyland (2008a) argued that differences could be due to less experienced writers being more dependent on formulaic language items than experienced writers. On the other hand, the formulaic

language items employed in the journal articles, never or seldom appear in the dissertations of master and PhD students. Hyland (2008a) assumed that the reason for this difference might be the different topics addressed by writers in the three genres. As for the structural differences, in all three genres, formulaic language items including prepositional phrases were used at a similar frequency. Lastly, for the functional differences, master's theses mainly included research-oriented items of formulaic language the reasons of which were assumed to be the preference of author anonymity and the impersonality feature of expository writing reinforced in the instruction while the items in the journal articles and PhD dissertations were basically participant and text oriented which was thought to be due to more sophisticated and more academic approach to engage their readers.

There are some other studies focussing on genres such as that of Wei and Lei (2011) and Eriksson (2012) which had similar results to that of Hyland (2008a). Eriksson (2012) compared the three-to-five-word formulaic language items in the writings of PhD students and the published articles of professional writers in the field of biochemistry and biotechnology. He found that there were more formulaic language items in the PhD corpus than in Biotech corpus of published articles. On the other hand, Wei and Lei (2011) compared most frequent 50 four-word formulaic language items in journal articles by professional writers and doctoral dissertations of Chinese EFL learners. The results demonstrated that greater number of formulaic language items with a wider variety was employed in doctorate dissertations. Although the formulaic language items used in both genres were basically characterised with research and text oriented formulaic language items, they were more varied and more passive structures in doctoral dissertations. Wei and Lei (2011) similar to the argument of Hyland (2008a)

stated that the reason of passive structures might be the writers' preference for the impersonality.

The qualitative study of Jalali and Ghayoomi (2010) explored the use of formulaic language items which were identified in research articles, in master's and doctoral theses as different genres of the same field-applied linguistics. The results indicated that formulaic language items used frequently in the research articles were employed considerably frequently in master's and doctoral theses, which runs counter to the findings of Hyland (2008a). Although in each genre some formulaic language items usage was dominant, considering the overall use of formulaic language, very little difference was found between the three genres. Jalali and Ghayoomi (2010) argued that frequent use of formulaic language items by the students could be due to previous exposure to these items many times in the literature during their study.

2.4.4. Studies concerned with the explicit teaching of formulaic language

Although considerable research has been carried out into the formulaic language in various discourses, there seems to be a dearth of studies on the effects of the explicit teaching of formulaic language within a pedagogical dimension. Most of the studies only focused on activities or techniques such as noticing-awareness raising (Boers et al., 2006), typographic salience (Bishop, 2004a; Peters, 2012), memorization (Wray, 2004; Wray & Fitzpatrick, 2008), rote rehearsal (Szudarski & Conklin, 2014), repetition (Alali & Schmitt, 2012; Webb et al., 2013; Peters, 2014), glossed sentence and cloze tasks (Webb & Kagimoto, 2009), concordance and corpus instruction (Sun & Wang, 2003; Chan & Liou, 2005) and contrastive analysis and translation (Laufer & Girsai, 2008), some of which have been utilized in the present study. On the other hand,

there are a few intervention studies with mixed results, on a targeted aim like fluent speaking (Wood, 2009; Dickinson, 2012) or academic writing (Jones & Haywood, 2004; Cortes, 2006; Peters & Pauwels, 2015), focussing on teaching formulaic language as a whole; that is, selecting the formulaic language items, teaching them with a variety of activities and techniques, having students practice and produce them and finally evaluating and giving feedback. However, those studies were case studies in general and the participants involved were so few that the results which may shed light on teaching of formulaic language, were suggested to be considered tentatively by the writers.

One of the pioneering studies to explore whether teaching of formulaic language can cause any improvement on the proficiency of the students was carried out by Jones and Haywood (2004) who used a variety of standard awareness raising exercises like highlighting of identified target formulaic language items in reading texts; deeper processing exercises like classifying them according to meaning or structure, discussing their usefulness for EAP writing, using concordance lines; lastly gap-fill exercises, writing assignments and discussion sessions for the experimental group of 10 students during ten teaching weeks. From lesson observations, interviews, and comparison of pre and post-treatment data Jones and Haywood (2004) found that they were largely successful in raising students' awareness of formulaic language, but this awareness did not lead to any general increase in the use of the items in students' later output. The researchers also noted that the formulaic language can be utilised to express technically complex ideas in an economical way not only to mark the stages in their discourse but also show the necessary level of formality.

In a study similar to that of Jones and Haywood (2004), Cortes (2006) adopted similar techniques with some refinements such as contextualized examples from corpus, paraphrasing activities and discussion sessions, with the participation of eight native speaker students who were taught formulaic language via five 20-minutes micro lessons in an intensive history writing class. Analysing the written assignments handed before, during and after the treatment, Cortes (2006) concluded that treatment raised the awareness of the students toward the use, frequency and function of formulaic language items in published articles. However awareness did not turn into success in the written production of the students, which corroborates the findings of the study by Jones and Haywood (2004).

Taken the study of Jones and Haywood (2004) as a basis and considering its limitations, Čolović-Marković (2012) designed her study by including more participants, extending the treatment duration and diversifying the activities. The results of her study indicated that the performance of the treatment group, in controlled situations (e.g. C-tests), was significantly higher than the group not receiving explicit instruction. However in uncontrolled situations, namely essays written as a sign of overall quality, the results were in line with the results of the study by Jones and Haywood (2004). Besides quantitative analysis, Čolović-Marković (2012) carried out qualitative analysis through individual interviews with some participants who stated that their focus was on expressing their ideas on the topic given at the beginning of the semester before the explicit instruction. However, when they were given the same topic at the end of the semester, which was after the treatment, their focus shifted to formulaic language more and they were observed to use formulaic language for not only what to say but also how to say it. The participants also expressed that to be given a list as a

reference would be much more helpful and they employed mostly memorization, highlighting, and conscious focus while reading to learn the formulaic language items during their self-study.

Two other studies focussing on the explicit teaching of formulaic language items were carried out by Peters and Pauwels (2015) and Alhassan and Wood (2015). For three weeks with 29 participants in an EFL class, Peters and Pauwels (2015) focussed on the teaching of certain formulaic language items by some activities that they categorised as recognition (underlining), cued output (fill in the gap, rephrase, use in a sentence) and recognition+cued output activities. They gathered the data from recognition, cued output and writing tests before and after the treatment and end-of-year assignments as well. Alhassan and Wood (2015) carried out their research over ten weeks with the participation of twelve mixed level students by using not only contextualized but also decontextualized activities and gathered the data from the pre-test, the post-test and the delayed post-test. On the contrary to the prior studies, both studies demonstrated that the explicit teaching of the targeted items was effective since the students receiving explicit instruction presented higher success and used a wider range of formulaic language items in the post-test and the delayed post-test. Moreover, Peters and Pauwels (2015) found that cued output activities caused more learning gains than recognition activities and Alhassan and Wood (2015) found that students could internalize the target items as there was no significant difference between the post-test and the delayed-post-test.

In addition to studies focussing on writing, some studies (Wood, 2009; Dickinson, 2012) explored to what extent the explicit teaching of formulaic language items has an effect on the speaking fluency of non-native students. Dickinson (2012)

who compiled a list of formulaic language items from previous research studies, presentation course books, and videos of relevant presentations, taught 10 Japanese students from a mixed proficiency level from false beginner to quite advanced, explicitly how to use these items, encouraged memorization and asked to complete video- and paper-based activities for two weeks. Wood (2009) asked an intermediate level Japanese student to produce narratives spontaneously in the university language laboratory on topics of personal relevance, with no preparation time or use of notes to prepare for the talk during a six-week workshop which included input stage, automatization stage and practice-production stage. Quantitative and qualitative analysis of both research showed that there was an increase in the use of formulaic language items regardless of their proficiency and the overall quality of the presentation/speech such as faster speech and longer runs between hesitations.

As can be understood from the literature, formulaic language has a very significant role in second language learning for learners to reach a native-like proficiency with the help of facilitative and processing advantages as well as discourse functions not only in production but also in comprehension. Although this importance is well-known, formulaic language is generally meant to be taught implicitly through exposure; however, literature has shown that there are some challenges, caused by certain factors such as materials, teachers, learners or formulaic language itself, which inhibit the learners from proficiently using formulaic language. Moreover, implicit teaching of formulaic language might take longer than limited classroom time allows, since it is proposed that even single words have to be encountered not fewer than eight times for the meaning to be learned (Waring, Takaki, 2003; Webb& Kagimoto, 2009; Webb et al, 2013). Thus, a considerable number of studies were carried out so as to

indicate to what extent the explicit teaching of formulaic language is affective, through some activities and techniques with the involvement of a few participants; however, many of them simply focussed on awareness raising, processing of it or the frequency of use rather than the overall quality that formulaic language provides to writing. Thus, this study aims to demonstrate the effectiveness of the explicit teaching of formulaic language to the overall argumentative quality of the academic writing of the university students, with the involvement of more participants than the previous studies, within a well-framed methodology rather than random activities and techniques which are embedded in an existing curriculum.

CHAPTER III

3. METHODOLOGY

3.1. Participants

This classroom based research took place in the Department of English Language Teaching at Yeditepe University in Istanbul, in Turkey. Two freshman classes consisting of 52 students and two sophomore classes consisting of 33 students, in total 85 students aged ranged between 18 and 21 were involved in the present research. Valid data could be gathered from only 71 students, though, since not all of the students appeared at all of the tests, which required writing an argumentative essay on a given topic. The experimental group was composed of one already established freshman class (35 students) and one sophomore class (18 students), in total 53 students, while the control group was another freshman class (17 students) and sophomore class (15 students), totalling 32 students. The students either took one year preparation class or were exempt from it by scoring 79 on the TOEFL, 6.5 on the IELTS or 60 on the

Yeditepe University Exemption Exam. Their proficiency in English, therefore, is estimated to be of upper-intermediate to advanced level. At the time of this research being carried out, freshmen were taking the course "English Composition-I" and sophomores "Academic Reading & Writing" courses. Moreover, the ones who took the preparation class also took an academic writing course in preparation class. In the research, there were three instructors different from the researcher; two of whom were teaching the classes in the control group based on the regular curriculum while the third one, who had been the advisor of this study, was teaching both classes in the experimental group.

Table 3.1
Details of the participants of the study

	Participants involved in the study		Participants from whom the data could be gathered	
	Experimental group	**Control group**	**Experimental group**	**Control group**
Freshmen	35 students	17 students	30 students	14 students
Sophomores	18 students	15 students	14 students	13 students
Total	53 students	32 students	44 students	27 students

3.2. Procedure

3.2.1. Selecting the formulaic language items

Prior to the teaching period, a reference list of formulaic language including 100 items (Appendix A) was available as a result of a lecture-called "Corpus" given to M.A. level students, by the advisor of this study, who has great experience in academic writing and corpus studies. The items included in the list was built up by reviewing the books prepared for academic writing such as Kaplan's TOEFL IBT 2009 4th Edition; Barron's How to prepare for the TOEFL, 11th edition; Sharpe, 2004; TOEFL IBT tips by ETS; Teaching academic ESL writing by Eli Hinkel 2004; English Grammar for the

Utterly confused, 2003; Better writing right now, Galko, 2001 and online teaching materials geared towards preparing students for TOEFL, IELTS, GRE, and academic writing, available on the internet, e.g. https://www.oxford-royale.co.uk/articles/words-phrases-good-essays.html, http://www.smart-words.org/linking-words/transition-words.html, http://www.dcielts.com/ielts-essays/vocabulary-a-list/, http://www.tesoltasks.com/ArgVocab.htm, http://www.slideshare.net/englishbites/useful-argumentative-essay-words-and-phrases

Due to insufficient time to teach all of the items included in the reference list, the items was then reduced to 40 which was thought to be those mostly used in argumentative writing. The 40-item-formulaic language list was composed of the items in the available reference list, intuitively, based on the requirements of the writing courses (English Composition-I and Academic Reading & Writing) conducted during the present study, while keeping in mind the usefulness and the relevance to the specific discourse functions intended to be taught. The frequency of these items were also checked through COCA and more frequent items were chosen.

Table 3.2
40-item-formulaic language list

Additional Support	In addition, In the same way,	Equally important,
Putting the same idea in a different way	In other words, To put it simply,	That is to say
Opposing words	By contrast, On the other hand,	On the contrary,
Giving examples	For example, For instance,	To illustrate, such as
Enumeration	First, Second,	Third, Finally,
Consequential words	As a result, Thus,	For this reason, In effect,
Certainty words	Without doubt, Undoubtedly,	Needless to say,

Comparison words	Nevertheless,	Nonetheless,
Introducing opposing ideas	It is often argued that…	Opposing views claim that…
Supporting opposing ideas or partly accepting to find a common ground	One cannot deny that… It could be argued that…	It is true that…
Refutation of opposing ideas	However this conclusion is not well supported, Nevertheless, this conclusion is flawed.	On second thoughts,
Negative words	unfounded questionable	oversimplified
Conclusion	In conclusion, To conclude,	In closing,

Besides, students were given an organizational template for argumentative writing and three thesis statement structures, which were constituted by reviewing the aforementioned books prepared for academic writing, to help them write more organized and cohesive essays. The students were assumed to be familiar with the rhetorical structure of argumentative writing and theses statement writing, as they attended academic writing classes prior to this study.

Table 3.3
Organizational template for argumentative writing

1st paragraph	Introduction Thesis statement Opposing view Topic Position Reasons
2nd paragraph	Opposing view Finding a common ground Refutation
3rd paragraph	My point Supporting evidence and examples
4th paragraph	Conclusion Thesis statement rephrased and restated. Connection is made to real life and broader world.

Table 3.4
Thesis Statement Structures for argumentative writing

Structure 1	Although many people claim that [opposing view], closer examination shows that [my view]
Structure 2	While [opposing view], [my view].
Structure 3	[my view] because [my reasons]

3.2.2. Preparing teaching activities and tests

After having decided what formulaic language items to be focused on, accordingly, some activities were prepared to have the participants practice formulaic language items in activities which were tried in the studies available in the literature and demonstrated to have positive effect on learning formulaic language; these are giving a reference list of target formulaic language items (Čolović-Marković, 2012) (See Appendix A and B), highlighting and using bold letters (Bishop, 2004a; Peters 2012) to make the student notice formulaic language; translation exercises (Laufer & Girsai, 2008) (Appendix E), fill in the blanks exercises (Jones & Haywood, 2004) and cloze tasks (Webb & Kagimoto, 2009) (See Appendix C) and discussions to make the students process deeply and practice formulaic language.

Finally, different topics for timed-argumentative essays were chosen (See Appendix D) for the participants' pre-test, post-test, final test and delayed post-test, in discussions with the instructors of the courses, based on the rationale that the students may have sufficient background knowledge about them, and also a quiz (See Appendix E) was designed for the experimental group. For each test, multiple topics (See Appendix D) were selected so that students can choose the one about which they know more, thus, they were aimed to feel more comfortable and less stressed during the 50 minute essay writing exam.

3.2.3. Instruction and data collection

As for the data collection, initially, the participants in both group were given a pre-test in which they were asked to write a timed-argumentative essay by choosing one of the topics given as alternatives (See Appendix D). The students in both groups were not informed about the pre-test in advance with the aim of diagnosing their current level in argumentative writing.

Following the pre-test, only the experimental group was given the full list of formulaic language items as a reference (See Appendix A) and the 40-item-formulaic language list (See Appendix B), which were instructed through explicit teaching. The instruction basically consisted of a presentation stage followed by a practice stage using aforementioned activities for four hours in total over two consecutive weeks and ended with a production stage in which the students were given a quiz (See Appendix E) including cued questions which push them to use the target formulaic language items.

The explicit instruction started with the presentation of the target formulaic language items through lists (See Appendix A and B) in which the targeted items were presented, then the importance, functions and features of them were explained to draw attention of the students as awareness-raising activities. In the practice stage, the participants were asked to do some activities such as fill in the blanks exercises, translation exercises and cloze exercises (See Appendix E and C). To provide a better understanding, the answers of the exercises were checked and discussed in the classrooms. Additionally, for their individual study, the students were encouraged to memorize the items which is a suggested technique by Wray (2004) and Wray and Fitzpatrick (2008) and also repeat the items orally and in writing as much as possible,

as suggested in the literature by Webb et al. (2013) and Alali and Schmitt (2012). Moreover, during the explicit instruction period the experimental group received a quiz the answers of which were later discussed and analysed in the classroom with the purpose of deep-processing as mentioned in the study of Jones and Haywood (2004). After the instruction was completed, immediately a post-test was carried out, in which the students were asked to write an argumentative essay by choosing one of the topics given. They were given the same topics of the pre-test, but asked to choose the one that they had not previously chosen for the pre-test, with the purpose to compare the achievements before and after the treatment.

While the experimental group was being instructed explicitly, the control group was doing some presentations related to academic writing and essay writing exercises, none of which were specifically related to the target formulaic language items. In other words, they followed the regular course syllabus. To compare the achievements of both groups, the control group was also assigned to write an argumentative essay for the post-test.

Following the post-test, two weeks later, both the experimental group and the control, took the final test as a requirement of the curriculum. The students had three alternative topics to write an argumentative essay for the final test (See Appendix D). Then, the experimental group students were given an assignment (See Appendix C), prepared by cloze test techniques for a consolidation. In order to track any permanent learning a delayed post-test was given to the students approximately one month later.
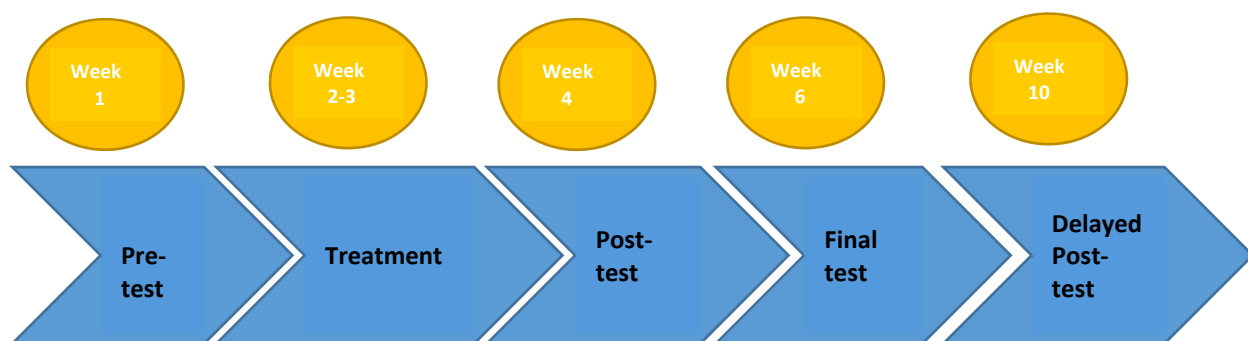
Figure 3.1 The process of the instruction and data collection

Table 3.5
Activities and tests carried out during the research

	Experimental group	**Control group**
Pre-test	✓	✓
The formulaic language list as a reference	✓	✗
40-item-formulaic language list as a study list	✓	✗
Explicit instruction	✓	✗
Noticing activities such as highlighting and using bold letters, with 40-item-formulaic language list	✓	✗
Translation activities	✓	✗
Fill in the blanks exercise	✓	✗
Quiz and discussion of the answers	✓	✗
Post-test	✓	✓
Final test	✓	✓
Cloze test exercise (assignment)	✓	✗
Delayed post-test	✓	✓

3.2.4. Raters and Scoring

Three raters participated in the study since many scholars recommend that multiple raters be used for a better interpretation of the results (Penny, et al., 2000; Graham et al. 2012), even refer to three raters as optimal (Bogartz, 2010). The raters were composed of a native speaker who has a master degree and great experience with ESL learners both in high school level and university level in different countries, especially in academic writing. The other two raters were Turkish, one of them is the writer of this study and the other rater is the advisor of this study with invaluable experience in teaching academic writing at almost all levels in ELT departments in different countries.

Each essay was evaluated holistically for the overall argumentative quality by these three raters separately, being scored between 1 and 4 via a rubric (Appendix F) which was inspired by the TOEFL Writing Scoring Guide in accordance with the aim of the research. The inter-rater reliability of the raters with the intra-class correlation coefficient was found .65 for the pre-test, .75 for the post-test, .63 for the final test and .85 for the delayed post-test, with a 95% confidence interval. Then the scores of the raters were averaged as there is a hot debate in the literature about the standards of reliability. For instance, many scholars (Rothbard & Edwards, 2003; McAllister & Bigley, 2002) state that the acceptable reliability level is .70 by referring to Jum C. Nunnally; however, Lance et al. (2006) calls it as "urban legend" by giving a quotation of Nunnaly "the satisfactory level of reliability depends on how a measure is being used". Moreover, Schmitt (1996), Kehoe (1995) and Trevethan (2016) state that .50 of alpha level can be acceptable depending on the test and interpretation and it does not seriously damage the validity. Another reason for preferring averaged scores

of the raters is that it is one of the mostly advised methods (Bogartz, 2010; Johnson, et al., 2000; Stemler & Tsai, 2008) and for TOEFL writing section the scores of two human raters and an e-rater are also averaged (https://www.ets.org/s/toefl/flash/33910_toefl-scoring-video_transcript.html , https://magoosh.com/toefl/2014/how-is-the-toefl-scored/).

The appropriate use of each target formulaic language item was evaluated independently from the overall argumentative essay. Each target formulaic language item used in each essay was automatically tagged by a computer programme developed by the advisor of this study. With the help of an edit-distance algorithm, utterances close to the target items but not exactly the same were also identified and tagged. For instance, as illustrated below, the target formulaic language item "*in other words*" was used appropriately by the students and the programme tagged it as correct. However in the second example, the target formulaic language item "*in conclusion*" was used inappropriately by the student as "*to conclusion*", thus the programme tagged it with two character differences.

[In other words,]**<In other words, :0>** any parents can go to the parks and amusement parks with their children for pleasure .

[To conclusion,]<**In conclusion, :2**> big cities have every facilities to bring up a child such as education , health facilities and activities .

Then manual checking was done to correct the wrong and missing tags. Some items were mistagged because the surface form was synonymous with the target formulaic language item as in the below example, in which the student used the target

formulaic language item like an adjective instead of an adverb and it was rejected by adding "r" to the tag.

To begin with the [first] <**first: 0r**> reason why native speakers of English are best teachers is that they know the whole functions and features of their native language, …

For each correct use of formulaic language items, a score of 1 was awarded and if a deviation from the original target formulaic language item was observed, a score of 0.5 was given. Each paper was assigned type and token frequency and percentage values based on the tags as shown below. To illustrate, the first number stands for the type of the formulaic language items used after the student's definition part highlighted with bold characters and the second number following it stands for the token of the formulaic language items, that is to say that StudentX of experimental group in the pre-test used only one target formulaic language item which was "for this reason" and he used it twice in his/her writing; StudentY of control group in the post-test used three different target formulaic language items (first, on the other hand, finally) and tokens for them in total four.

pretest_experimentalGroup_StudentX, 1, 2, for this reason,
posttest_controlGroup_StudentY, 3, 4, first, on the other hand, finally
delayedposttest_experimentalGroup_StudentZ, 4, 4, such as, on the other hand, for instance, to conclude

Lastly, each essay was analysed manually by the researcher to identify the thesis statement structures and organisational template that the students used. Thesis statement structures and organisational template used by the students were coded so

that they can be analysed statistically later. Accordingly six different uses for the thesis statement and seven different uses for the organisational template were defined as shown below.

Table 3.6
Categorisation of thesis statement use of the participants

Category 1	Although many people claim that **[opposing view]**, closer examination shows that **[my view].**
Category 2	While **[opposing view]**, **[my view]**.
Category 3	**[my view]** because **[my reasons].**
Mixed category (2-3)	While **[opposing view]**, **[my view]** because **[my reasons]**.
TS is absent.	
TS exists but out of category.	

Table 3.7
Categorization of organizational template use of the participants

Category A	perfectly follows the template taught
Category B	has all elements of the template taught but a different order
Category C	opposing view and its reasons missing
Category D	opposing view and its refutation missing
Mixed category (C-D)	both of C and D missing
Category E	own point and supporting examples missing
Category F	poorly follows the template taught (more than two elements missing)

3.3. Data analysis

This study sought to describe the data gathered from the related participants using descriptive and inferential statistics as well as content analysis. After data

43

collection process details of which were mentioned above, the data were analysed using the Statistical Package for Social Sciences (SPSS) 21st edition software, which determines any significant difference or relationships between/among variables. Besides, some manual analysis such as identifying essay organisation and thesis statement structures, was required as well as the Perl language computer programme designed by the advisor, using count the formulaic language items, identify deviations of them and define text length.

Before proceeding to analyses, all of the essays written by the participants were typed and converted to .txt files and only spelling mistakes were corrected so as not to cause the programme to fail to identify the target formulaic language items. Then, the scores of each rater for the overall argumentative quality of essays, the scores calculated automatically by the computer programme for the use of the target formulaic language items and the codes for thesis statement structures and organisational template were entered in SPSS.

The inter-rater reliability of the raters with the intra-class correlation coefficient was checked first. Then the descriptive analysis were conducted. As a prerequisite of inferential analyses, the variables of each research question was checked for incorrect, out-of-range and missing values, and so were assumptions recommended by parametric tests in order to ease the interpretation of the findings. Any incorrect, out-of-range and missing values were corrected either manually or by the help of SPSS and also the outliers and the assumption of normal distribution were checked by different methods such as histograms, boxplots, Skewness-Kurtosis values and Kolmogrov-Smirnov and Shapiro-Wilk Test. Accordingly, for each question, it was decided whether a parametric test could be run or not. Wherever it was not possible to run a parametric test because

of terminating the assumption of the test, a non-parametric equivalent test was preferred.

CHAPTER IV

4. RESULTS

This chapter provides the major findings of the data analyses carried out to answer the research questions of this study. In order to answer the research questions on a reliable basis, both descriptive and inferential analyses were run in the Statistical Package for Social Sciences (SPSS) 21st edition software.

4.1. Findings related to the first research question

The first research question sought to determine whether there are any differences in the overall argumentative quality scores of the essays written by the students who are explicitly taught formulaic language items and that of those written by the students who are not explicitly taught. After the data set was checked for any missing values or out-of-range values, descriptive statistics, which are illustrated in Figure 4.1., regarding the mean scores taken by each test by both groups, were gathered. It was clearly seen that the experimental group students averaged higher scores than the control group students in each test.
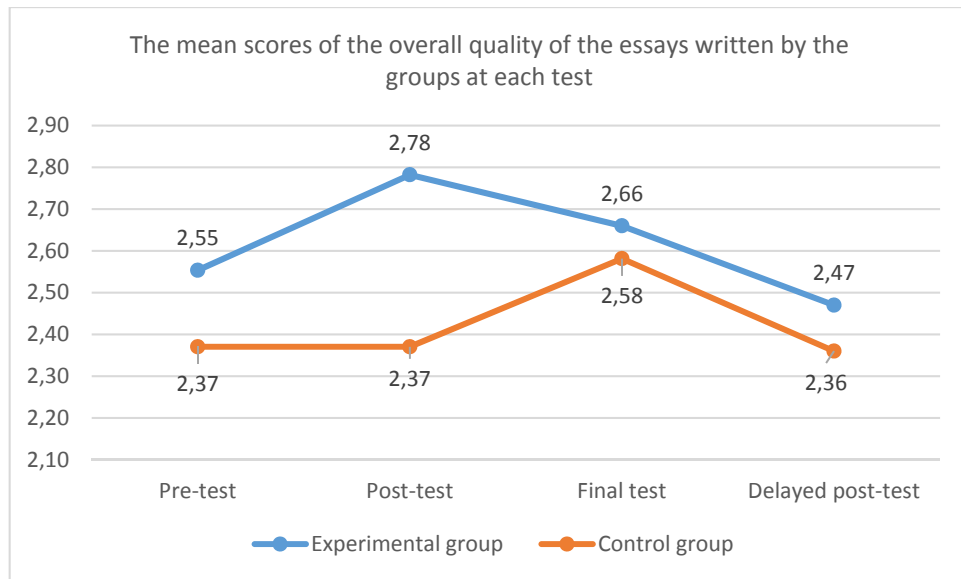
Figure 4.1 The line chart for the mean scores of the overall argumentative quality of the essays written by the groups at each test

In Table 4.1., detailed descriptive statistics indicated that before any explicit instruction took place, although the achievements of the experimental group in the overall quality of the writing was higher than that of the control group, they were close to each other (M EG pre-test=2.55; M CG pre-test=2.37). However, after the explicit instruction to the experimental group, there was an increase in the overall argumentative quality scores of the writing on their post-test (M EG post-test=2.78) while the control group's achievement stayed stable (M CG post-test=2.37). Although the mean for the experimental group decreased in the final test (M EG final test=2.66), compared to their post-test, it was still higher than their pre-test. On the other hand the control group students increased their scores in the final test (M CG final test=2.58). As for the delayed post-test (M EG delayed post-test=2.47; M CG delayed post-test=2.36) the mean of both groups decreased compared to their final and post-test scores, even they were slightly below their pre-test scores.

Table 4.1
Descriptive statistics for the overall quality scores of the essays written by the groups at each test

		N	Min.	Max.	M	SD.
Pre-test	Experimental group	44	1.33	4.00	2.55	.48
	Control group	27	1.33	3.33	2.37	.52
Post-test	Experimental group	44	1.67	3.67	2.78	.48
	Control group	27	1.33	3.67	2.37	.69
Final test	Experimental group	44	1.33	4.00	2.66	.51
	Control group	27	1.33	3.00	2.58	.34
Delayed post-test	Experimental group	26	1.33	3.33	2.47	.66
	Control group	10	1.00	3.33	2.36	.69

Besides the descriptive statistics, some inferential analyses were carried out as well. First, the assumptions of the test were checked. The data set was controlled to determine whether there are any outliers or not for each group and unfortunately some outliers were found (See Appendix G).

Thus, skewness and kurtosis values were also consulted to figure out how the normality was affected by these outliers. While the skewness values were changing in a range between ±1, the kurtosis values were ±3, which seems to be an acceptable rule of thumb (Hair et al., 2006), excluding the values in the domain of the final test (Skewness -1.99; Kurtosis 6.22) (See Table 4. 2.). The histogram graphs (See Appendix H) were also utilized and they indicated that the distribution for the experimental group was leptokurtic for the pre-test, slightly platykurtic for the post-test, almost normal for the final test and negatively skewed for the delayed post-test. On the other hand, for the control group, the distribution was not symmetrical, especially the delayed post-test was negatively skewed (See Appendix H).

Table 4.2
Skewness and Kurtosis values for the overall quality scores of the essays written by the groups at each test

			Skewness		Kurtosis	
		N	Statistic	SD.	Statistic	SD.
Pre-test	Experimental group	44	-.14	.35	2.11	.70
	Control group	27	-.38	.44	-.43	.87
Post-test	Experimental group	44	-.10	.35	.82	.70
	Control group	27	.08	.44	-.74	.87
Final test	Experimental group	44	-.50	.35	1.47	.70
	Control group	27	-1.99	.44	6.22	.87
Delayed post-test	Experimental group	36	-.76	.39	-.38	.76
	Control group	10	-.78	.68	-.14	1.33

As the distribution was observed to satisfy the normality assumption only in a few domains, lastly a Kolmogorov-Smirnov and Shapiro-Wilk test was carried out since it provides a more sensitive analysis for checking normality. The results of the Shapiro-Wilk test indicated that the data significantly deviated from a normal distribution (*p<.05*), excluding the control group pre-test and post-test results (*p>.05*) (See Table 4.3.).

Table 4.3
Test of normality for the overall quality scores of the essays written by the groups at each test

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Pre-test	Experimental group	.23	44	.00	.89	44	.00
	Control group	.17	27	.03	.94	27	.14
Post-test	Experimental group	.25	44	.00	.86	44	.00
	Control group	.11	27	.20[*]	.94	27	.18
Final test	Experimental group	.20	44	.00	.91	44	.00
	Control group	.34	27	.00	.74	27	.00
Delayed post-test	Experimental group	.33	26	.00	.73	26	.00
	Control group	.30	10	.00	.78	10	.00

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Hence, for a safe and grounded interpretation of the analysis, a non-parametric Mann-Whitney U test was utilised to find an answer to the first research question (See Table 4.4.). It could be clearly seen from the table that in each test the experimental group students had higher scores than the students in the control group. However, statistical findings revealed that there was a significant difference in the overall argumentative quality scores of only the post-test essays written, by the students who had been explicitly taught formulaic language items and the ones who had not been (*Z=-2.63; p=.00<.05*).

Table 4.4
Mann-Whitney U test results to compare the overall quality scores of the essays written by the groups at each test

		N	Mean Rank	Sum of Ranks	U	Z	p
Pre-test	Experimental group	44	38.52	1695.00	483.00	-1.36	.17
	Control group	27	31.89	861.00			
Post-test	Experimental group	44	40.89	1799.00	379.00	-2.63	.00
	Control group	27	28.04	757.00			
Final test	Experimental group	44	37.74	1660.50	517.50	-.95	.33
	Control group	27	33.17	895.50			
Delayed post-test	Experimental group	26	18.73	487.00	124.00	-.23	.81
	Control group	10	17.90	179.00			

4.2. Findings related to the second research question

As for the second question which sought out to find an answer whether there is a difference in the argumentative quality of the essays before and after the treatment for the experimental group, a non-parametric Friedman Test of differences among repeated measures was conducted since the data was observed to have a non-normal distribution as stated in Table 4.4. above. The results of Friedman Test (See Table 4.5.) revealed that there was a statistically significant difference in the overall argumentative quality

scores of the essays depending on the tests taken before and after the instruction ($x^2=9.50; p=.02<.05$).

Table 4.5
Friedman Test Results to compare the overall quality scores of the essays written by the experimental group before and after the explicit instruction

	N	Percentiles			Mean Rank	Chi-Square	df	Asymp. Sig.
		25th	50th (Median)	75th				
Pre-test		2.33	2.66	2.66	2.15			
Post-test		2.66	2.66	3.33	2.94			
Final test	26	2.58	2.66	3.00	2.79	9.50	3	.02
Delayed post-test		1.91	2.67	3.00	2.12			

To examine where the differences actually occurred, a post hoc analysis with Wilcoxon signed-rank test was conducted on the different combinations of related variances and also a Bonferroni adjustment was used, since multiple comparisons were made, so as not to make a Type I error while interpreting the results. Bonferroni adjustment was calculated by dividing the initial significance level (.05) by the number of the tests (6) to be run, hence, the new significance level was found as 0.008. From Table 4.6., it could be concluded that at *p<.008* significance level, only the overall argumentative quality scores of the essays between the post-test and the delayed post-test was statistically significantly different (*Z=-2.732; p=.00*).

Table 4.6
Wilcoxon signed-rank test results to compare the overall quality scores of the essays written by the experimental group before and after the explicit instruction

		N	Mean Rank	Sum of Ranks	Z	Asymp. Sig. (2-tailed)
Pre-test – Post-test	Negative Ranks	10[a]				
	Positive Ranks	23[b]	13.80	138.00	-2.56[s]	.01
	Ties	1[c]	18.39	423.00		
	Total	44				
Final test – Pre-test	Negative Ranks	1[d]				
	Positive Ranks	23[e]	19.64	216.00	-1.40[s]	.16
	Ties	10[f]	16.48	379.00		
	Total	44				
Delayed post-test – Pre-test	Negative Ranks	13[g]				
	Positive Ranks	12[h]	14.23	185.00	-.60[t]	.54
	Ties	1[i]	11.67	140.00		
	Total	26				
Final test – Post-test	Negative Ranks	18[j]				
	Positive Ranks	12[k]	18.39	331.00	-2.04[t]	.04
	Ties	14[l]	11.17	134.00		
	Total	44				
Delayed post-test – Post-test	Negative Ranks	17[m]				
	Positive Ranks	9[n]	16.65	283.00	-2.73[t]	.00
	Ties	0[o]	7.56	68.00		
	Total	26				
Delayed post-test – Final test	Negative Ranks	18[p]				
	Positive Ranks	7[q]	13.58	244.50	-2.21[t]	.02
	Ties	1[r]	11.50	80.50		
	Total	26				

[a] Post-test < Pre-test
[b] Post-test > Pre-test
[c] Post-test = Pre-test
[d] Final test < Pre-test
[e] Final test > Pre-test
[f] Final test = Pre-test
[g] Delayed post-test < Pre-test
[h] Delayed post-test > Pre-test
[i] Delayed post-test = Pre-test
[j] Final test < Post-test

[k] Final test > Post-test
[l] Final test = Post-test
[m] Delayed post-test < Post-test
[n] Delayed post-test > Post-test
[o] Delayed post-test = Post-test
[p] Delayed post-test < Final test
[q] Delayed post-test > Final test
[r] Delayed post-test = Final test
[s] Based on negative ranks.
[t] Based on positive ranks.

Although the control group did not receive any explicit instruction, the overall argumentative quality scores of the essays written for the pre-test and for the post-tests were analysed by Friedman test of differences among repeated measures as the data did not satisfy the assumptions of parametric tests as explained in the previous research

question. The test results at Table 4.7. showed mixed results regarding the overall argumentative quality scores of the essays by examining the mean ranks, however, this difference was not statistically significant ($x^2=1.20; p=.75>.05$).

Table 4.7
Friedman Test Results to compare the overall quality scores of the essays written by the control group at the pre-test and the post-tests

| | N | Percentiles | | | Mean Rank | Chi-Square | df | Asymp. Sig. |
		25th	50th (Median)	75th				
Pre-test		1.91	2.33	2.66	2.40			
Post-test	10	1.58	2.33	2.66	2.30	1.20	3	.75
Final test		2.33	2.66	2.75	2.85			
Delayed post-test		2.00	2.33	3.00	2.45			

4.3. Findings related to the third research question

The third question investigated whether there is a difference in the use of the target formulaic language items between the pre-test essays and the post-test essays by the experimental group, and also between the experimental and the control group. The use of the target formulaic language items was analysed under two subtitles regarding the types of the target formulaic language items (each type of formulaic language item was counted only once) and the tokens of the target formulaic language items (each occurrence of formulaic language items regardless of its repetition).

4.3.1. Differences within the experimental group: Types of the target formulaic language items used and the tokens of the target formulaic language item used

First, any missing values in the experimental group data were checked for the use of the target formulaic language items in each essay and the missing data due to the low attendance to the delayed post-test were excluded. Then, the descriptive statistics were gathered and the results shown at Table 4.8. indicated that the tokens of the target formulaic language items used in each essay was more than as of pre-test ($M_{\text{pre-test}}$=1.85, $M_{\text{post-test}}$=3.45, $M_{\text{final test}}$=7.27; $M_{\text{delayed post-test}}$=3.73). Moreover, the experimental group students used more varied target formulaic language items in the post, final and delayed post-test compared to the pre-test ($M_{\text{pre-test}}$=1.64, $M_{\text{post-test}}$=3.40, $M_{\text{final test}}$=6.67; $M_{\text{delayed post-test}}$=3.34).

Table 4.8
The types and tokens of the target formulaic language items used by the experimental group in each test

		N	Min.	Max.	M	SD.
Pre-test	types of the target formulaic language items used	44	.00	6.50	1.64	1.30
	tokens of the target formulaic language items used		.00	6.50	1.85	1.52
Post-test	types of the target formulaic language items used	44	.00	10.00	3.40	2.37
	tokens of the target formulaic language items used		.00	11.00	3.45	2.53
Final test	types of the target formulaic language items used	44	.50	16.50	6.67	3.75
	tokens of the target formulaic language items used		.50	18.00	7.27	4.28
Delayed post-test	types of the target formulaic language items used	26	.00	9.50	3.34	2.54
	tokens of the target formulaic language items used		.00	12.00	3.73	3.08

The figure below also illustrated the increase in both the type and the tokens of the target formulaic language items used by the experimental group students in each essay.
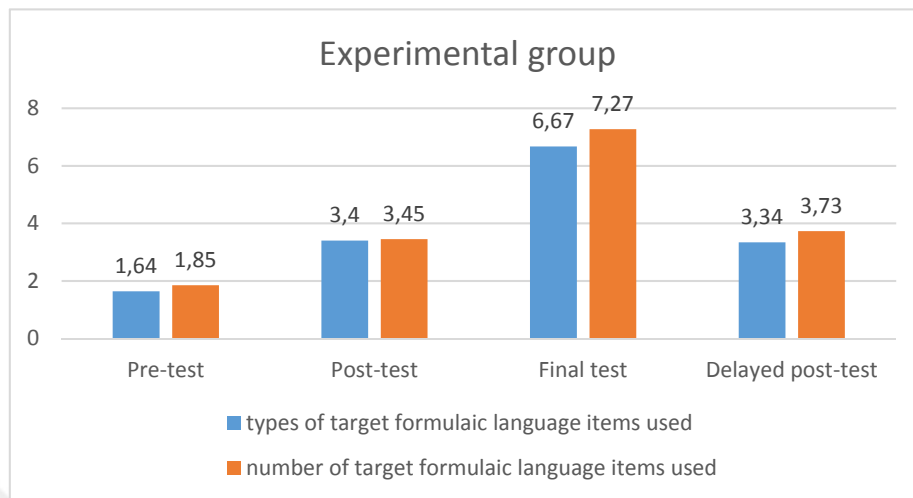
Figure 4.2 The types and tokens of the target formulaic language items used by the experimental group in each test

As for the inferential statistics, skewness and kurtosis values were checked first and the results shown at Table 4.9. were almost indicators of a normal distribution except for the tokens of the target formulaic language items used in the pre-test (Skewness 1.34; Kurtosis 3.28) and the types of target language items used in the delayed post-test (Skewness 1.11).

Table 4.9
*Skewness and Kurtosis values for the types and tokens of the target formulaic
language items used by the experimental group in each test*

			Skewness		Kurtosis	
		N	**Statistic**	**SD.**	**Statistic**	**SD.**
Pre-test	types of the target formulaic language items used	44	1.00	.35	1.05	.70
	tokens of the target formulaic language items used		1.34	.35	3.28	.70
Post-test	types of the target formulaic language items used	44	.72	.35	.28	.70
	tokens of the target formulaic language items used		.65	.35	-.06	.70
Final test	types of the target formulaic language items used	44	.70	.35	.23	.70
	tokens of the target formulaic language items used		.58	.35	.16	.70
Delayed post-test	types of the target formulaic language items used	26	1.11	.45	.71	.88
	tokens of the target formulaic language items used		.86	.45	.02	.88

As some values were found to fall outside the range from -1 to +1 for skewness
and from -3 to +3 for kurtosis, the Kolmogorov-Smirnov and Shapiro-Wilk test, as a
normality test, was run as well. The results of the normality test indicated (See Table
4.10.) that only the types of the target formulaic language items used in the final test
were distributed normally (*p=.10>.05*).

Table 4.10

Test of Normality results for the types and tokens of the target formulaic language items used by the experimental group in each test

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Pre-test	types of the target formulaic language items used	.18	44	.00	.88	44	.00
	tokens of the target formulaic language items used	.18	44	.00	.90	44	.00
Post-test	types of the target formulaic language items used	.11	44	.14	.94	44	.03
	tokens of the target formulaic language items used	.12	44	.06	.93	44	.01
Final test	types of the target formulaic language items used	.08	44	.20[*]	.95	44	.10
	tokens of the target formulaic language items used	.09	44	.20[*]	.94	44	.04
Delayed post-test	types of the target formulaic language items used	.24	26	.00	.91	26	.02
	tokens of the target formulaic language items used	.21	26	.00	.88	26	.00

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

4.3.1.1. Types of the target formulaic language items used

As the data did not meet the normality assumption, a non-parametric Friedman Test of differences among repeated measures was conducted. The results (See Table 4.11.) indicated that there was a statistically significant difference in the types of the target formulaic language items used before and after the instruction ($x^2=50.36$; $p=.00<.05$). Median for the types of the target formulaic language items used for the

pre-test, the post-test, the final and the delayed post-test were 1.5 (.00 to 2), 3.5 (2.50

to 6.12), 8.50 (5.25 to 10) and 2 (1.50 to 5) respectively.

Table 4.11
*Friedman Test results for the types of the target formulaic language items used
among the tests taken by the experimental group*

| | N | Percentiles | | | Mean Rank | Chi-Square | df | Asymp. Sig. |
		25th	50th (Median)	75th				
Pre-test		.00	1.50	2.00	1.44			
Post-test	26	2.50	3.50	6.12	2.69	50.36	3	.00
Final test		5.25	8.50	10.00	3.83			
Delayed post-test		1.50	2.00	5.00	2.04			

Although test results showed that there was an overall difference among the

tests regarding the types of the target formulaic language items, it could not be able to

pinpoint which test in particular differed from the others, thus, a post hoc analysis with

Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied,

resulting in a significance level set at *p<.008*. The results (See Table 4.12.) revealed

that except for the delayed post-test and the post-test *(Z=-1.53; p=.12>.008)*, there was

a statistically significant difference in the type of the target formulaic language item

used between the pre-test and the post-test *(Z=-4.48; p=.00<.008)*, the pre-test and the

final test *(Z=-5.77; p=.00<.008)*, delayed post-test and the pre-test *(Z=-2.90;

p=.00<.008)*, the final test and the post-test *(Z=-4.95; p=.00<.008)*, also between the

delayed post-test and the final test *(Z=-4.27; p=.00<.008)*.

Table 4.12

Wilcoxon signed-rank test results for the types of the target formulaic language items used among the tests taken by the experimental group

		N	Mean Rank	Sum of Ranks	Z	Asymp. Sig. (2-tailed)
Pre-test – Post-test	Negative Ranks	7[a]				
	Positive Ranks	33[b]	11.07	77.50	-4.48[s]	.00
	Ties	4[c]	22.50	742.50		
	Total	44				
Final test – Pre-test	Negative Ranks	0[d]				
	Positive Ranks	44[e]	.00	.00	-5.77[s]	.00
	Ties	0[f]	22.50	990.00		
	Total	44				
Delayed post-test - Pre-test	Negative Ranks	6[g]				
	Positive Ranks	14[h]	4.58	27.50	-2.90[s]	.00
	Ties	6[i]	13.04	182.50		
	Total	26				
Final test– Post-test	Negative Ranks	3[j]				
	Positive Ranks	38[k]	16.17	48.50	-4.95[s]	.00
	Ties	3[l]	21.38	812.50		
	Total	44				
Delayed post-test– Post-test	Negative Ranks	18[m]				
	Positive Ranks	8[n]	13.08	235.50	-1.53[t]	.12
	Ties	0[o]	14.44	115.50		
	Total	26				
Delayed post-test– Final test	Negative Ranks	24[p]				
	Positive Ranks	2[q]	14.31	343.50	-4.27[t]	.00
	Ties	0[r]	3.75	7.50		
	Total	26				

[a] Post-test < Pre-test
[b] Post-test > Pre-test
[c] Post-test = Pre-test
[d] Final test < Pre-test
[e] Final test > Pre-test
[f] Final test = Pre-test
[g] Delayed post-test < Pre-test
[h] Delayed post-test > Pre-test
[i] Delayed post-test = Pre-test
[j] Final test < Post-test

[k] Final test > Post-test
[l] Final test = Post-test
[m] Delayed post-test < Post-test
[n] Delayed post-test > Post-test
[o] Delayed post-test = Post-test
[p] Delayed post-test < Final test
[q] Delayed post-test > Final test
[r] Delayed post-test = Final test
[s] Based on negative ranks.
[t] Based on positive ranks

4.3.1.2. The tokens of the target formulaic language items used

The results of Friedman Test of differences among repeated measures, conducted to compare the tokens of the target formulaic language items used in each test by the experimental group, demonstrated (See Table 4.13.) that there was a

statistically significant difference in the tokens of the target formulaic language items used before and after the instruction ($x^2=48.86$; $p=.00<.05$).

Table 4.13
Friedman Test results for the tokens of the target formulaic language items used among the tests taken by the experimental group

| | N | Percentiles | | | Mean Rank | Chi-Square | df | Asymp. Sig. |
		25th	50th (Median)	75th				
Pre-test		.00	1.50	3.00	1.54			
Post-test		2.12	3.50	6.62	2.52			
Final test	26	5.87	8.50	10.62	3.85	48.86	3	.00
Delayed post-test		1.87	2.25	5.62	2.10			

In order to define the differences among the tokens of the target formulaic language items used in each test, a post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at *p<.008*. The results (See Table 4.14.) revealed that there was a statistically significant difference between the pre-test and the post-test *(Z=-4.48; p=.00<.008)*, the pre-test and the final test *(Z=-5.71; p=.00<.008)*, the final test and the post-test *(Z=-5.28; p=.00<.008)*, also between the delayed post-test and the final test *(Z=-4.23; p=.00<.008)*, except for the delayed post-test and the post-test *(Z=-.54; p=.58>.008)* and also between the delayed post-test and the pre-test *(Z=-2.51; p=.01>.008)*.

Table 4.14
Wilcoxon signed-rank test results for the tokens of the target formulaic language
items used among the tests taken by the experimental group

		N	Mean Rank	Sum of Ranks	Z	Asymp. Sig. (2-tailed)
Pre-test – Post-test	Negative Ranks	9[a]			-4.04[s]	.00
	Positive Ranks	31[b]	12.28	110.50		
	Ties	4[c]	22.89	709.50		
	Total	44				
Final test – Pre-test	Negative Ranks	0[d]			-5.71[s]	.00
	Positive Ranks	43[e]	.00	.00		
	Ties	1[f]	22.00	946.00		
	Total	44				
Delayed post-test - Pre-test	Negative Ranks	7[g]			-2.51[s]	.01
	Positive Ranks	14[h]	6.21	43.50		
	Ties	5[i]	13.39	187.50		
	Total	26				
Final test– Post-test	Negative Ranks	3[j]			-5.28[s]	.00
	Positive Ranks	39[k]	9.83	29.50		
	Ties	2[l]	22.40	873.50		
	Total	44				
Delayed post-test– Post-test	Negative Ranks	16[m]			-.54[t]	.58
	Positive Ranks	10[n]	12.31	197.00		
	Ties	0[o]	15.40	154.00		
	Total	26				
Delayed post-test– Final test	Negative Ranks	24[p]			-4.23[t]	.00
	Positive Ranks	2[q]	14.25	342.00		
	Ties	0[r]	4.50	9.00		
	Total	26				

[a] Post-test < Pre-test
[b] Post-test > Pre-test
[c] Post-test = Pre-test
[d] Final test < Pre-test
[e] Final test > Pre-test
[f] Final test = Pre-test
[g] Delayed post-test < Pre-test
[h] Delayed post-test > Pre-test
[i] Delayed post-test = Pre-test
[j] Final test < Post-test

[k] Final test > Post-test
[l] Final test = Post-test
[m] Delayed post-test < Post-test
[n] Delayed post-test > Post-test
[o] Delayed post-test = Post-test
[p] Delayed post-test < Final test
[q] Delayed post-test > Final test
[r] Delayed post-test = Final test
[s] Based on negative ranks.
[t] Based on positive ranks

4.3.2. Differences between the experimental and the control groups: Types and the tokens of the target formulaic language items used

In order to define any differences between the experimental and the control group, the descriptive statistics were utilised after checking the missing values. Figure 4.3. and 4.4. illustrated that both of the groups increased the use of the target formulaic language items. The experimental group had higher mean scores for both the type and the tokens of the target formulaic language items, though. Although a decrease was observed in the delayed post-test for both groups, regarding the types and the tokens of the target formulaic language used, the percentages were still higher than as of the pre-test.
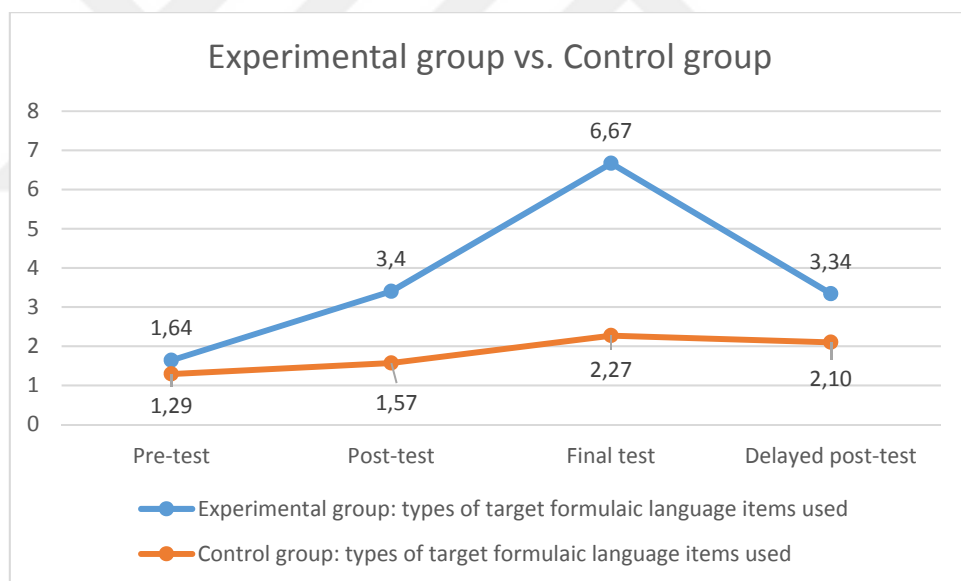
Figure 4.3 The line chart for the types of the target formulaic language items used by the groups at each test
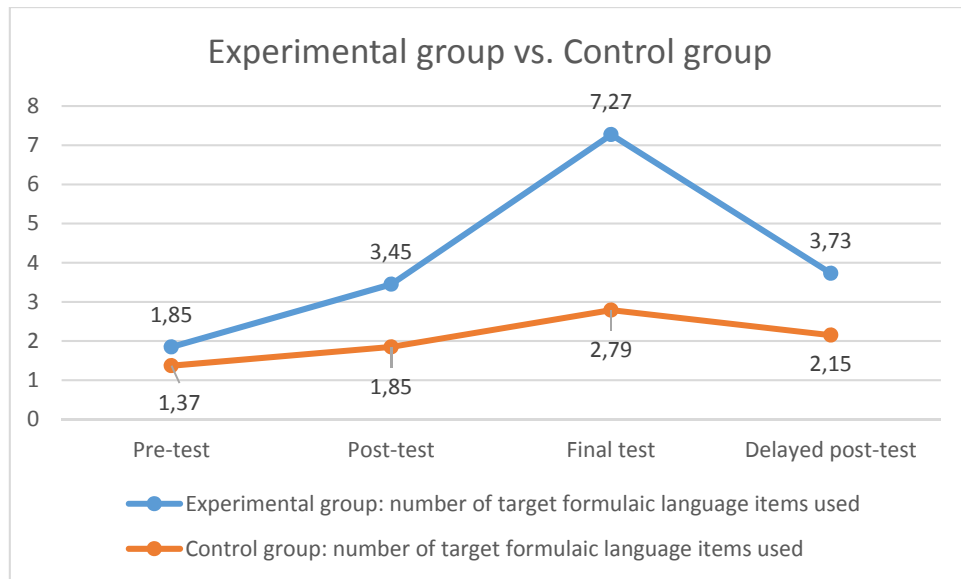
Figure 4.4 The line chart for the tokens of the target formulaic language items used by the groups at each test

The results shown at Table 4.15. indicated that the experimental group used greater number of the target formulaic language items at each test. This gap was small in the pre-test (*M* $_{EG\ pre\text{-}test\ target\ formulaic\ language\ use}$=1.85; *M* $_{CG\ pre\text{-}test\ target\ formulaic\ language\ use}$=1.37), however with the post-test it almost doubled (*M* $_{EG\ post\text{-}test}$=3.45; *M* $_{CG\ post\text{-}test}$=1.85) and in the final test the experimental group used almost three times more target formulaic language items than the control group (*M* $_{EG\ final\ test}$=7.27; *M* $_{CG\ final\ test}$=2.79) and in the delayed post-test the use of formulaic language items of the experimental group was higher (*M* $_{EG\ delayed\ post\text{-}test}$=3.73; *M* $_{CG\ delayed\ post\text{-}test}$=2.15) .

Moreover, the descriptive statistics revealed that experimental group students used more varied target formulaic language items than the control group students in each test. In the pre, post and final test the experimental group's mean for the types of the target formulaic language items used, increased gradually (*M* $_{EG\ pre\text{-}test}$=1.64; *M* $_{EG\ post\text{-}test}$=3.40; *M* $_{EG\ final\ test}$=6.67) while it was falling and rising for the control group (*M*

62

$_{CG\ pre\text{-}test}$=1.29, M $_{CG\ post\text{-}test}$=1.57 and M $_{CG\ final\ test}$=2.27). As for the delayed post-test, the variety of the target formulaic language items used decreased for both groups, compared to the final test, but the means were still higher than as of the pre-test (M $_{EG}$ $_{delayed\ post\text{-}test}$ =3.34; M $_{CG\ delayed\ post\text{-}test}$ =2.10).

Table 4.15
The comparison of the target formulaic language items used by the groups in each test

		The types of the target formulaic language items used					The tokens of the target formulaic language items used				
		N	Min.	Max	M	SD.	N	Min.	Max.	M	SD.
Pre-test	Experimental Group	44	.00	6.50	1.64	1.30	44	.00	6.50	1.85	1.52
	Control Group	27	.00	3.00	1.29	1.04	27	.00	4.00	1.37	1.25
Post-test	Experimental Group	44	.00	10.00	3.40	2.37	44	.00	11.00	3.45	2.53
	Control Group	27	.00	4.00	1.57	1.28	27	.00	6.50	1.85	1.70
Final Test	Experimental Group	44	.50	16.50	6.67	3.75	44	.50	18.00	7.27	4.28
	Control Group	27	.50	5.00	2.27	1.22	27	.50	5.50	2.79	1.44
Delayed post-test	Experimental Group	26	.00	9.50	3.34	2.54	26	.00	12.00	3.73	3.08
	Control Group	10	.50	5.00	2.10	1.35	10	.50	4.50	2.15	1.35

In order to utilise the inferential statistics, the test of normality was conducted first and the results shown at Table 4.16. were the indicators of a non-normal distribution *(p<.05,* except the types *(p =.19>.05)* and the tokens *(p =.34>.05)* of the target formulaic language items used in the delayed post-test.

Table 4.16
Test of Normality results for the types and tokens of the target formulaic language items used by the control group

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	**df**	**Sig.**	**Statistic**	**df**	**Sig.**
Pre-test	types of the target formulaic language items used	.19		.01	.88		.00
			27			27	
	tokens of the target formulaic language items used	.16		.06	.89		.00
Post-test	types of the target formulaic language items used	.18		.01	.89		.01
			27			27	
	tokens of the target formulaic language items used	.16		.04	.89		.00
Final test	types of the target formulaic language items used	.21		.00	.92		.03
			27			27	
	tokens of the target formulaic language items used	.22		.00	.89		.01
Delayed post-test	types of the target formulaic language items used	.23		.14	.89		.19
			10			10	
	tokens of the target formulaic language items used	.20		.20[*]	.91		.34

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

4.3.2.1. Types of the target formulaic language items used

Since the data regarding the control group was not normally distributed, as it was for the experimental group, the non-parametric Mann-Whitney U test was utilised (See Table 4.17.) in order to compare the experimental and the control group regarding their use of the target formulaic language items in variety. The results revealed that there was a statistically significant difference in the type of the target formulaic language items used in the post-test *(Z=-3.40; p=.00<.05)* and the final test *(Z=-5.25; p=.00<.05)*, between the students who had been explicitly instructed and the ones who had not. The mean rank for the post-test and the final test of the experimental group students were respectively 42.5 and 46.06 while they were 25.41 and 19.61 for the control group students. The analysis showed no statistically significant difference between the pre-tests of the experimental group and the control group *(Z=-.85; p=.39>.05)* and between the delayed post-tests of the both groups *(Z=-1.28; p=.19>.05)*. However, an examination of the mean ranks of each test scores demonstrated that the students in the experimental group attained higher scores after the explicit instructions when compared to their peers in the control group.

Table 4.17
Mann-Whitney U test results to compare the types of the target formulaic language items used by the groups in each test

		N	Mean Rank	Sum of Ranks	U	Z	p
Pre-test	Experimental group	44	37.60	1654.50	523.50	-.85	.39
	Control group	27	33.39	901.50			
Post-test	Experimental group	44	42.50	1870.00	308.00	-3.40	.00
	Control group	27	25.41	686.00			
Final test	Experimental group	44	46.06	2026.50	151.50	-5.25	.00
	Control group	27	19.61	529.50			
Delayed post-test	Experimental group	26	19.88	517.00	94.00	-1.28	.19
	Control group	10	14.90	149.00			

4.3.2.2. The tokens of the target formulaic language items used

In order to compare the experimental and the control group regarding the tokens of the target formulaic language items that they used in each essay, the Mann-Whitney U test was utilised (See Table 4.18.), because the data belonging to control group did not distribute normally as it was in the experimental group. The analysis resulted in a statistically significant difference between the control and the experimental group in the tokens of the target formulaic language items used in the post-test *(Z=-2.73; p=.00<.05)* and the final test *(Z=-4.90; p=.00<.05)*. The mean rank for the post-test and the final test of the experimental group students were respectively 41.22 and 45.36 while they were 27.50 and 20.74 for the control group students. The results indicated that there was not a statistically significant difference between the pre-tests of the experimental group and the control group *(Z=-.96; p=.33>.05)* and between the delayed post-tests of the both groups *(Z=-1.26; p=.20>.05)*. However, an examination of the mean ranks for each test demonstrated that the students in the experimental group attained higher scores after the explicit instructions when compared to their peers in the control group.

Table 4.18
Mann-Whitney U test results to compare the tokens of the target formulaic language items used by the groups in each test

		N	Mean Rank	Sum of Ranks	U	Z	p
Pre-test	Experimental group	44	37.27	1640.00	494.00	-.96	.33
	Control group	27	32.50	845.00			
Post-test	Experimental group	44	41.22	1813.50	364.50	-2.73	.00
	Control group	27	27.50	742.50			
Final test	Experimental group	44	45.36	1996.00	182.00	-4.90	.00
	Control group	27	20.74	560.00			
Delayed post-test	Experimental group	26	19.87	516.50	94.50	-1.26	.20
	Control group	10	14.95	149.50			

4.4. Findings related to the fourth research question

The fourth research question sought to determine any relationship between the tokens of the target formulaic language items used and the overall argumentative quality of the essays, as well as between the type of the target formulaic language items used and the overall quality scores of the essays.

The data regarding the essay scores and the type and the tokens of the target formulaic language items used in each essay were checked for any missing values. After the missing data for the delayed post-test, caused by the low attendance, were excluded, 249 essays were analysed. Before inferential analyses, descriptive statistics were gathered and the test of normality was conducted. The mean score was 2.28 for the type of the target formulaic language items used, 3.35 for the tokens of the target formulaic language items used and it was 3.3 over 4 for the overall argumentative quality of the essays.

Table 4.19
Descriptive statistics for the types and tokens of the formulaic language items used in all of the essays

	N	Min.	Max.	M	SD.
type of the target formulaic language items used		.00	6.50	2.28	.93
tokens of the target formulaic language items used	249	.00	18.00	3.35	3.21
overall argumentative quality scores of the essays		.00	16.50	3.34	2.69

The results of normality tests and boxplots (See Appendix G) indicated that the tokens and the type of the target formulaic items used and the overall argumentative quality scores of the essays were observed to distribute non-normally *(p=.00<.05)*.

Thus, the non-parametric Spearman rank-order correlation coefficient test was conducted.

Table 4.20
Test of Normality results for the use of target formulaic language use and the overall quality scores of the essays for both groups

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
type of the target formulaic language items used	.18		.00	.91		.00
tokens of the target formulaic language items used	.18	249	.00	.83	249	.00
overall argumentative quality scores of the essays	.22		.00	.82		.00

a. Lilliefors Significance Correction

The results of the Spearman's rank-order correlation test (See Table 4.20.) which was run to define any relationship between the overall argumentative quality scores of the essays and the tokens of the target formulaic language items indicated that there was a strong relationship between them ($r_s =.79, p=.00<.05$); however, there was not a relationship between the overall argumentative quality scores of the essays and the type of the target formulaic language used ($r_s =.10, p=.09>.05$).

Table 4.21
Spearman rank-order correlation coefficient test to define the relationship between the overall argumentative quality scores of the essays and the use of the target formulaic language items

			type of the target formulaic language items used	tokens of the target formulaic language items used	overall argumentative quality scores of the essays
Spearman's rho	type of the target formulaic language items used	Correlation Coefficient	1.00	.41**	.10
		Sig. (2-tailed)	.	.00	.09
		N	249	249	249
	tokens of the target formulaic language items used	Correlation Coefficient	.419**	1.00	.79**
		Sig. (2-tailed)	.00	.	.00
		N	249	249	249
	overall argumentative quality scores of the essays	Correlation Coefficient	.10	.79**	1.00
		Sig. (2-tailed)	.09	.00	.
		N	249	249	249

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

4.5. Findings related to the fifth research question

The last research question investigated to what extent the explicit teaching of organisational template and the structures of thesis statement affected the organisation of the essays. The results gathered by utilizing the percentages for the students who used the categories defined for both organizational template and thesis statement structures, indicated to what extent the participants used the organizational template and thesis statement structures recommended.

4.5.1. The findings regarding the explicit teaching of organisational template

The results at Table 4.21., indicate that before the intervention, much higher percentage of the students in the control group (*14.8%, N=4*) than the experimental group (*11.4%, N=5*), followed the recommended template perfectly on the pre-test which took place prior to explicit teaching. After explicit teaching, however, the students in the experimental group (*36.4%, N=16*) were observed to perfectly follow the recommended template, three times more than the control group (*11.1%, N=3*). Moreover, the percentage of the experimental group (*38.6%, N=17*) at following the recommended template perfectly increased in the final test while the percentage of the control group decreased (*3.7%, N=1*), even more when compared to the pre-test and the post-test. Even though fewer students ($N_{EG}=26; N_{CG}=10$) of both groups participated in the delayed post-test, the percentages indicated that 7.7% *(N=2)* of the experimental group followed the recommended template perfectly while no one of the control group.

As for the second category, while 38.6 % (*N=17*) of the students in the experimental group had all elements of the organisation template in a different order, for the pre-test, this percentage was 11.4% (*N=5*) for the post-test, 36.4% (*N=16*) for the final test and 11.5% *(N=3)* for the delayed post-test. On the other hand, 11.1% (*N=3*) of the control group students had all elements of the organisational template in a different order for both the pre-test and the post-test, however this percentage increased to 59.3% (*N=16*) for the final test and decreased again for the delayed post-test (*30%, N=3*).

Another analysis was implemented to find out the missing elements of the organisational template when the students did not follow the recommended template perfectly. Some of the students of both groups did not use the elements regarding opposing views, its reasons and its refutation or their own views with supporting examples. To illustrate, although the percentage was changing at a range between 2.3% ($N=1$) and 6.8% ($N=3$) for each test, in total for all tests 22.8% of the students in the experimental group did not include one of the elements recommended (opposing view and its reasons, its refutation or own point with supporting examples) in their writing, while this percentage was 18.5% ($N=5$) for the control group in total for all tests, changing between 3.7% ($N=1$) and 7.4% ($N=2$). Moreover, quite a high percentage of the student in both groups, included neither opposing view and its reasons nor refutation of an opposing view at the same time. The lack of these two elements was gradually decreased for the experimental group ($EG_{pre-test}=31.8\%$ ($N=14$); $EG_{post-test}=29.5\%$ ($N=13$); $EG_{final\ testt}=13.6\%$ ($N=6$) until the delayed post-test at which 61.5% ($N=16$) of the experimental group students had these two elements missing together. For the control group students who did not include these two elements together in their essays, mixed results were found ($CG_{pre-test}=48.1\%$, $N=13$; $CG_{post-test}=55.6\%$, $N=15$; $CG_{final\ test}=18.5\%$, $N=5$; $CG_{delayed\ post-test}=20\%$, $N=2$).

Lastly, the percentage of the students, in the control group ($CG_{pre-test}=14.8\%$, $N=4$; $CG_{post-test}=22.2\%$, $N=6$; $CG_{final\ test}=11.1\%$, $N=3$; $CG_{delayed\ post-test}=50\%$, $N=5$), having more than two missing elements was at a higher as a percentage than the ones in the experimental group for each test ($EG_{pre-test}=11.4\%$, $N=5$; $EG_{post-test}=13.6\%$, $N=6$; $EG_{final\ test}=4.5\%$, $N=2$; $EG_{delayed\ post-test}=19.2\%$, $N=5$).

Table 4.22
Frequency and percentage for the organisational template used by the groups

	Experimental group				Control group			
	f %	*f* %	*f* %	*f* %	*f* %	*f* %	*f* %	*f* %
	Pre-test $N=44$	Post-test $N=44$	Final test $N=44$	Delayed post-test $N=26$	Pre-test $N=27$	Post-test $N=27$	Final test $N=27$	Delayed post-test $N=10$
Category A (perfectly follows the template taught)	5 (11.4%)	16 (36.4%)	17 (38.6%)	2 (7.7%)	4 (14.8%)	3 (11.1%)	1 (3.7%)	0 (0%)
Category B (has all elements of the template taught but a different order)	17 (38.6%)	5 (11.4%)	16 (36.4%)	3 (11.5%)	3 (11.1%)	3 (11.1%)	16 (59.3%)	3 (30%)
Category C (opposing view and its reasons missing)	1 (2.3%)	3 (6.8%)	3 (6.8%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Category C-D (opposing view, its reasons and its refutation missing)	14 (31.8%)	13 (29.5%)	6 (13.6%)	16 (61.5%)	13 (48.1%)	15 (55.6%)	5 (18.5%)	2 (20%)
Category D (opposing view and its refutation missing)	1 (2.3%)	1 (2.3%)	0 (0%)	0 (0%)	2 (7.4%)	0 (0%)	1 (3.7%)	0 (0%)
Category E (own point and supporting examples missing)	1 (2.3%)	0 (0%)	0 (0%)	0 (0%)	1 (3.7%)	0 (0%)	1 (3.7%)	0 (0%)
Category F (poorly follows the template taught -more than two elements missing)	5 (11.4%)	6 (13.6%)	2 (4.5%)	5 (19.2)	4 (14.8%)	6 (22.2%)	3 (11.1%)	5 (50%)

4.5.2. The findings regarding the explicit teaching of thesis statement structures

The analysis of the second part of the research question regarding the structures of the thesis statement was demonstrated at Table 4.22.. Accordingly, it can be seen that in the pre-test, more than half of the thesis statements written by both groups (*EG 52.3%, N=23); CG 63%, N=17*) fell in a category that does not exist, yet for both groups out of category thesis statement percentages showed a gradual decrease for the post and the final test while it increased again for the delayed post-test. The percentages were still higher for the control group students (*post-test 59.3%, N=16; final test 44.4%, N=12; delayed post-test 80%, N=8*) compared to their counterparts in the experimental group (*post-test 36.4%, N=16; final test 34.1%, N=15, delayed post-test 61.5%, N=16*), though. See below examples;

Student A:

"Even though English is not the mother tongue of some countries, can they be as good as the native speakers?"

Student B:

"Teaching anything or any language is not easy, you have to feel comfortable about teaching and you should know the details and the including of the topic well."

Moreover, a few students from the both group organised the essays without writing a thesis statement; as a percentage it was 14.5 (*pre-test 4.5%, N=2; post-test 2.3%, N=1; final test 0%, N=0; delayed post-test 7.7%, N=2*) for the experimental

group and 21.1% (*pre-test 0%, N=0; post-test 11.1%, N=3; final test 0%, N=0; delayed post-test 10%, N=1*) for the control group.

Additionally, from the control group, not a single student did write a thesis statement from the first category (*Although many people claim that **[opposing view]**, closer examination shows that **[my view].***), but from the experimental group three students (*6.8%*) at the post-test, nine students (*20.5%*) at the final and one student at the delayed post-test *(3.8%)* preferred to write a thesis statement from the first category. The below example, is a perfect match of the first category that was taught.

Student C:

"Although, some people claim that **[native speakers cannot teach well]**, the closer examination shows that **[native talking teachers of English are the ones who can teach English best]**."

Category 1: Although many people claim that **[opposing view]**, closer examination shows that **[my view].**

As for the second category *(While **[opposing view]**, **[my view].**)*, both groups increased the use of this category gradually but decreased in the delayed post-test. The percent of the experimental group students using the second category was 15.9% *(N=7)* for the pre-test, 20.5% *(N=9)* for the post-test, 20.5% *(N=11)* in the final test and 11.5% *(N=3)* in the delayed post-test. On the other hand, the percentages of the control group students, which are lower for each essay, were 3.7% *(N=1)* in the pre-test, 7.4% *(N=2)* in the post-test, 14.8% *(N=4)* in the final test and 10% *(N=1)* in the delayed post-test. See example below;

Student D:

"While some people think that **[raise a child in a small town is ideal]**, I believe that **[the most ideal place to bring up a child is big cities]**."

Category 2: While **[opposing view]**, **[my view]**.

The third category usage *([my view] because [my reasons].)*, on the other hand, revealed mixed results not only for the experimental group (*pre-test 27.3%, N=12; post-test 29.5%, N=13; final test 13.6%, N=6; delayed post-test 11.5%, N=3*) but also for the control group (*pre-test 33.3%, N=9; post-test 22.2%, N=6; final test 40.7%, N=11; delayed post-test 0%, N=0*). See examples below;

Student E:

"In my opinion **[cell phone usage should be allowed in classrooms]** for many reasons such as: **[allows student gather information from the internet, act as a calculator and cell phone and specific sites should be used to make class more fun]**."

Category 3: **[my view]** because **[my reasons].**

Lastly, the experimental group was observed to use a new mixed category surprisingly *(While [opposing view], [my view] because [my reasons].)* by blending the second and third category (*post-test 4.5%, N=2; final test 6.8%, N=3; delayed post-test 3.8%, N=1*), as it is in the below example.

Student F:

"While some people claim that **[having excellent knowledge of the subject being taught is more important]**, **[the most significant issue is the teachers ' ability to relate well with students]** because **[student wouldn't want to study with a teacher who doesn't care about their students and teachers who have poor relationship with their student do not want to teach eagerly as it would be just to gain money]**."

Mixed category (2-3): While **[opposing view]**, **[my view]** because **[my reasons]**

Table 4.23
Frequency and the percentage for the thesis statement structures used by the groups

	Experimental group				Control group			
	f %	*f* %	*f* %	*f* %	*f* %	*f* %	*f* %	*f* %
	Pre-test N=44	Post-test N=44	Final test N=44	Delayed post-test N=26	Pre-test N=27	Post-test N=27	Final test N=27	Delayed post-test N=10
1.category Although many people claim that **[opposing view]**, closer examination shows that **[my view].**	0 (0 %)	3 (6.8 %)	9 (20.5 %)	1 (3.8 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
2.category While **[opposing view]**, **[my view]**.	7 (15.9 %)	9 (20.5 %)	11 (25 %)	3 (11.5 %)	1 (3.7 %)	2 (7.4 %)	4 (14.8 %)	1 (10 %)
3.category **[my view]** because **[my reasons].**	12 (27.3 %)	13 (29.5 %)	6 (13.6 %)	3 (11.5 %)	9 (33.3 %)	6 (22.2 %)	11 (40.7 %)	0 (0 %)
Mixed category (2-3) While **[opposing view]**, **[my view]** because **[my reasons]**.	0 (0 %)	2 (4.5 %)	3 (6.8 %)	1 (3.8 %)	0 (0 %)	0 (0 %)	0 (0 %)	0 (0 %)
At least one recommended category used	19 (43.2 %)	27 (61.5%)	29 (65.9%)	8 (30.6%)	10 (37%)	8 (29.6%)	15 (55.5%)	1 (10%)
No thesis statement **TS is absent.**	2 (4.5 %)	1 (2.3 %)	0 (0 %)	2 (7.7 %)	0 (0 %)	3 (11.1 %)	0 (0 %)	1 (10 %)
Thesis statement out of category **TS exists but out of category.**	23 (52.3 %)	16 (36.4 %)	15 (34.1 %)	16 (61.5 %)	17 (63.0 %)	16 (59.3 %)	12 (44.4 %)	8 (80 %)

CHAPTER V

5. DISCUSSION and CONCLUSION

In this chapter, all findings were discussed through summaries with respect to the main findings of each research questions formulated for the study. It is organised in five sections. The first section begins with a summary and discussion of the main findings addressing each of the research questions. The second section considers the implications for practice. It is followed by the third section in which the limitations of the study were discussed and the fourth one recommends for further studies. Finally, conclusion is presented.

5.1. Summary and discussion of the findings

The present paper investigated the effects of the explicit teaching of formulaic language on academic writing, specifically argumentative writing. In some ways, the findings in the present paper are in compliance with the previous studies and contradicts in some issues. Additionally, it contributes notably to the literature providing an understanding of how the explicit teaching affects the use formulaic language on academic writing.

5.1.1. The overall argumentative quality of the writing

The findings demonstrated that after the treatment, the overall argumentative quality scores of the essays written in the post-test by the experimental group students increased sharply while it neither increased nor decreased, but stayed stable for the control group. Moreover, statistical findings also confirmed the significant difference between two groups, and also the difference in the quality of the essays before and after the treatment. The results did not correspond to the findings of Čolović-Marković

(2012) who compared the in-class 40-minute argumentative essays of the students before and after the treatment and her study resulted in no significant difference between the students receiving treatment and the ones who did not. The reasons of these differences between the studies might be varied. For instance, different rubrics were employed to evaluate the overall quality of the essays. The rubric in the present paper was a holistic one while Čolović-Marković (2012) employed a more analytical rubric focussing on five aspects of writing separately. Thus, the participants in her study might have been scored less. Another factor might have been the individual differences between the participants of the studies such as language aptitude, motivation and learning strategies. Even though student interviews empirically were not conducted in the present study, some discussions took place during the class-time to receive the reflections of the students. The reflections gathered in the present study and the interviews Čolović-Marković (2012) conducted with the participants of her study might have contributed to the argument of individual differences. To illustrate, the participants involved in her study were multinational and enrolled in an Intensive English Programme in the USA while the ones in the present study was enrolled in an ELT bachelor programme in Turkey, and mostly Turkish. The multinational participants might have had some difficulty during the treatment due to their educational background, both in terms of majors and the education systems. For example, one of the low achieving student in her study stated that his major was business and he was not much interested in formulaic sequences, on the other hand the participants of the present paper were studying to be a language teacher, which was an important motivating factor. Similar to the ones in the present paper, high achieving participants of her study stated they invested more effort and time outside the class by employing

different learning strategies such memorization and reading with specific focus on formulaic language items.

After the post-test, the overall argumentative quality scores of the essays of the experimental group students declined in the final test but still higher than as of the control group and their own pre-test average. On the other hand, the control group surprisingly did their best scores higher than their post-test, whose reason might have been that the final test had more weight than other assignments and tests, on the evaluation of the course and this might have motivated them to write better quality essays. As for the experimental group, they may not have given enough importance to their argumentative writing for the final test as they already had higher scores in the previous tests.

The last findings on the overall argumentative quality scores of the essays were that the delayed post-test scores of both groups dropped; bewilderingly, the mean scores in the delayed post-test were even lower than the ones in the pre-test. In this respect, one might argue that the explicit teaching of formulaic language items may not have been directly linked to the overall quality of the writing and there are some other aspects like organisation, mechanics, and language use. Nevertheless, it is a pertinent remark to state the explicit teaching of formulaic language is effective to improve the overall argumentative quality of the writing as the experimental group students received higher scores for the overall argumentative quality of the essays once they used more formulaic language items in each test than their counterparts in the control group. This result is also supported with the studies of Ohlrogge (2009), Kennedy and Thorp (2007), Read and Nation (2006) and Bonk (2001) who examined either written or oral productions of the learners taking part it high-stakes proficiency exams such as IELTS, ECCE and

81

TOEFL. They compared the overall scores of the candidates to the formulaic language items used and resulted in that the more formulaic language items the candidates used the higher score they received by the raters. Additionally, one might note that the positive effects of an intervention conducted for a short period, were short term as well since the gains decreased in the delayed post-test. In order to obtain permanent improvement, one might think to embed explicit teaching of formulaic language into the syllabus in the long run.

5.1.2. Target formulaic language items

The findings showing that the experimental group students used more type and tokens of the target formulaic language items after the treatment, indicated that the explicit teaching helped them to utilise formulaic language items progressively in their writing. Similarly, Peters and Pauwels (2015) who researched the recognition and spontaneous use of formulaic sequences, comparing the pre-test and post-test writing of 29 participants, observed an increase of %13 in the types and %11 in the tokens of formulaic language items after the treatment which lasted three weeks. While the increase was humble in their study, in the present paper it almost doubled in the post-test and three times more in the final test. This huge difference in increase between two studies might have been caused by the treatment content. While Peters and Pauwels (2015) included activities required underlining and filling the gaps with the help of given cues, in the present study, besides similar activities some more activities were employed such as translation and cloze exercises, giving a reference list and discussion of the answers of the quiz. On the contrary, Jones and Haywood (2004) in their study in which they both aimed to analyse the awareness and production of formulaic sequences, interpreted the results regarding the production of formulaic sequences of

their participants as inconclusive since except one student, the other six students achieved no improvement in the number of formulaic sequences in their post-test essays. Although the participants in their study received more hours of instruction and be exposed to more reading materials for ten weeks, some other factors affecting the results were reported such as the small number of the participants (ten experimental, eleven control group) and their absenteeism to both treatment and the post-test, time constraints pushing them to conduct the post-test only after two weeks instead of four and choosing a topic related to the texts in the class for the pre-test. Moreover, the situation of the only student used more formulaic sequences in the post-test was revealed by the interviews after the pre-test. She had attended English language courses, been instructed in English in Ph.D. courses and taken English proficiency exams, thus she was quite aware of the importance of formulaic sequences and already defined a learning strategy for herself such as recording formulaic sequences in a vocabulary notebook and repeating them. If Jones and Haywood (2004) had been able to conduct the interviews at the end of the study as they expected, the results and the causes of the lack of improvement on the production of formulaic sequences could have been more explicatory. Similar to the findings of Jones and Haywood (2004), Cortes (2006) did not observe any progress in the number of the formulaic items in the written assignments of the participants who were native speakers of English, after the treatment including five 20-minute micro lessons during ten weeks. In her study, Cortes (2006) attributed the reasons to the short instruction period and the activities which may not be appropriate to activate autonomous use of formulaic language items by the students. Although she conducted the instruction more often than the present study, as she stated, total duration of instruction was less than as of the present study which was approximately four hours during two consecutive weeks.

In the present paper, differently than the below mentioned studies, a delayed post-test was conducted as well in order to seek out the effects of the explicit teaching of the target formulaic language items, if any, in the long run. The results indicated that there was a decrease in the delayed post-test; however, the mean scores were still higher than as of the pre-test. The difference between the mean scores of the students in the post-test and the delayed post-test was not found statistically significant. The reason of the decrease might have been that the students were not given any instruction or advised to revise the items during the duration between the post-test and the delayed post-test. On this point, although there was not a statistical significant difference, it could be suggestive to argue that explicit teaching fosters the students' use of formulaic language items by raising awareness, and without explicit teaching, unconsciously encountered formulaic language items might be ignored or forgotten in time with fewer gains. This is in line with the results of the study conducted by Alhassan and Wood (2015) who taught formulaic language items to twelve participants over ten weeks. After the treatment, they asked students to write a response paragraph to the graphs provided as in the IELTS examination. The three outputs of each student, which are 20 minutes timed paragraphs were analysed and the results showed that the tokens of the formulaic language items used, increased in the post-test and decreased in the delayed post-test; however, it was higher than the pre-test, the same as in the present paper. Similarly, the participants involved in the study of Alhassan and Wood (2015), compared to their pre-test, successfully used different types of the target formulaic language items in the post and delayed post-test instead of repeating them over and over. Moreover, between the post-test and the delayed post-test, they did not find any significant difference, either.

Comparison of the use of the target formulaic language items by experimental group and control group also supports that the explicit teaching might be promising and a good opportunity for the students to make use of formulaic language in their writing. To illustrate, in each essay the experimental group students used more types of the target formulaic language items instead of relying the same items and more tokens of the target formulaic language items than the students in the control group who were not exposed to explicit teaching. The results can be partially linked to the study of Peters and Pauwels (2015), because they did not have two treatments group at the beginning of their study. However, at the end of the term they decided that it was worth to compare the end of year assignment of the participants involved in their study to the ones of a class of students who was not a part of their study at the beginning. They found that the students receiving treatment used much more formulaic language items, which is similar to the findings of the present study. Due to the design of their study, Peters and Pauwels (2015) did not analyse the gain, if any, of the students not receiving treatment, for each test, but the present paper also indicated that the control group also increased their use of the target formulaic language items in small numbers; however, this increase was never as remarkable as the experimental group did. There might be many reasons of this modest increase such as their prior knowledge, peer learning, and unconscious exposure during the courses or in their social life while reading, watching or listening. The inferential statistics, supportively, indicated that there was a significant difference between the experimental and control group, regarding the type and tokens of the target formulaic language items used in the post-test and the final test conducted after the treatment, but not for the pre-test which was before the treatment and for the delayed post-test which was conducted one month after the final test. Taken together, these statistics could be accepted as an indication of the fact that the students gained

much improvement in the use of the target formulaic language items through the explicit instruction, but in the long run there might be some decrease in the tokens unless the explicit instruction was provided regularly.

5.1.3. The use of formulaic language and the overall argumentative quality of the writing

As there might a possibility for the scores given to the overall argumentative quality of the essays to be affected by the use of the target formulaic language items, a correlation test was conducted and the results indicated that there was a strong relationship between the tokens of the target formulaic language used and the overall argumentative quality scores; however, there was not a relationship between the overall argumentative quality scores of the essays and the type of the target formulaic language used. Although the raters did not receive any training and there was no instruction about the formulaic language in the rubric they were given to score the overall argumentative quality, it seems that they tend to score higher when greater number of formulaic language items are used in the essay. In other words, the number of the formulaic language items used in the essays caused the raters to score higher for the overall argumentative quality scores of the essays. The results are partially compatible with the study of Alhassan and Wood (2015). They analysed each rater individually in their study and resulted in the evaluation of two raters for the overall quality was strongly correlated with the variety and the repetition of the formulaic language items, but not the third rater.

5.1.4. The organisational template and thesis statement structures

As mentioned earlier, the experimental group students were also given an organisational template and three thesis statement structures.

5.1.4.1. The organisational template

The analyses showed that the explicit teaching of the organisational template affected the organisation of the essays in a positive way since the percentage of the experimental group students following the recommended template gradually increased after the treatment while it declined steadily for the control group. Kaminski, et al. (1993), similarly, found that after the explicit teaching of organisational structures to the fourth grade students, the number of the students who wrote their composition well organized and more sophisticated was higher than the ones as of the control group. While the explicit teaching of organisational structures was provided through a template in the present paper, Kaminski et al. (1993) taught the structure using graphics during ninety minutes instruction. Another study was conducted by Miri (2014) who used different approaches to teach the organisation of an argumentative paper during five-week explicit instruction. Her findings showed that the approach in which students were provided with models of argumentative papers and explained the ways of expressing opposing and defending points for five weeks, helped students write better quality argumentative papers. The result of the present paper that the percentage of the students following the recommended template decreased for both groups in the delayed post-test might have been due to the fact that they did not receive the explicit teaching during the time until the delayed post-test.

The students were also observed to have used all the items of the recommended templates but in a different order. The results were quite varied for both the experimental group and the control group. After the treatment there was a fall in the percentage of the experimental group students following this different order, then in the final test it rose again and in the delayed post-test it declined. There might have been many factors causing this rise and fall in the percentages such as an effort to internalize the new knowledge or ignorance of it. Moreover, the students might have been using prior learnings, or unaware of the importance of such a template to produce better organised and more persuasive argumentative papers. The control group students' following this different order neither increased nor decreased in the post-test, but surprisingly in the final more than half of the students in the group employed this new order, but the percentage fell in the delayed post-test. As the control group did not receive any explicit teaching about the organisational template, they were simply believed to make use of their previous learning or write their argumentative papers by not focussing on argumentation aspect.

Analyses also revealed that some students had, in their argumentative papers, at least one missing element as of those recommended for the organisational template such as opposing view, the reasons to oppose the view, refutation of the opposing view and own view with supporting details. The results indicated that the mostly missed elements were mainly related to the opposing points for both of the groups, which were similar to the findings of the studies conducted by Qin and Karabacak (2010) and Rusfandi (2015). The participants of both studies had involved in academic writing classes before and for the mentioned studies they did not receive any writing instruction or given a template as it was in the present study, but asked to write argumentative papers on the

given topics. Rusfandi (2015) stated the reasons of this lack of counter-argument as low proficiency in English and the possibility to be perceived as rude in Indonesian culture while opposing someone. However, the argument of language proficiency was refuted by the analysis of the essays written by the same students in Indonesian language, which also were deprived of counter-arguments, and also by the present paper since the participants of this paper were all from ELT department which required certain high level language proficiency. Another finding of this paper worth to discuss was that the experimental group student was observed to use counter-arguments more in their essays after the treatment; however, in the delayed post-test missing of counter-arguments increased. This might be related to the cognitive demand required to write counter-argument statements as highlighted in the study of Qin and Karabacak (2010) and the necessity of the explicit teaching proposed as in the study of Gersten and Baker (2001) who analysed many intervention studies conducted in the field of teaching writing to learners with disabilities and Ka-kan-dee and Ka-ur (2015) who interviewed with writing EFL lecturers for the strategies of teaching argumentative writing.

5.1.4.2. The thesis statement structures

The results also indicated that the explicit teaching of the thesis statement structures improved the organisational structure of the essays. For instance, before the intervention 52.3% of the experimental group students structured the thesis statements in a way that is not in compliance with any of the three recommended structures. Miller and Pesoa (2016), who analysed the essays written by the students in a history course which required the students to write argumentative essays to the prompts, similarly, found that the students had challenges in writing thesis statements in argumentative writing. Although the participants of their study did not receive any instruction during

the semester, they improved their thesis statement writing through exposure to reading materials; however, this improvement was not observed among all students. That's why they emphasized the importance of explicit teaching to foster the writing development. The importance of explicit teaching was also supported by the present paper, because the percentage of the students who wrote a thesis statement rather than the recommended structures increased in the delayed post-test in which twenty-six students participated and it was conducted approximately six weeks after the treatment was over. Moreover, the percentage of employing the recommended thesis statement structures was higher for the experimental group students in each test compared to the control group students.

Another indication of the improvement was that the experimental group students formed a new structure of the thesis statement rather than the recommended ones after the treatment in the post-test, the final test and the delayed post-test, by mixing the second and third category. However, such a structure was never observed with the essays of the control group students for any of the tests. This new mix category was believed to emerge as a result of internalizing the structures recommended.

Lastly, the percentage of the students who did not write a thesis statement at all is lower in the experimental group compared to the ones in the control group, which is partially in line with the study of Alhassan and Wood (2015) who conducted a content analysis on the introductory and concluding sentences of the paragraphs written by the participants. Alhassan and Wood (2015) provided model introductory and concluding sentences to the students during the training, different but similar to the training in the present paper, which provided three thesis statement structures instead of models. They found that the students not writing a thesis statement in the pre-test, wrote a similar

thesis statement to the model provided during the training. Another study conducted by Owusu and Adade-Yeboah (2014), analysing the thesis statement of business students in their different genre papers also revealed that 87% of the papers did not include a thesis statement. Taken together, one might argue that writing thesis statements to written productions is important for the organisation and the quality of the written work and providing models and organisational structures through explicit teaching may have positive effect on the organisation and the quality of the writing as well.

5.2. Implication of the study

The present study, focussing on the importance of formulaic language, sought out the effects of the explicit teaching of formulaic language on the academic writing, specifically argumentative writing. Based on the literature and the findings of the present study, several pedagogical implications were proposed to provide insights into the explicit teaching of formulaic language.

First of all, the present paper indicated that more focus should be given and more efforts should be dedicated to the formulaic language since it plays a vital role in academic writing as the formulaic items serve specific functions and nativelikeness. Thus, language teachers should take some actions in that sense. First and foremost action that a language teacher should take is to raise awareness of the students about the frequency, use and functions of the formulaic language as it is generally lacking in salience in the input. It should be noted that raising awareness should be supported with examples and activities by employing some techniques instead of just explaining how prevalent formulaic language items are and what their functions are. It is also essential

for the teachers to have background knowledge such as Schmidt's Noticing Hypothesis to use raising awareness techniques (Schmidt, 2001).

Another important implication that can be drawn from the present study is that explicit teaching is effective for the students to improve their use of formulaic language and the overall argumentative quality of their writings. Thus, the language teachers who want to foster the formulaic language use and the overall argumentative quality of the essays of the students may wish to consult to explicit teaching. However, they should be meticulous while employing explicit teaching by following some steps such as noticing, retrieval and generation which are the tree principles of vocabulary learning (Nation, 2001). To illustrate, teachers should compile a reference list empirically, for instance by using concordancers and according to the aims of the course, then provide different materials in which the formulaic language items are made noticeable and provide activities in which the students will encounter the formulaic language several times and find opportunities to exercise. In this way, it is believed to be more helpful in some ways rather than implicit instruction. For instance, having an impression of what is learnt, or even whether it is learnt, may take long time, when formulaic language, which is such a difficult element of language learning due to its transparency, plethora and varied structures, is aimed to be taught implicitly only through exposure. Moreover, students may feel more confident to use formulaic language when they are taught explicitly, because through implicit instruction they may not comprehend the crucial functions such as idiomaticity and discourse functions of formulaic language and so they do not want to take the risk of making mistakes by using it.

The finding that decrease in the tokens of the target formulaic language items used and in the overall argumentative quality scores of the essays in the delayed post-

test, also revealed the importance of the continuity of the explicit instruction. Thus, teaching formulaic language could be integrated to the writing classes, especially to the genre-based classes, as they have distinctive characteristics across genres (Hyland, 2008a; Ellis et al., 2008). For instance, as in the present paper, if the students are required to write an argumentative paper, then the formulaic language items which function to discuss contrasting ideas, defend own position, provide examples and so on, should be chosen and associated with the organisational structure of argumentative writing. Another finding of the present paper supporting the integration of the formulaic language teaching to the writing classes is the positive correlation found between the tokens of the target formulaic language items and the overall quality scores of the essays. That is to say that the increase in the use of formulaic language items in the essays caused an increase in the overall argumentative quality scores of the essays.

Although empirical interview sessions were not conducted in the present study, the class discussions provided insights about the possible individual differences such as language aptitude, language learning styles, and motivation and so on. Thus, a final implication might be that teachers should keep in mind these differences and try to provide a great variety of materials such as listening scripts, videos, concordancing programmes, games, books and academic papers rich in formulaic language so that each student can employ the one appealing him/her. Moreover, it could be effective to motivate the students not only with curricular activities but also extra-curricular activities such as introducing them exchange programmes, summer schools and projects by which the students can directly get in touch with native speakers and providing native utterances since some researchers (Dörnyei et al., 2004; Ortaçtepe, 2013; Li & Schmitt, 2009) found strong relationship between the success in acquiring

formulaic language and students' active involvement in social communities in which the target language is spoken natively.

All in all, the findings of the present study demonstrated that utilising the explicit teaching of formulaic language might be promising, so as to foster formulaic language learning and improve the overall quality of the writing. Thus, it can lead all the stakeholders having a role in the process of teaching to draw a number of conclusions pertaining to the explicit teaching of formulaic language.

5.3. Limitations of the study

There are some certain limitations relevant to this study such as time constraints and absenteeism of the participants. Before proceeding to limitations, it should be noted that this classroom based research was primarily concerned with investigating the effects of the explicit teaching of formulaic language on the academic writing of the university students which were restricted to freshmen and sophomores of English Language Department at Yeditepe University. Thus, the results should be considered tentatively.

The first limitation is concerned with time constraints, that is, the experiment and data gathering had to take place in during the academic year which means that there was a curriculum which had to be implemented until certain deadlines for the exams. Thus, only four-hour intervention during two weeks could be allocated to the study, which was not ample. Although the results showed significant differences after the treatment, if more time could have been allocated for the explicit instruction, it could be much more possible to discuss the results unwaveringly. Another limitation revealed because of time constraint is pertaining to the methodology. For instance, the target

formulaic language items, in accordance with the participants' needs, were gathered intuitively by the instructor of the experimental group from the materials including lists which were comprised of through corpus methodologies. Instead a more detailed corpus study could have been carried out to select the target items; however, it would be a new research study on its own and there would be no time for the present study. Moreover, due to time constraints the target items had to be limited to forty although a hundred items had been chosen initially. This might have caused the raters to perceive the paper as memorized clichés and ordinary since they saw the same formulaic language items over and over. The limited number of formulaic language items also might have restricted the creativity of the students. They might have only focussed on the target items instead of using similar formulaic sequences or generated utterances. Thus, if more formulaic language items had been targeted, the findings might have been resulted in different ways. Final limitation caused by the time constraints was that a small number of type activities were employed. In order to appeal the attention of more students, more varied types of activities could have been employed.

Another limitation of the study was that the research was conducted as a classroom based study with the participation of 71 ELT students at Yeditepe University. Some data could not be included in the study since not all participants were available on the day to gather the data for the study. The participants were supposed to write four essays in total as a pre-test, a post-test, a final test and a delayed post- test. However, some students were absent during the delayed post-test. Thus, only the participants who wrote all of the essays were included in the analysis of some research questions linked with the delayed post-test. Although the researcher could gather sufficient data to carry out statistical tests, the generalization of the findings might not be possible due to the

number of the participants. Including greater number of participants from different majors even from different institutions can help to draw more reliable and generalizable results.

Lastly, the essays written at different points during the study had different weights on students' grades pertaining to the course evaluation. This might have affected the motivation of the student and therefore the overall quality of the essays and the results of the study. Thus in order to control this kind of external factors, one might weight the assignment or test equal and take some other cautions for the external factors.

5.4. Recommendation for future research

The formulaic language has been found to be worth investigating by many researchers for considerable years, due to the benefits of it, yet experimental research with formulaic language is still in its infancy and still draws attention of researchers. Thus, this study has humbly proposed several recommendations as a guide to other researchers who would like to delve into the similar problems in greater detail.

First, this study is limited to the sample population selected from freshman and sophomore classes at the Department of English Language Teaching of Yeditepe University which is a private institution in İstanbul, Turkey. In further studies, more students from different classes and from different universities, private and state, even from other cities, can be involved so as to reach more generalizable results. Moreover, the participants can be chosen from other backgrounds such as academicians, native speakers, bilingual or trilingual speakers etc. to make a contrastive analysis.

The experimental and control group which were constituted of freshman and sophomore classes, were already taught academic writing in their previous years. Thus, they may have some knowledge of formulaic language and academic writing in advance which might affect the results of the study. For further studies, the students who will meet academic writing for the first time in their academic life, for instance the preparatory class students, can be selected, thus, any doubt of having previous knowledge of academic writing might be eliminated.

In the current study, the demographic features of the students were not taken into consideration; however, to mind them would be beneficial for further studies since they may affect the results of the study. For instance, some students might have been abroad before and exposed to the taught items, some may have foreign parents or acquaintances from whom they already learned the target items, some may keen on reading books, watching movies/series or listening to music in English which provide rich input in formulaic language and some students might have goals of having an academic career which may motivate them more to learn formulaic language. Thus, in further researchers, these may be taken into consideration while interpreting the results.

The study explored the effects of the explicit teaching of formulaic language on academic writing, specifically argumentative writing. The same can be replicated for other types of academic writing such as comparative writing and research papers or other fields in social sciences or natural sciences. Thus, the effects of explicit teaching on different fields or genres can be better illustrated.

Lastly and more importantly, the time allocated for the instruction was not much for the study as there was a syllabus to be followed. A longitudinal study, in which the

formulaic language is integrated to the course syllabus and instruction is focussed more with varied exercises, can be conducted, so as to see the long-term effects of the explicit teaching of formulaic language on academic writing. Moreover, the reflection of the students can be acquired in a more systematic and empirical way like using regular reports, interviewing or a survey.

5.5. Conclusion

The main purpose of this study was to find out the effects of the explicit teaching of formulaic language on academic writing, specifically on argumentative writing.

The results of the descriptive and inferential analyses revealed significant differences between the experimental and the control group, regarding both the overall quality scores of the essays and the target formulaic language items used. The students receiving explicit teaching, compared to their counterparts, wrote better quality essays, greater number and more varied formulaic language items in their writing after the treatment. However, this short-term intervention did not lead to an improvement on the overall quality of the students' later production. Even though the students' essays did not receive higher scores in the long run, their use of formulaic language items were relatively remarkable.

The trials of the students on using formulaic language items and the organisational template and the thesis statement structures recommended, indicated that the explicit teaching provided them an opportunity to practice formulaic language items and these structures, thus they internalised both in short term. Moreover, a positive correlation was revealed by the qualitative analysis between the number of the formulaic language items used and the overall quality scores of the essays. Thus, in

order to write better quality essays, formulaic language items might be utilised especially as discourse organisers.

All in all, both the qualitative and the quantitative analyses demonstrated that the awareness of the students was raised through explicit instruction which resulted in better quality essays written and greater number and more varied of formulaic language items utilised after the treatment. Without instructional intervention, the students might not have noticed formulaic language items in such a short time due to the characteristics of formulaic language such as transparent semantics, incomplete structure and irregular syntax. Thus, the present study is believed to be a promising initiative to utilise explicit teaching in order to teach formulaic language items.

REFERENCES

Ädel, A., Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes*, *31*, 81-92.

Ahmadi, H. S., Ghonsooly, B., Fatemi, A. H. (2013). An analysis of lexical bundles in research article abstracts by Iranian and native English-speaking authors of applied linguistics articles. *The Asian ESP Journal, 9 (1)*, 5-27.

Alali, F. A., Schmitt, N. (2012). Teaching formulaic sequences: The same as or different from teaching single words? *TESOL Journal, 3 (2)*, 153-180. DOI: 10.1002/tesj.13

AlHassan, L., Wood, D. (2015). The effectiveness of focused instruction of formulaic sequences in augmenting L2 learners' academic writing skills: A quantitative research study. *Journal of English for Academic Purposes, 17*, 51-62. http://dx.doiorg/10.1016/j.jeap.2015.02.001

Alipour, M., Zarea, M. (2013). A disciplinary study of lexical bundles: The case of native versus non-native corpora. Taiwan International ESP Journal, 5 (2), 1-20.

Altenberg, B., Granger, S. (2001). The grammatical and lexical patterning of make in native and non-native student writing. *Applied Linguistics, 22 (2),* 173-195.

Arnon, I., Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62,* 67-82. DOI:10.1016/j.jml.2009.09.005

Bannard, c., Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science, 19 (3),* 241-248.

Bardovi-Harlig, K. (2009). Conventional expressions as a pragmalinguistic resource: Recognition and production of conventional expressions in L2 pragmatics. *Language Learning, 59 (4),* 755-795.

Biber, D., Barbieri, F. (2007). Lexical bundles in university spoken and written registers. *English for Specific Purposes*, 26, 263-286. DOI:10.1016/j.esp.2006.08.003

Biber, D., Conrad, S., Cortes, V. (2004). If you look at … : Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25 (3),* 371-405.

Bishop, H. (2004a). The effect of typographic salience on the look up and comprehension of unknown formulaic sequences. In N. Schmitt (Ed.), *Formulaic Sequences: Acquisition, Processing and Use* (pp. 227-248). Amsterdam/Philadelphia: John Benjamins Publishing.

Bishop, H. (2004b). Noticing sequences-a problem of measuring the subjective. *LSO Working Papers in Linguistics 4,* 15-19.

Boers, F., Eychmans, J., Kappel, J., Stengers, H., Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency. Putting a lexical approach to the test. *Language Teaching Research, 10 (3),* 245-261. DOI: 10.1191/1362168806lr195oa

Boers, F., Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics, 32,* 83-110. DOI: 10.1017/S0267190512000050

Bogartz, R.S. (2010). *Interrater agreement and combining ratings* [PDF document]. Retrieved from http://people.umass.edu/~bogartz/Interrater%20Agreement.pdf

Bonk, W. J. (2001). Testing ESL learners' knowledge of collocations. In T. Hudson, J. D. Brown (Eds.), *A focus on language test development: expanding the language proficiency construct across a variety of tests*. (pp. 113- 142). Honolulu: University of Hawaii, Second Language Teaching and Curriculum Center.

Byrd, P., Coxhead, A. (2010). On the other hand: Lexical bundles in academic writing in the teaching of EAP. *University of Sydney Papers in TESOL, 5*, 31-64.

Chan, T., Liou, H. (2005). Effects of web-based concordancing instruction on EFL students' learning of verb-noun collocations. *Computer Assisted Language Learning, 18 (3)*, 231-251. DOI: 10.1080/09588220500185769

Chen, L. (2008). *An investigation of lexical bundles electrical engineering introductory textbooks and ESP textbooks* (Unpublished Master of Arts dissertation). Carleton University, Canada.

Chen, Y., Baker, P. (2010). Lexical bundles in L1 and L2 academic writing. *Language Learning & Technology, 14 (2),* 30-49.

Čolović-Marković, J. (2012). *The effects of explicit instruction of formulaic sequences on second-language writers* (Unpublished Doctor of Philosophy dissertation). The University of Utah: the United States of America.

Conklin, K., Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and non-native speakers? *Applied Linguistics, 29 (1)*, 72-89. DOI:10.1093/applin/amm022

Conklin, K., Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics, 32*, 45-61. DOI: 10.1017/S0267190512000074

Conrad, S. M., Biber, D. (2005). The frequency and use of lexical bundles in conversation and academic prose [PDF document]. Retrieved from Applied Linguistics Faculty Publications and Presentations: http://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1002&context=ling_fac

Cortes, V. (2002). *Lexical bundles in academic writing in History and Biology* (Unpublished Doctor of Philosophy dissertation). Northern Arizona University, the United States of America.

Cortes, V. (2004). Lexical bundles in academic writing in History and Biology. *English for Specific Purposes, 23*, 397-423. DOI:10.1016/j.esp.2003.12.001

Cortes, V. (2006). Teaching lexical bundles in the disciplines: An example from a writing intensive history class. *Linguistics and Education, 17*, 391-406. DOI:10.1016/j.linged.2007.02.001

DeCock, S. (2000). Repetitive phrasal chunkiness and advanced EFL speech and writing. In C. Mair & M. Hundt (Eds.), *Corpus linguistics and linguistic theory* (pp. 51−68). Amsterdam: Rodopi.

DeCock, S. (2004). Preferred sequences of words in NS and NNS speech. *BELL : Belgian journal of English language and literatures, 2*, 225-246

Dickinson, P. (2012). Improving second language academic presentations with formulaic sequences. *Niigata University of International and Information studies (*新潟国際情報大学情報文化学部紀要*), 15, 25-36.*

Dörnyei, Z., Durow, V., Zahran, K. (2004). Inividual differences and their effects on formulaic sequence acquisition. In N. Schmitt (Ed.), In N. Schmitt (Ed.), *Formulaic Sequences: Acquisition, Processing and Use* (pp. 87-106). Amsterdam/Philadelphia: John Benjamins Publishing.

Durrant, P. (2008). *High frequency collocations and second language learning* (Unpublished Doctor of Philosophy dissertation). University of Nottingham, the United Kingdom.

Durrant, P., Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *IRAL - International Review of Applied Linguistics in Language Teaching, 47 (2)*, 157–177, DOI: 10.1515/iral.2009.007

Durrant, P., Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research, 26 (2),* 163-188. DOI: 10.1177/0267658309349431

Ellis, N. C. (2012). Formulaic language and second language acquisition: Zipf and the phrasal teddy bear. *Annual Review of Applied Linguistics, 32,* 17-44. DOI: 10.1017/S0267190512000025

Ellis, N. C., Simpson-Vlach, R., Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics and TESOL. *TESOL Quarterly, 42 (3),* 375-396.

Eriksson, Andreas. 2012. 'Pedagogical perspectives on bundles: Teaching bundles to doctoral students of biochemistry'. In J. Thomas & A. Boulton (Eds.), *Input, Process and product: Developments in teaching and language corpora* (pp. 195-211). Brno: Masaryk University Press.

Erman, B., Warren, B. (2000). The idiom principle and the open choice principle. *Text - Interdisciplinary Journal for the Study of Discourse, 20 (1)*, 29-62.

Foster, P. (2001). Rules and routines: A consideration of their role in the task based language production of native and non-native speakers. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 75-93). Harlow: Pearson Education Limited.

Gersten, R., Baker, S. (2001). Teaching expressive writing to students with learning disabilities: a meta-analysis. *The Elementary School Journal, 101 (3),* 251-272.

Graham, M., Milanowski, A., Miller, J. (2012). *Measuring and promoting inter-rater agreement of teacher and principal performance ratings*. The Center for Educator and Compensation Reform (CECR) [PDF document] Retrieved from http://files.eric.ed.gov/fulltext/ED532068.pdf.

Graham, D., Osment, C. (2014). Interactive web-based learning of corpus generated phrases. *AsiaCALL Online Journal, 9,* 1-13.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). Multivariate data analysis (6 ed.). New Jersey: Pearson Prentice Hall.

Henriksen, B. (2013). Research on L2 learners' collocational competence and development – a progress report. In C. Bardel, B. Laufer & C. Lindqvist (Eds.), *L2 vocabulary acquisition, knowledge and use* (pp. 29-56). Erosla.

Herbel-Eisenmann, B., Wagner, D. &Cortes, V. (2010). Lexical bundle analysis in mathematics classroom discourse: the significance of stance. *Educational Studies in Mathematics, 75 (1)*, 23-42. DOI 10.1007/s10649-010-9253-6

Hyland, K. (2008a). Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics, 18 (1),* 41-62.

Hyland, K. (2008b). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes, 27,* 4-21. DOI:10.1016/j.esp.2007.06.001

Jablonkai, R. (2009). "In the light of": A corpus-based analysis of lexical bundles in two EU-related registers. *WoPalp, 3*, 1-27.

Jalalai, H., Ghayoomi, S. (2010). A comparative qualitative study of lexical bundles in three academic genres of applied linguistics. *MJAL, 2 (4),* 323-333.

Jiang, N., Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal, 91 (3),* 433-445.

Johnson, R.L., Penny, J.A. Gordon, B. (2008). *Assessing Performance*. New York NY: Guilford Press.

Jones, M., Haywood, S. (2004). Facilitating the acquisition of formulaic sequences: an exploratory study in an EAP context. In N. Schmitt (Ed.), *Formulaic Sequences: Acquisition, Processing and Use* (pp. 269-292). Amsterdam/Philadelphia: John Benjamins Publishing.

Ka-kan-dee, M., Kaur, S. (2015). Teaching strategies used by Thai EFL lecturers to teach argumentative writing. *Procedia- Social and Behavioral Sciences, 208,* 143-156. DOI: 10.1016/j.sbspro.2015.11.191

Kaminski, R., Bean, L., Lazar, B. (1993). *Students' ability to apply and reflect on organizational structures used in composing*. Paper presented at a Joint Table at the 4rd Annual meeting of the National Reading Conference. Charleston, SC.

Kecskes, I., (2007). Formulaic language in English Lingua Franca. In I. Kecskes & L. R. Horn, (Eds.), *Explorations in Pragmatics* (pp. 191-218). Germany: Mouton de Gruyter. DOI (Chapter): 10.1515/9783110198843.3.191

Kehoe, J. (1995). Basic item analysis for multiple-choice tests. P*ractical Assessment, Research & Evaluation, 4 (10)*. Available online. http://files.eric.ed.gov/fulltext/ED398237.pdf

Kennedy, C., & Thorp, D. (2007). A corpus investigation of linguistic responses to an IELTS Academic Writing task. In L. Taylor & P. Falvey (Eds.), *IELTS collected paper: research in speaking and writing assessment* (pp. 316–378). Cambridge: Cambridge University Press.

Kuiper, K. (2004). Formulaic performance in conventionalised varieties of speech. In N. Schmitt (Ed.), *Formulaic Sequences: Acquisition, Processing and Use* (pp. 37–54). Amsterdam/Philadelphia: John Benjamins Publishing.

Lance, C.E., Butts, m.M., Michels, L.C. (2006) The sources of four commonly reported cutoff criteria: what did they really say? *Organizational Research Methods, 9 (2),* 202-220. DOI: 10.1177/109448105284919

Laufer, B, Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning, 61 (2)*, 647-672. DOI: 10.1111/j.1467-9922.2010.00621.x

Laufer, G., Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics, 29 (4)*, 694-716. DOI:10.1093/applin/amn018

Lee, D. Y. W. (2001). Genres, registers, text types, domains, and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning & Technology, 5 (3),* 37-72. http://llt.msu.edu/vol5num3/lee/

Li, J., Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing, 18*, 85-102. DOI:10.1016/j.jslw.2009.02.001

Lin, P. M. S. (2013). Investigating the validity of internet television as a resource for acquiring L2 formulaic sequences. *System, 42*, 164-176. http://dx.doiorg/10.1016/j.system.2013.11.010

Liu, D. (2011). The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes, 31*, 25-35. DOI:10.1016/j.esp.2011.07.002

Liu, W., Huo, Y. (2011). On the role of formulaic sequences in second language acquisition. *US-China Foreign Language, 9 (11)*, 708-713.

Martinez, R., Schmitt, N. (2012). A phrasal expression list. *Applied Linguistics, 33 (3)*, 299-320. DOI:10.1093/applin/ams010

McAllister, D.J., Bigley, G.A. (2002). Work context and the definition of self: how organizational care influences organization-based self-esteem. *Academy of Management Journal, 45*, 894-904.

Meunier, F. (2012). Formulaic language and language teaching. Annual Review of *Applied Linguistics, 32*, 111-129. DOI: 10.1017/S0267190512000128

Millar, N. (2011). The processing of malformed formulaic language. *Applied Linguistics, 32 (2)* , 129-148. DOI:10.1093/applin/amq035

Miller, R.T., Pessoa, S. (2016). Where's your thesis statement and what happened to your topic sentences? Identifying organizational challenges in undergraduate student argumentative writing. *TESOL Journal, 7 (4),* 847-873.

Miri, T. (2014). Toward Finding an Approach for Improving Rhetorical Organization of EFL Learners' Argumentative Writing. *International Journal of Applied Linguistics & English Literature, 3 (3),* 164-170. DOI:10.7575/aiac.ijalel.v.3n.3p.164

Nekrasova, T. M. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Language learning, 59 (3)*, 647-686.

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

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied linguistics, 24 (2)*, 223-242.

Ohlrogge A (2009). Formulaic expression in intermediate EFL writing assessment. In, R. Corrigan, E. Moravcisk, H. Ouali, K. Wheatley (Eds.), *Formulaic Language: Acquisition, Loss, Psychological Reality, and Functional Explanations*, (pp. 375-386) Philadelphia: John Benjamins Publishing Company.

Ortaçtepe, D. (2013). Formulaic language and conceptual socialization: The route to becoming nativelike in L2. *System, 41*, 852-865. http://dx.doiorg/10.1016/j.system.2013.08.006

Owusu, E., Adade-Yeboah, A. (2014). Thesis statement: a vital in expository essays. *Journal of Langugae Teaching and Research, 5 (1),* 56-62. DOI:10.4304/jltr.5.1.56-62

Pang, W. (2010). Lexical bundles and the construction of an academic voice: A pedagogical perspective. *Asian EFL Journal, 47*, 30-43.

Penny, J., Johnson, R.L., Gordon, B. (2000). Using rating augmentation to expand the scale of an analytical rubric. *The Journal of Experimental Education, 68 (3)*, 269-287.

Peters, E. (2012). Learning German formulaic sequences: The effect of two attention-drawing techniques. *The language Learning Journal, 40 (1)*, 65-79. http://dx.doiorg/10.1080/09571736.2012.658224

Peters, E. (2014). The effects of repetition and time of post-test administration on EFL learners' form recall of single words and collocations. *Language Teaching Research, 18 (1)*, 75-94. DOI: 10.1177/1362168813505384

Peters, E., Pauwels, P. (2015). Learning academic formulaic sequences. *Journal of English for Academic Purposes, 20*, 28-39. http://dx.doiorg/10.1016/j.jeap.2015.04.002

Ping, P. (2009). A study on the use of four-word lexical bundles in argumentative essays by Chinese English-majors-a comparative study based on WECCL and LOCNESS. *CELEA* (中国英语教), *32 (3)*, 25-45.

Qi, Y., Ding, Y. (2011). Use of formulaic sequences in monologues of Chinese EFL learners. *System, 39*, 164-174. DOI:10.1016/j.system.2011.02.003

Qin, J., Karabacak, E. (2010). The analysis of of Toulmin elements in Chinese EFL university argumentative writing. *System*, *38*, 444-46. DOI:10.1016/j.system.2010.06.012

Read, J., Nation, P. (2004). Measurement of formulaic sequences. In N. Schmitt (Ed.), *Formulaic Sequences: Acquisition, Processing and Use* (pp. 23-35). Amsterdam/Philadelphia: John Benjamins Publishing.

Read, J., Nation, P. (2006). An investigation of the lexical dimension of the IELTS speaking test. *International English Language Testing System (IELTS) Report (6)*, 1-25.

Rothbard, N.P., Edwards, J.R. (2003). Investment in work and family roles: a test of identity and utilitarian motives. *Personnel Psychology, 56*, 699-729.

Römer, U., Arbor, A. (2009). English in academia: Does nativeness matter? *Anglistik: International Journal of English Studies, 20 (2)*, 89–100.

Rott, S. (2009). The effects of awareness-raising on the use of formulaic constructions. In R. Corrigan, E. A. Moravcsik, h. Ouali & K. M. Wheatley (Eds.), *Formulaic Language: Volume 2: Acquisition, Loss, Psychological Reality, and Functional*

Explanations (pp.405-422). Amsterdam/Philadelphia: John Benjamins Publishing.

Rusfandi, R. (2015). Argument-counterargument structure in Indonesian EFL learners' English argumentative essay: a dialogic concept of writing. *RELC Journal, 46 (2),* 181-197. DOI: 10.1177/0033688215587607

Schmidt, R. W. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction*. (pp.3-32) Cambridge: Cambridge University Press.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8 (4),* 350-353. DOI: 10.1037/1040-3590.8.4.350

Schmitt, N., Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic Sequences: Acquisition, Processing and Use* (pp. 1-22). Amsterdam/Philadelphia: John Benjamins Publishing.

Simpson-Vlach, R., Ellis, N. C. (2010). An academic formula list: New methods in phraseology research. *Applied Linguistics, 31 (4)*, 487-512.

Siyanova, A., Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *The Canadian Modern Language Review, 64 (3)*, 429-458.

Siyanova-Chanturia, A., Conklin, K., Schmitt, N. (2011a). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research, 27 (2)*, 251-272. DOI: 10.1177/0267658310382068

Siyanova-Chanturia, A., Conklin, Van Heuven, W. J. B. (2011b). Seeing a phrase "time and again" matters: the role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory and Cognition, 37 (3)*, 776-784. DOI: 10.1037/a0022531

Staples, S., Egbert, J., Biber, D., McClair, A. (2013). Formulaic sequences and EAP writing development: Lexical bundles in the TOEFL IBT writing section. *Journal of English for Academic Purposes*, 12, 214–225. http://dx.doiorg/10.1016/j.jeap.2013.05.002

Stemler, S.E., Tsai, J. (2008). Best practices in interrater reliability three common approaches. In J.W. Osborne (Ed.), *Best Practices in Quantitative Methods* (pp. 29-49). Thousand Oaks, California: Sage Publications Ltd. DOI: 10.4135/9781412995627.d5

Sun, Y., Wang, L. (2003). Concordancers in the EFL classroom: Cognitive approaches and collocation difficulty. *Computer Assisted Language Learning, 16 (1)*, 83-94. http://dx.doiorg/10.1076/call.16.1.83.15528

Szudarski, P, Conklin, K. (2014). Short- and long-term effects of rote rehearsal on ESL learners' processing of L2 collocations. *TESOL Quarterly, 48 (4)*, 833-842. DOI: 10.1002/tesq.201

Szudarski, P., Carter, R. (2014). The role of input flood and input enhancement in EFL learners' acquisition of collocations. *International Journal of Applied Linguistics, 26 (2)*, 245-265. DOI: 10.1111/ijal.12092

Tremblay, A., Derwing, B. Libben, G. Westbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning, 21 (2)*, 569-613. DOI: 10.1111/j.1467-9922.2010.00622.x

Trevethan, R. (2016). Intraclass correlation coefficients: clearing the air, extending some cautions, and making some requests. *Health Services and Outcomes Research Method*, 1-17. DOI: 10.1007/s10742-016-0156-6

Underwood, G., Schmitt, N. Galphin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic Sequences: Acquisition, Processing and Use* (pp. 153-172). Amsterdam/Philadelphia: John Benjamins Publishing.

Valsecchi, M., Künstler, V., Saage, S., White, B. J., Mukherjee, J., Gegenfurtner, K. R. (2013). Advantage in reading lexical bundles is reduced in non-native speakers. *Journal of Eye Movement Research, 6 (5)*, 1-15.

Van Lancker-Sidtis, D., Rallon, G. (2004). Tracking the incidence of formulaic expressions in everyday speech: methods for classification and verification. *Language & Communication, 24,* 207–240. DOI:10.1016/j.langcom.2004.02.003

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

Webb, S., Kagimoto, E. (2009). The Effects of Vocabulary Learning on Collocation and Meaning. *TESOL Quarterly, 43 (1)*, 55-77.

Webb, S., Newton, J. Chang, A. (2013). Incidental learning of collocation. *Language Learning, 63 (1),* pp. 91–120. DOI: 10.1111/j.1467-9922.2012.00729.x

Wei, Y., Lei, L. (2011). Lexical Bundles in the Academic Writing of Advanced Chinese EFL Learners. *RELC Journal, 42 (2)*, 155–166. DOI: 10.1177/0033688211407295

Wood, D. (2002). Formulaic language in acquisition and production: implications for teaching. *TESL Canada Journal, 20 (1)*, 1-15. DOI: http://dx.doiorg/10.18806/tesl.v20i1.935

Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *The Canadian Modern Language Review, 63 (1),* 13-33.

Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics, 12 (1)*, 39-57.

Wood, D., Appel, R. (2014). Multiword constructions in first year business and engineering university textbooks and EAP textbooks. *Journal of English for Academic Purposes, 15*, 1-13. http://dx.doiorg/10.1016/j.jeap.2014.03.002

Wray, A. (1999). Formulaic language in learners and native speakers. *Language Teaching, 32 (4)*, 213-231. http://dx.doiorg/10.1017/S0261444800014154

Wray, A. (2000). Formulaic sequences in second language teaching: principle and practice. *Applied Linguistics, 21 (4)*, 463-489.

Wray, A. (2004). Here's one I prepared earlier: Formulaic language learning on television. In N. Schmitt (Ed.), *Formulaic Sequences: Acquisition, Processing and Use* (pp. 249–268). Amsterdam/Philadelphia: John Benjamins Publishing.

Wray, A. (2008). *Formulaic Language: Pushing the boundaries*. Oxford, The United Kingdom: Oxford University Press.

Wray, A., (2009). Identifying formulaic language: persistent challenges and new opportunities. In R. Corrigan, E. Moravcsik, H. Oulali & K. Wheatley (Eds.), *Formulaic Language: Volume 1: Distribution and Historical Change* (pp. 27-52). Amsterdam/Philadelphia: John Benjamins Publishing.

Wray, A., Fitzpatrick, T. (2008). Why can't you just leave it alone? Deviations from memorized language as a gauge of nativelike competence. In F. Meunier & S. Granger (Eds.), *Phraseology in Foreign language Learning and Teaching* (pp. 123-148). Amsterdam/Philadelphia: John Benjamins Publishing.

Wray, A., Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication, 20*, 1-28.

APPENDICES

Appendix A. *Reference list of formulaic language items*

Additional support	Moreover, Furthermore, In addition, Similarly, In the same way, With respect to	as well as Besides, Equally important, In particular, More important, Worst,
Putting the same idea in a different way	In other words, To put it simply, To put it differently,	That is to say In view of this, With this in mind,
Opposing words	By contrast, On the other hand, In comparison,	On the contrary, Instead, though,
Giving examples	For example, For instance, To illustrate, To demonstrate,	such as Particularly, Specifically, Notably
Enumeration	First, Second,	Third, Finally,
Consequential words	As a result, Accordingly, Thereupon, Thus,	Because For this reason, In effect, Hence,
Certainty words	without doubt, Most importantly, Indubitably,	Undoubtedly, needless to say
Comparison words	in comparison, however, likewise, even so,	Nevertheless, Nonetheless, all the same
Coordinating conjunction	[…], so […].	e.g. I did not eat breakfast this morning, so I am a little hungry.
Introducing opposing ideas	It is often argued that...	Opposing views claim...
Supporting opposing ideas or partly accepting to find a common ground	Admittedly, Certainly, granted, Of course, At the same time...	to be fair, One cannot deny that... it could be argued that, ıt is true that…
Refutation of opposing ideas	However, this conclusion seems unwarranted. However, the information provided does not justify this conclusion. This conclusion is not well supported / fails to convince/ is flawed.	a stark contrast, as contrasted with, on second thoughts Nevertheless, However, On the other hand, But...
Negative words	unfounded, doubtful, questionable, oversimplified, problematic, unconvincing, unacceptable,	biased, unreliable, defective, presumptuous, vague, unwarranted, weak

Conclusion	In conclusion,	To conclude,
	In sum,	In closing,
	In summary,	To summarize,
	Therefore,	On the whole

Appendix B. *40-item-formulaic language list*

Additional Support	In addition, In the same way,	Equally important,
Putting the same idea in a different way	In other words, To put it simply,	That is to say
Opposing words	By contrast, On the other hand,	On the contrary,
Giving examples	For example, For instance,	To illustrate, such as
Enumeration	First, Second,	Third, Finally,
Consequential words	As a result, Thus,	For this reason, In effect,
Certainty words	Without doubt, Undoubtedly,	Needless to say,
Comparison words	Nevertheless,	Nonetheless,
Introducing opposing ideas	It is often argued that…	Opposing views claims that…
Supporting opposing ideas or partly accepting to find a common ground	One cannot deny that… It could be argued that…	It is true that…
Refutation of opposing ideas	However this conclusion is not well supported, Nevertheless, this conclusion is flawed.	On second thoughts,
Negative words	unfounded questionable	oversimplified
Conclusion	In conclusion, To conclude,	In closing,

Appendix C. *Cloze Test Assignment*

Most people want to get married and have children. Perhaps they want to leave a trace or offspring after them. **1** T_____ _____ _____ s_____, they want to become somehow "immortal". **2** I_____ o_____ a_____ t_____ children are the ultimate bliss in our lives. Many people agree with this statement. **3** N_n_____ , **4** o_____ v_____ c_____ t_____ it is not true.

5 F_____ , it is always claimed that having children brings happiness and meaning to one's life. **6** _____ e_____ , to raise them well becomes an ideal for parents. **7** H_____, t_____ c_____ _____ n_____ w_____ s_____, because children are also a financial and psychological burden for parents. **8** I_____ a_____ , to have children is to have responsibility. **9** T_____ _____ _____ s_____, parents spend all their time and money on their children, **10** f_____ i_____, on their education, clothing and other needs. **11** A_____ a r_____, they do not have much time or money left for them.

12 S_____, **13** I_____ _____ t_____ t_____ many people assert that it is in human nature to procreate and to see a part of them continue to live. **14** N_____ , t_____ c_____ _____ f_____ since many do not believe that it is necessarily instinctive. **15** B_____ c_____ , there are a lot of people who do not have any children and they do not have a tendency to see a part of them continue to live. **16** I_____ c_____ _____ a_____ t_____ having children, **17** f_____ t_____ r_____, it is very old-fashioned. **18** O_____ c_____ d_____ t_____ people can have other accomplishments in life **19** s_____ a_____ success and fame.

20 T_____ , children are, **21** u_____ , believed to be insurance for old age by many people. **22** T_____ , people think that when they get older their children will take care of them. **23** N__v___, they are no guarantee against loneliness in old age. **24** W_____ d_____, children, whether they are married or not, leave their homes. **25** I_____ e_____, people should not rely on much their children.

26 F_____, it is very selfish to have children just for emotional satisfaction or future worries. **27** T_____ i_____, a person should not only think that a child makes him/her happy. **28** I_____ _____ s_____ w_____, he/she should think whether he/she can make his/her child happy. To think for a child to have parents is sufficient is a rather **29** o_____ s_____ d idea. **30** O_____ _____ c_____ , if one has a child, he/she has to accept the fact that he/she is also an individual and has his/her own needs. **31** O_____ s_____ t_____ , people can understand the idea that their children will ever live with them is **32** u__f____ and **33** n_____ _____ s_____, they will eventually have their own life separately.

Clues:

1-putting the same idea in a different way
2-introducing opposite ideas
3-comparison words
4-introducing opposite ideas
5-enumeration
6-giving examples
7-refutation of opposing ideas
8-additional support
9-putting the same idea in a different way
10-giving examples
11-consequential words
12-enumeration
13-supporting opposing ideas or partly accepting to find a common ground
14-refutation of opposing ideas
15-opposing words
16-supporting opposing ideas or partly accepting to find a common ground
17-consequential words
18- supporting opposing ideas or partly accepting to find a common ground
19-giving examples
20- enumeration
21-certainty words
22- consequential words
23- comparison words
24- certainty words
25- consequential words
26- enumeration
27-giving examples
28-additional support
29-negative words
30-opposing words
31-refutation of opposing ideas
32-negative words
33- certainty words

34 _____ c_____, for most people it might be the biggest happiness to have children. **35** O_____ _____ _____ h_____ , for many people it is q____a_____ . **37** I_____ _____ w_____ , to have children is not everything in life. **38** E_____ i_____ , people can also live full and accomplished lives without children.

34-conclusion
35-opposing ideas
36-negative words
37-putting the same idea in a different way
38-additional support

Answer key:

1) To put it simply
2) It is often argued that
3) Nonetheless
4) Opposing views claim that
5) First
6) For example
7) However, this conclusion is not well supported
8) In addition
9) That is to say
10) For instance
11) As a result
12) Second
13) It is true that
14) Nevertheless, this conclusion is flawed
15) By contrast
16) It could be argued that
17) For this reason
18) One cannot deny that
19) Such as

20) Third
21) Undoubtedly
22) Thus
23) Nevertheless
24) Without doubt
25) In effect
26) Finally
27) To illustrate
28) In the same way
29) Oversimplified
30) On the contrary
31) On second thoughts
32) Unfounded
33) Needless to say
34) To conclude/In conclusion/In closing
35) On the other hand
36) Questionable
37) In other words
38) Equally important

Appendix D. *Essay topics for, pre-test, post-test, final test and delayed post-test*

<u>**Pre-test**</u>

Giving examples and explanations, write a 300-400 word essay in response to one of the following prompts. (50 mins).

-First, take 5 mins to organize your ideas.

-Have position on the issue and state this position in your thesis statement.

-Make sure you have separate paragraphs for: intro, support, and conclusion. You can have more than one supporting paragraph, of course.

Prompt: Do you agree or disagree with the following statement?

Group A: "Native speakers of English are the ones who can teach English best."

Group B: "A teacher's ability to relate well with students is more important than excellent knowledge of the subject being taught."

Use specific reasons and examples to support your ideas.

**………………………………………………………………………………………………………
……………………………………………**

<u>**Post-test**</u>

Giving examples and explanations, write a 300-400 word essay in response to one of the following prompts. (50 mins). Choose the topic that you have not written about it before in the previous test.

-First take 5 mins to organize your ideas.

-Have position on the issue and state this position in your thesis statement.

-Make sure you have separate paragraphs for: intro, support, and conclusion. You can have more than one supporting paragraph, of course.

Prompt: Do you agree or disagree with the following statement?

Group A: "Native speakers of English are the ones who can teach English best."

Group B: "A teacher's ability to relate well with students is more important than excellent knowledge of the subject being taught."

Use specific reasons and examples to support your ideas.

**………………………………………………………………………………………………………
……………………………………………**

Choose one of the topics below and write a well-organized argumentative essay.

Do you agree or disagree with the following statement? Use specific reasons and examples to support your ideas.

1. Big cities are ideal places to bring up a child.
2. Students should not be allowed to use their mobile phones during the class.
3. When choosing a career, financial gain should be the most important consideration.

Use specific reasons and examples to support your ideas.

……

………………………………………………………

Delayed Post-test

Giving examples and explanations, write a 300-400 word essay in response to the following prompts. (50 mins).

-First take 5 mins to organize your ideas.

-Have position on the issue and state this position in your thesis statement.

-Make sure you have separate paragraphs for: intro, support, and conclusion. You can have more than one supporting paragraph, of course.

Prompt: Do you agree or disagree with the following statement?

 "Teachers use technology well when teaching a course."

Use specific reasons and examples to support your ideas.

……

………………………………………………………

Appendix E. *Quiz*

Scenario: Some believe that the use of cell phones is dangerous. But you have been reading some sources and you are convinced that cell phones are not as harmful as people claim, but rather useful. Your reasons are: 1) recent studies showed that there seems no link between cell phone use and cancer; 2) they are great medium for research during class, as students can look up words and formulas; and 3) in case of emergency, cell phones can be used to call for help.

1. **Template A**

   A_ _ _ _ _ _ _ m _ _ _ _ p _ _ _ _ _ c _ _ _ _ t _ _ _ _ [o _ _ _ _ _ _ _ v _ _ _], c _ _ _ _ _ e _ _ _ _ _ _ _ _ _ s _ _ _ _ t _ _ _ [m _ v _ _ _].

2. **Template B**

   W _ _ _ _ [o _ _ _ _ _ _ _ v _ _ _], [m _ v _ _ _].

3. **Template C**

   [m _ v _ _ _] b _ _ _ _ _ _ [m _ r _ _ _ _ _ _].

B. Translate the formulaic language items into English.

Additional Support	
1. Buna ek olarak	1. -- -------
2. Aynı şekilde	2. -- --- ---- ---,
3. Yine bir o kadar önemli	3. ------- ---------,
Putting same idea in a different way	
4. Başka bir deyişle	4. -- ----- -----,
5. Basit bir şekilde açıklarsak	5. -- --- -- ------,
6. Bu demek oluyor ki	6. ---- -- -- ---
Opposing words	
7. Farklı olarak	7. -- --------,
8. Diğer taraftan	8. -- --- ---- ---,
9. Aksine	9. -- --- --------,
Giving examples	
10. Örneğin	10. --- -------,
11. Örneğin	11. --- --------,
12. Örnek vermek gerekirse	12. -- ----------,
13. mesela	13. ,---- --,
Enumeration	
14. İlk olarak	14. -----,
15. İkincisi	15. ------,
16. Üçüncüsü	16. -----,
17. Son olarak	17. -------,
Consequential words	
18. Bunun neticesinde	18. -- - ------,
19. O yüzden	19. ----,
20. Bu sebeple	20. --- ---- ------,
21. Sonuç olarak, bir bakıma	21. -- ------,
Certainty words	
22. Şüphesiz	22. ------- ----,
23. Hiç kuşkusuz	23. ----------,
24. Tabii ki	24. -------- -- ---,

Comparison words	
25. Buna rağmen, yine de	25. -----------,
26. Buna rağmen, yine de	26. -----------,
Introducing opposite ideas	
27. Genelikle şöyle bir kanı vardır	27. -- -- ----- ------ ---- …
28. Karşıt görüştekilerin şöyle bir iddiası vardır	28. -------- ----- ----- ---- …
Supporting opposing ideas or partly accepting to find a common ground	
29. –dığını kimse inkar edemez	29. --- ------ ---- ---- …
30. –diği iddia edilebilir	30. -- ----- -- ------ ---- …
31. Şurası bir gerçektir ki	31. -- -- ---- ---- …
Refutation of opposing ideas	
32. Ancak, bu varılan sonuç iyi desteklenmiyor.	32. -------, ---- ---------- -- --- - --- --------.
33. Buna rağmen, bu varılan sonuç hatalıdır.	33. -----------, ---- ---------- -- - -----.
34. Biraz düşününce,	34. -- ------ --------,
Negative words	
35. Temelsiz, asilsiz	35. ---------
36. Şüpheli, kesin olmayan, tartışmaya açık	36. ------------
37. Basite kaçan	37. --------------
Conclusion	
38. özetlersek; sonuç olarak	38. -- ----------,
39. özetlersek; sonuç olarak	39. -- --------,
40. konuyu kapatırken, özetlersek; sonuç olarak	40. -- --------,

C. Fill in the blanks with one of the formulaic language items in the box.

*however, these conclusions are not well supported   *second   *needles to say

*it is often argued that   *equally important   *finally

*such as   *in addition   *first   *to conclude

MIGRANT LABOR

Migrant labor is the movement of people from one place to another in order to find jobs. When you first hear this definition you may think a few people moving out of the country does not matter. However, you have to consider immigration may involve large numbers of people. I do not support the idea of labor migration at all, because movement of such large numbers may cause a lot of changes in the home country, host-country and the laborer's health.

(1) …………… migrant labor has a positive effects (2) …………… improvement on the economy of the home country and improvement on the immigrant's standard of living. According to these people, laborers invest in improved technology on family farms or set up new enterprises in their home-countries with the money they have earned. (3) ……………, the supporters of labor migration advocate that labor integration enables people to have a new life in new areas and the migrants also help their relatives to have a higher standard of living.

(4) ……………, migration causes a decrease in the number of people living in an area. Consequently, labor force decreases and the companies in the home-country collapse, which affects the economy of this country very badly. (5) ……………, there are hard social consequences of immigration. (6) ……………, the immigrants are split from their families, a condition which is almost impossible to stand. (7) ……………, they are thrown into a difficult cultural and linguistic environment which they are totally strange to. (8) ……………, conflicts with host-population occur, which may influence the immigrant's psychological health badly.

(9) ……………, labor migration has several bad effects, but still, some people support this movement. (10) ……………, these people do not consider the situation in all aspects. All things considered, I believe that the spread of migrant labor should be stopped before it gets out of hand.

Answer Key:
A. Based on the following scenario, write a thesis statement using the three templates given on the hand-out previously.

1) **Template A:** Although many people claim that [opposing view], closer examination shows that [My view].
2) **Template B:** While [opposing view], [my view].
3) **Template C:** [my view] because of [my reasons].

B. Translate the formulaic language items into English.

1. In addition,
2. In the same way,
3. Equally important,
4. In other words,
5. To put it simply,
6. That is to say
7. By contrast,
8. On the other hand,
9. On the contrary,
10. For example,
11. For instance,
12. To illustrate,
13. such as
14. First,
15. Second,
16. Third,
17. Finally,
18. As a result,
19. Thus,
20. For this reason,
21. In effect,
22. Without doubt,
23. Undoubtedly,
24. Needless to say,
25. Nevertheless,
26. Nonetheless,
27. It is often argued that …
28. Opposing views claim that …
29. One cannot deny that…
30. It could be argued that…
31. It is true that…
32. However, this conclusion is not well supported.
33. Nevertheless, this conclusion is flawed.
34. On second thoughts,
35. Unfounded
36. Questionable
37. oversimplified
38. In conclusion,
39. To conclude,
40. In closing,

C. Fill in the blanks with one of the formulaic language items in the box.

Answer Key

1) It is often argued that
2) such as
3) In addition
4) However, these conclusions are not well supported
5) Equally important
6) First
7) Second
8) Finally
9) To conclude
10) Needless to say

Appendix F. *Rubric used for scoring the essays*

Score of Four An essay at this level:	• shows effective writing skills • is well organized and well developed • uses details clearly and properly to support a thesis or illustrate an idea • displays ability in the use of the language • shows variety in sentence structure and range of vocabulary
Score of Three An essay at this level:	• addresses the writing topic adequately but does not meet all of the goals of the task • is adequately organized and developed • uses some details to support a thesis or illustrate an idea • shows adequate but possibly inconsistent ability with sentence structure • may contain some usage errors that make the meaning unclear
Score of Two An essay at this level may reveal one or more of the following weaknesses:	• inadequate organisation or development • poor choice of details or does not provide enough details to support or illustrate generalizations • a noticeably improper choice of words or word forms • numerous errors in sentence structure and/or usage
Score of One An essay at this level is seriously flawed by one or more of the following weaknesses:	• serious disorganisation or underdevelopment • little or no detail, or irrelevant specifics • serious and frequent errors in sentence structure or usage • serious problems with focus, may be off-topic

Appendix G. *Boxplots to check any outliers in the data*

Boxplots to define any outliers in the overall quality scores of the essays written by the groups in each test
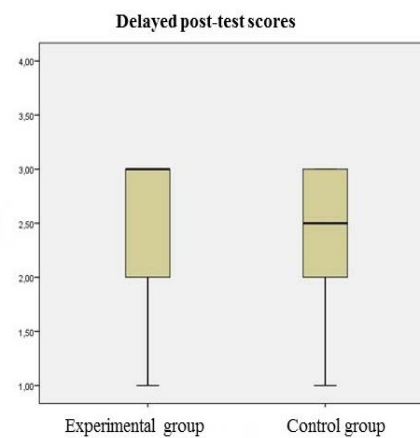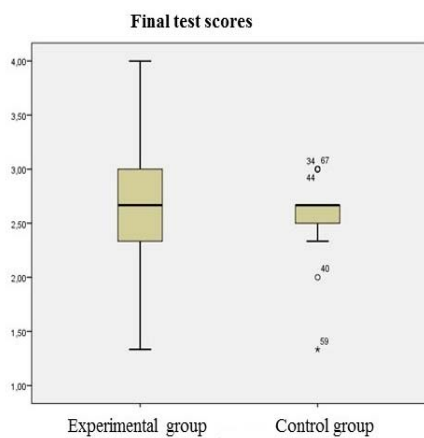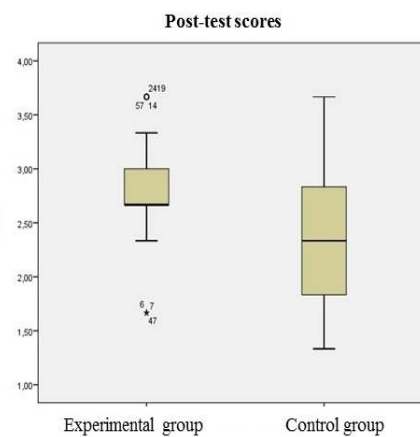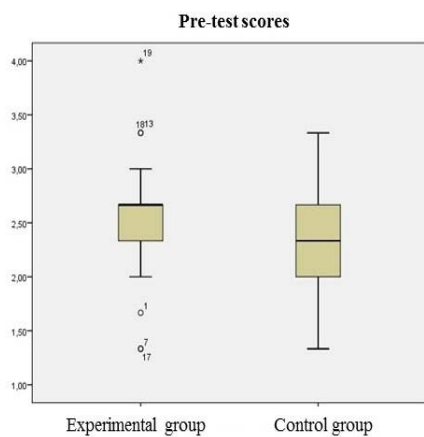
Boxplots to define any outliers regarding the types and tokens of the formulaic language items used in all of the essays
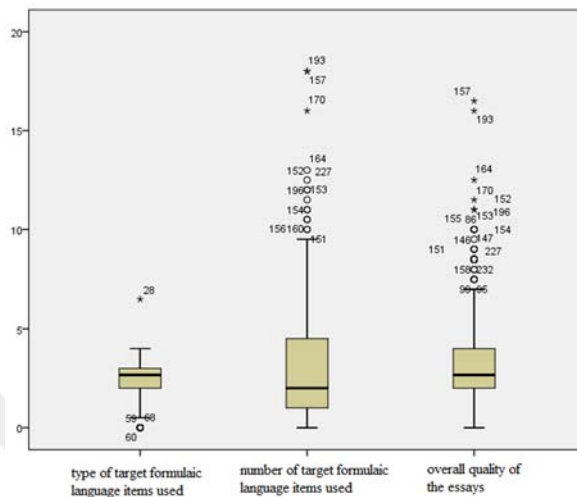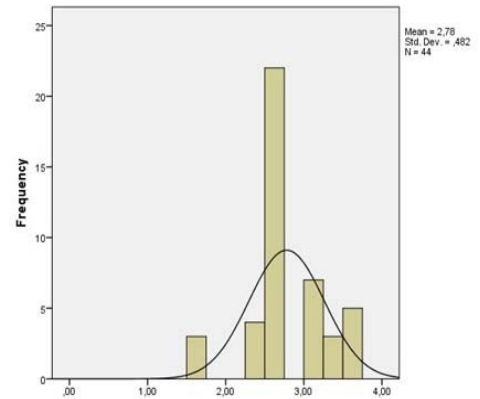
Appendix H. *Histogram graphs to check whether the data is normally distributed*

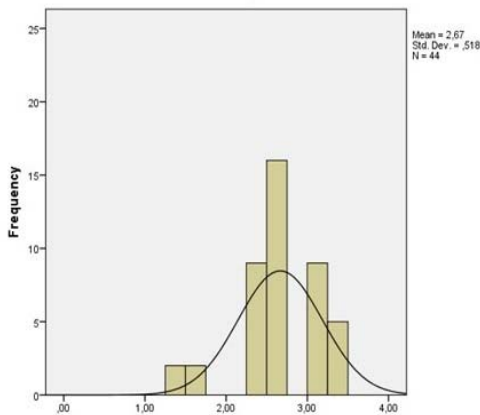**Histogram graphs for the overall quality scores of the essays written by the experimental group in each test**

Experimental Group

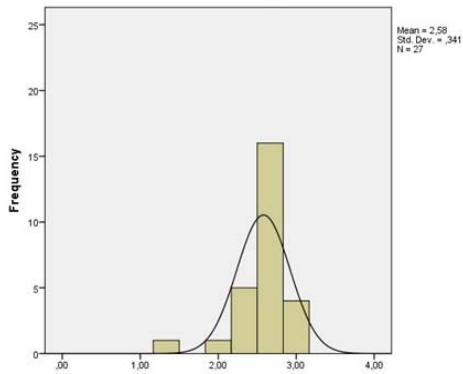**Histogram graphs for the overall quality scores of the essays written by the control group in each test**
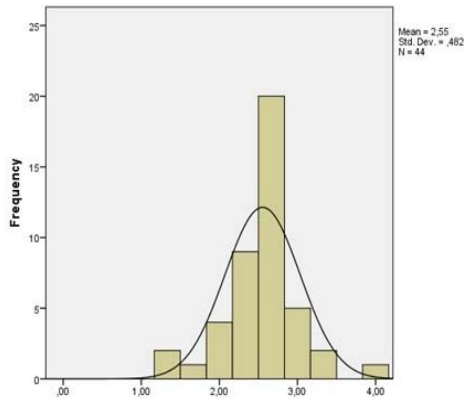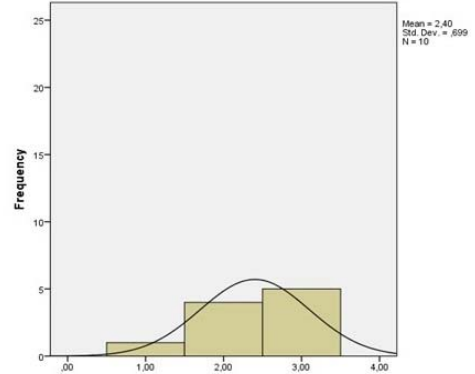

Control Group
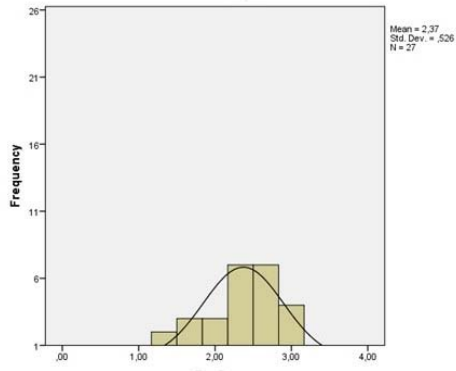
Pre-test scores
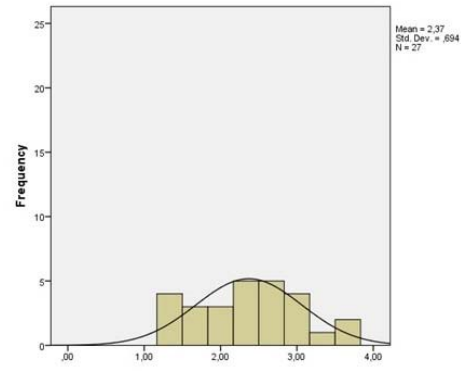
Post-test scores

Final test scores

Delayed post-test scores

ARMAGAN BAKILI AKKOC 03.03.2017

%15	%13	%9	%5
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

1	www.thesis.bilkent.edu.tr Internet Source	%1
2	www.ets.org Internet Source	%1
3	Submitted to University of Plymouth Student Paper	<%1
4	www.academypublisher.com Internet Source	<%1
5	"Overcoming Students' Misconceptions in Science", Springer Nature, 2017 Publication	<%1
6	www.norbertschmitt.co.uk Internet Source	<%1
7	www.ifets.info Internet Source	<%1
8	digitalcommons.sacredheart.edu Internet Source	<%1
9	www.psychology.nottingham.ac.uk Internet Source	<%1