EFFECTS OF TASK COMPLEXITY AND RHETORICAL MODE ON WRITING PERFORMANCE OF EFL LEARNERS

Mine YILDIZ

Doctorate Dissertation
Department of Foreign Languages Teaching
Assist. Prof. Dr. Savaş YEŞİLYURT
2017

T.C.
ATATÜRK ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
YABANCI DİLLER EĞİTİMİ ANABİLİM DALI
İNGİLİZCE EĞİTİMİ BİLİM DALI

GÖREV KARMAŞIKLIĞI VE ANLATIM TÜRÜNÜN İNGİLİZCE YABANCI DİL ÖĞRENCİLERİNİN YAZMA PERFORMANSINA ETKİSİ

(Effects of Task Complexity and Rhetorical Mode on Writing Performance of EFL Learners)

DOKTORA TEZİ

Mine YILDIZ

Danışman: Yrd. Doç. Dr. Savaş YEŞİLYURT

ERZURUM
Şubat, 2017

KABUL VE ONAY

Yrd.Doç.Dr. Savaş YEŞİLYURT danışmanlığında, Mine YILDIZ tarafından hazırlanan "Görev Karmaşıklığı ve Anlatım Türünün İngilizce Yabancı Dil Öğrencilerinin Yazma Performansına Etkisi" başlıklı çalışma 24/02/2017 tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Yabancı Diller Eğitimi Anabilim Dalı'nda Doktora Tezi olarak kabul edilmiştir.

Başkan/Danışman:Yrd.Doç.Dr.Savaş YEŞİLYURT İmza: .....

Jüri Üyesi: Prof. Dr. Mehmet TAKKAÇ İmza: .....

Jüri Üyesi: Doç. Dr. Erdinç PARLAK İmza: .....

Jüri Üyesi: Yrd. Doç. Dr. Oktay AKARSU İmza: .....

Jüri Üyesi: Yrd. Doç. Dr. Servet ÇELİK İmza: .....

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.

23.02.2017

Prof. Dr. Mustafa SÖZBİLİR
Enstitü Müdürü

TEZ ETİK VE BİLDİRİM SAYFASI

Doktora Tezi olarak sunduğum "GÖREV KARMAŞIKLIĞI VE ANLATIM TÜRÜNÜN İNGİLİZCE YABANCI DİL ÖĞRENCİLERİNİN YAZMA PERFORMANSINA ETKİSİ" başlıklı çalışmanın, tarafımdan, bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurmaksızın yazıldığını ve yararlandığım eserlerin kaynakçada gösterilenlerden olduğunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve onurumla doğrularım.

Lisansüstü Eğitim-Öğretim yönetmeliğinin ilgili maddeleri uyarınca gereğinin yapılmasını arz ederim.

24./02./2017

Mine YILDIZ

ÖZET

DOKTORA TEZİ

GÖREV KARMAŞIKLIĞI VE ANLATIM TÜRÜNÜN İNGİLİZCE YABANCI DİL ÖĞRENCİLERİNİN YAZMA PERFORMANSINA ETKİSİ

Mine YILDIZ

2017, 199 sayfa

Görev tabanlı dil eğitimi ve bilişsel görev karmaşıklığı alan yazınına dayalı olarak gerçekleştirilen bu çalışma, görev karmaşıklığı ve anlatım türünün İngilizce yabancı dil öğrencilerinin dilbilimsel zorluk, tutarlılık, uyum ve genel yazma başarısı açısından yazma performansları üzerine etkisini incelemektedir. Bu çalışmada, üç anlatım türünde yazma çalışması yapıldı: betimsel, öyküleyici ve sebep-sonuç ilişkisi anlatım. Her anlatım türü için ise karmaşık ve basit diye tanımlanan iki adet yazma çalışması yapıldı. Öğrenciler, her anlatım türünde basit görevi, içerik ve süreç planlaması yapmaları için özel zaman (15 dakika) ayrılan stratejik görev öncesi planlamayı yaparak tamamladılar. Zaman sınırlaması olmadan görev esnasında plan eşliğinde yapılacak karmaşık görevleri için ise ne yazma görevlerini tamamlamak için zaman sınırlaması ne de plan yapmak için özel zamana sahiplerdi. Tekrarlanan ölçümler yöntemi kullanılan çalışmada veriler, 41 Türk İngilizce yabancı dil öğrencisinden toplandı. Her öğrenci üç anlatım türündeki iki görev için de bir yazma çalışması yaptı ve böylece, toplamda 246 yazma çalışması elde edildi. Yazma çalışmaları, dilbilgisel zorluk, sözcüksel zorluk ve uyum açısından bilgisayar programları aracılığıyla değerlendirildi. Analitik rubrik kullanan iki değerlendirici, tüm yazma çalışmalarını tutarlılık ve genel yazım kalitesi açısından inceledi.

Tekrarlanan ölçümlerde iki yönlü MANOVA testi sonuçları, öğrenci yazma çalışmalarının anlatım türüne göre dilbilgisel ve sözcüksel karmaşıklık açısından anlamlı düzeyde farklılık gösterirken görev karmaşıklığı açısından herhangi bir fark olmadığını göstermektedir. Ayrıca, görev karmaşıklığı yalnızca bir uyum türü (referanssal uyum) üzerinde olumsuz etkiye sahipken anlatım türü uyum üzerinde orta seviyede etkiye neden olmaktadır. Tekrarlanan ölçümlerde iki yönlü ANOVA testi sonuçlarına göre, öğrencilerin yazma çalışmalarındaki tutarlılık, anlatım türünden ve

görevin stratejik görev öncesi plan ya da zaman sınırlaması olmaksızın görev esnası plana göre uygulanmasından etkilenmemektedir. Ayrıca, tekrarlanan ölçümlerde iki yönlü ANOVA testi hem anlatım türünün hem de görev karmaşıklığının genel yazma başarısı üzerinde düşük seviyede etkisini göstermektedir. Görev karmaşıklığının dilbilimsel karmaşıklık üzerine etkisiyle ilgili bulgular, Bilişsel Hipotez'i destekler niteklikte değildir. Sonuçlardan yola çıkarak bu çalışma, yazma eğitimi ve değerlendirmesi ve görev tabanlı araştırmaya yönelik öneriler sunmaktadır.

Anahtar Sözcükler: Görev tabanlı dil eğitimi, Görev karmaşıklığı, Anlatım türü, İngilizce yabancı dil öğrencileri, Yazma performansı, Dilbilgisel zorluk, Sözcüksel zorluk, Tutarlılık, Uyum, Genel yazma başarısı

ABSTRACT

DOCTORATE THESIS

EFFECTS OF TASK COMPLEXITY AND RHETORICAL MODE ON WRITING PERFORMANCE OF EFL LEARNERS

Mine YILDIZ

2017, 199 pages

Based on task-based language teaching and cognitive task complexity literature, this study examined the effects of task complexity and rhetorical mode on EFL learners' writing performance in terms of linguistic complexity, coherence, cohesion, and overall writing quality. Three rhetorical modes were studied: descriptive, narrative, and cause-and-effect writing. For each rhetorical mode, two writing tasks described as complex and simple were performed. Students were required to conduct simple task in each rhetorical mode under strategic pre-task planning during which special time (15 minutes) was given to make a plan of content and process. For their complex task carried out through unpressured on-line planning, they had neither time-pressure to complete their writing task nor special time to make a plan. In the study following a repeated-measures design, data were collected from 41 Turkish EFL learners at tertiary level. Each student wrote an essay for the two tasks in three rhetorical modes; thus, a total of 246 essays were obtained. Essays were assessed by automated analysis tools for syntactic and lexical complexity, and cohesion. Using an analytic rubric, two trained-raters evaluated essays in terms of coherence and overall writing quality.

Two-way repeated measures MANOVA tests revealed that although significant differences were seen in essays according to rhetorical mode, task complexity had no effect on the results of syntactic and lexical complexity. Furthermore, small to medium effect of rhetorical mode was observed on both types of cohesion whereas complexity of task had an adverse effect on just one of them (referential cohesion). According to two-way repeated measures ANOVA tests, coherence of students' writing was not influenced by the rhetorical mode and whether the task was performed under strategic planning or unpressured on-line planning. In addition, two-way repeated measures ANOVA tests indicated a small effect of both rhetorical mode and task complexity on

overall writing quality results. The finding regarding effects of task complexity on linguistic complexity do not have any support for the Cognition Hypothesis. In the light of these results, this study proposes some implications for writing instruction and assessment and task-based research.

Keywords: Task-based language teaching, Task complexity, Rhetorical mode, EFL learners, Writing performance, Syntactic complexity, Lexical complexity, Coherence, Cohesion, Overall writing quality

ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without the support of many individuals. To these people, my gratitude is beyond the words.

Firstly, I would like to thank my advisor Assist. Prof. Dr. Savaş YEŞİLYURT for his special guidance, constant patience, motivation, and encouragement throughout my dissertation stage and all my PhD study process. Without his guidance and support, this study would not have come to fruition. I would also extend my deep gratitude to the members of my thesis committee: Prof. Dr. Mehmet TAKKAÇ and Assist. Prof. Dr. Oktay AKARSU for sharing with me their insightful comments during my PhD study. I am also profoundly grateful to the rest of my thesis jury, Assoc. Prof. Dr. Erdinç PARLAK and Assist. Prof. Dr. Servet ÇELİK for their valuable insights and feedback.

I also want to thank Assist. Prof. Dr. Yasemin TAŞ who has given hours of her time to help me for statistical analysis. She had to be interrupted by my endless questions many times and answered each question with her patience and immense knowledge. I owe thanks to all faculties who taught the PhD courses I took and particularly to Prof. Dr. Mustafa SÖZBİLİR from whom I learned not only theoretical knowledge but also how to be a successful academic. I also feel grateful to all the participants in this study. They kindly devoted their time and energy to complete their tasks.

I am also indebted to my office-mate and close friend, Dr. Merve GEÇİKLİ for her encouragement and support in all processes of my PhD study. She always listened to me patiently and provided me with the opportunity to discuss each stage of dissertation and to get her priceless insights. My special thanks also go to my friend, Şafak MÜJDECİ who has never missed her support despite the distance and her workload.

I would like to thank TÜBİTAK (The Scientific and Technological Council of Turkey) for the financial support during the process of my entire PhD study.

Besides individuals supporting me academically, there are surely many other individuals having great importance in completion of this study. My family may not have provided feedback on my study but their emotional support is always with me. I would like to thank each member of my family individually and particularly to my

brother. Their constant support and unfailing confidence in me had great role in this accomplishment.

I cannot leave this page without expressing my deepest gratitude to my husband for his encouragement, patience and support. Last but not least, my hearty thanks go to my little naughty daughter, Elif Ece. Besides thanks, I would also like to apologize you for stealing your time we would spend together.

Erzurum-2017 **Mine YILDIZ**

TABLE OF CONTENTS

CHAPTER THREE

CHAPTER FOUR

CHAPTER FIVE

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

CAF : Complexity, Accuracy, and Fluency

EFL : English as a Foreign Language

ESL : English as a Second Language

L2 : Second Language

TBLT : Task-Based Language Teaching

CHAPTER ONE

1. INTRODUCTION

1.1. Background

Recent years in SLA research have witnessed an increasing attention to task-based language teaching (TBLT). Main focus of TBLT is to foster learning a language through the use of tasks (Larsen-Freeman & Anderson, 2011) that involve meaningful, pragmatic, and communicative activities (Willis, 1996). That is, the task is the core unit of planning, instruction, and assessment in task-based language teaching. In literature, the studies on TBLT explicitly revolve on the effects of cognitive complexity on language production particularly in terms of complexity, accuracy, and fluency (CAF) (Ellis & Yuan, 2004, 2005; Kawauchi, 2005; Kuiken, Mos, & Vedder, 2005; Ruiz-Funes, 2015; Skehan & Foster, 2005; Tavakoli & Skehan, 2005; Yang, 2014; Yang, Lu, & Weigle, 2015; Yuan & Ellis, 2003). There are two main competing models that focus on cognitive complexity effects, Skehan's Limited Attentional Capacity Model (Skehan & Foster, 1999, 2001) and Robinson's Triadic Componential Framework, (Robinson, 2001, 2003, 2005; Robinson & Gilabert, 2007), which will be briefly explained in this section as they are to be discussed in depth in Chapter 2.

Limited Attentional Capacity Model provides three dimensions for task complexity: (Skehan, 1998; Skehan, 2003; Skehan & Foster, 1997, 1999, 2001) code complexity, cognitive complexity, and communicative stress. Whereas the linguistic demands a task involves are in domain of code complexity, cognitive complexity focuses on the content of the task and the structure of a task material. On the other hand, communicative stress, the third dimension, mainly focuses on the condition in which the task is performed. Basically, according to Limited Attentional Capacity Model, basically humans have a limited cappacity to process information and learners are required to use more attentional resources while performing a task which is manipulated as complex; therefore, such a task is believed to result in trade-off effects among the three aspects of language production, CAF (Skehan, 2009).

The Triadic Componential Framework of Robinson (2001, 2003, 2005) stands in contrast to the model of Skehan in terms of complexity task output. Whereas Skehan (1998) suggests that due to limited attentional resources, learners have to prioritize between three dimensions of complexity, accuracy, and fluency, according to Robinson (2001), learners enhance their performance on all three of these dimensions (CAF). Similarly, task complexity in Robinson's model was divided into three dimensions: task complexity, task conditions and task difficulty. The first dimension, task complexity refers to information processing demands that a pedagogic task requires in terms of memory, attention and reasoning (Robinson, 2001). This dimension, characterized as "the intrinsic cognitive demands of a task which contribute to between task variation in spoken and other kinds of performance for any one learner performing a simple and a more complex version [of a task]" (Robinson, Cadierno, & Shirai, 2009, p. 535), has two types of cognitive task features as resource-directing and resource-depleting variables. The latter is renamed as resource-dispersing in the later version of Cognition Hypothesis (Robinson & Gilabert, 2007).

The main feature of these variables in the framework is that they can be manipulable and are believed to have influence on language performance and learning in different ways. Whereas resource-directing dimensions accounting for presence or absence of few elements to be compared (+/- few elements), events in the past or present, or things far or near (+/- here-and- now), presence or absence of reasoning demands imposed on the learner (+/- reasoning) make *cognitive* and *conceptual* demands, resource-dispersing dimensions that include possession of planning time allotted to learners (+/- planning), structure of a task single or multiple task (+/- single task), and the presence or absence of prior knowledge (+/- prior knowledge) that could aid in the completion of the task make *performative* and *procedural* demands on learners (Robinson, 2001). These variables were expanded in the later version of the framework (Robinson & Gilabert, 2007). The second dimension, task conditions, focuses on participant variables and participation factors such as flow of information or communicative factors; on the other hand, task difficulty, the third dimension, is concerned with ability variables as much as affective variables. Based on these models, the present study was firstly situated around planning (absence or presence) which is a resource-dispersing variable of Robinsons' Triadic Componential Framework.

Planning is presented in two main categories as: "*pre-task planning"* and "*within-task planning"* (Ellis, 2005) depending on whether it is performed before or during task. Both types of planning has also two sub-categories. Pre-task planning is differentiated according to whether learners are provided with an opportunity to perform the task before main task performance, called as 'rehearsal', or whether the learners engage in preparing for the task performance by considering the content and the way of expressing that content, called as 'strategic planning'. On the other hand, whether planning is performed under time-pressure ('pressured') or no time-pressure ('unpressured') defines the type of within-task planning, also called as 'on-line planning' (Ellis, 2005; Ellis & Yuan, 2004; Yuan & Ellis, 2003). In the light of these definitions, two types of planning were adopted in the current study. Accordingly, while one of two writing tasks in the three rhetorical modes was carried out with strategic planning during which the students were given extra time to make planning before writing, the other was conducted with careful on-line planning during which the students had neither special time to plan what and how to write nor time pressure to complete their writing performances.

The studies that investigate the effects of planning on task performance in terms of accuracy, complexity, and fluency appear to reach a three-fold conclusion: planning leads to higher fluency by decreasing on-line cognitive load and thus communicative stress; it provides the learners with an opportunity to produce a more complex language since they have the chance to use their lexical and structural knowledge at maximum level; and planning results in performance with more accurate language as the learners pay more attention to form (Kawauchi, 2005). Like many other task-based research studies that focus on oral performance, those studies investigating the effect of planning were also mostly concerned with the oral production of learners (e.g, Ahangari & Abdi, 2011; Ellis, 2009; Kawauchi, 2005; Skehan & Foster, 2005; Yuan & Ellis, 2003). However, Ellis and Yuan (2005) investigated the effects of careful on-line planning (unpressured within-task planning) on writing production of learners besides their oral performance since writing, due to its nature, is probably more influenced by careful within-task planning that provides more time to produce their text and control all processes of writing more successfully. Current research issued two types of planning by Ellis (2005), on-line planning, or with its other name within-task planning, and

strategic planning, a type of pre-task planning. Furthermore, the condition of on-line planning in the current study can be described as careful planning (Ellis & Yuan, 2005) as participants had no time pressure to perform writing tasks.

Rhetorical mode is the second dimension of the current study which is believed to cognitively affect the process and outcome of writing (Blair & Crump, 1984; Crowhurst & Piche, 1979; Engelhard, Gordon, & Gabrielson, 1992; Nemati, 1999; Prater, 1985; Prater & Padia, 1983; Shavelson & Stern, 1981; Yang, 2014). For instance, Shavelson and Stern (1981) found out that writers got the highest score for their narrative writing that was followed by descriptive and then expository writing. However, unlike our first dimension, rhetorical mode does have no clear representation in the two fundamental cognitive frameworks in the TBLT.

1.2. Focus of the Study

This study engages in exploring the effect of task complexity under two dimensions, planning and rhetorical mode of writing, on syntactic complexity, lexical complexity, coherence, cohesion, and general writing achievement of EFL learners' written production. In this study, the task is conceptualized as a goal-oriented activity in which learners are expected to pay attention to meaning rather than form and thus goes beyond the cognitive processes. Furthermore, a task can be analyzed in terms of three fundamental elements as goals, outcomes, and procedures. In this sense, students are expected to perform writing tasks to produce essays (goal and outcome) in the light of the instructions (procedure). Furthermore, this study considers the complexity of the task rather than just examining the tasks in basic terms and aims examining the effects of task complexity on students' production. However, most task-based research regarding the effects of task complexity on language production has focused on oral task performance (Adrian & Mangado, 2015; Ahangari & Abdi, 2011; Ahmadian, Abdolrezapour, & Ketabi, 2012; R. Ellis, 2009; Iwashita, 2006; Kawauchi, 2005; Rahimpour & Hazar, 2007; Salimi & Dadaspour, 2012; Yuan & Ellis, 2003); therefore, the issue of how task-based research is related to writing performance has yet to be answered.

The importance of writing, which is considered as an aid to learning a language, (Hedge, 1988), for EFL learners cannot be ignored; furthermore, a great place should be allocated to writing in syllabuses of language teaching and learning. Besides urging writers to engage in constructing knowledge telling, writing is simply a way of making meaning. Moreover, writing which requires linguistic and cognitive engagement enables learners to create a meaningful world through the language. Almost all writing activities carried out by learners can be described as 'task' since they meet the fundamental features a task should have; however, the studies on task-based research particularly examining the effect of task complexity seem to disregard writing performance of learners. In this sense, among major aims of the present study, the most salient one is to fill in a missing piece to the picture by addressing the relationship between task complexity and written task performance.

Moreover, most of the studies on written task performance focused on writing production of L2 learners (Bulté & Housen, 2014; Ellis & Yuan, 2005; Kawauchi, 2005; Kuiken & Vedder, 2007a; Ryshina-Pankova, 2015; Vyatkina, 2015); however, they do not shed direct light on the effects of task complexity on task performance of EFL writers. Unlike second language that is a medium of instruction besides the native language of learners and thus seen in a natural-like context, foreign languages are just seen classroom context and occur just as a school subject. Therefore, due to the context of language use, great difference may appear between writings of EFL learners and production of L2 writers (Ortega, 2005). In this regard, like Genç (2012), Malicka and Levkina (2012), and Ruiz-Funes (2015) setting their studies on foreign language writing, the present study focused on task performance of Turkish EFL writers.

Furthermore, there is very limited research investigating the effects of task complexity on written production conducted in Turkey. For instance, when the literature was reviewed, the only study found in Turkish context was by Genç (2012) that investigates the effects of strategic planning on the accuracy of EFL learners' both oral and written narrative task performances. Therefore, this study aims at probing into the effects of cognitive task complexity on written production of EFL learners in Turkey. In this sense, this study will probably contribute to fill the research gaps in the effects of task complexity, particularly in terms of written task performance, in Turkish context.

Particularly based on Robinsons' Triadic Componential Framework and classification of planning by Ellis (2005), we first manipulated task complexity along one of the resource dispersing variables, planning. In this sense, two types of planning, strategic planning and careful on-line planning, were applied in this study in order to investigate their effects on linguistic complexity, cohesion, coherence, and overall writing quality of EFL writers' production. Whereas their writing tasks conducted through strategic planning were identified as simple, those carried out under unpressured on-line planning were described as complex tasks. The basic explanation of why the task with strategic planning was simple, the other with unpressured on-line planning was complex was the cognitive overload participants would have during task performance. That is, writers were expected to use more their attentional resources in planning they would make while writing; however, making planning of the content of what to write before beginning to write would decrease the need of cognitive demand.

The studies investigating the effect of planning on task performance, we reviewed, mostly focused on the oral performance of L2 learners. However, planning may have great impact on writing, by its nature, which is influenced by the conditions under which it is produced. Furthermore, providing writers more time for the production of the text, easing the load of on-line process and thereby offering greater control over the process, on-line planning is particularly supposed to have impact on writing production (Ellis & Yuan, 2005). In this respect, the current study was expected to shed light on exploring the effect of strategic and on-line planning on written performance of EFL learners.

The other independent variable of our study was rhetorical mode of writing operationalized at three levels as descriptive, narrative, and cause-and-effect writing. Although it has no clear place in task complexity model by Robinson, Skehan and Foster' Limited Attentional Capacity Model includes a variable as "familiarity of discourse genre" in cognitive familiarity which is a category of cognitive complexity (Skehan, 1998, p. 99). Accordingly, whereas descriptive writing in this study involving more personal information and thus more familiar was identified as the simplest rhetorical mode (Way, Joiner, & Seaman, 2000), cause-and-effect writing was the most complex mode as it was more information based. Rhetorical mode of writing described as a particular way of presenting and organizing ideas regarding the audience and

purpose is believed to have an impact on writing performance of L2 learners. However, in the light of studies in the literature, there still exists an open question to be answered.

Another striking point that makes this study important is its dependent variables. Based on the basic assumptions of task-based research philosophy, the studies in the literature focused on three components of task performance: complexity, accuracy, and fluency. However, there are also some other components a successful performance is required to involve. Another way of stating this, besides complexity, accuracy, and fluency, other components such as coherence and cohesion may be a good predictor of a successful task performance, particularly writing performance. Therefore, unlike the studies on task-based language teaching, besides complexity, we also issued new dimensions of writing to be explored such as coherence, cohesion, and general writing achievement.

In short, we carried out this study to see whether increasing complexity of writing task along planning and the rhetorical mode in which students produce their writing has effects on some dimensions of their writing consisting of lexical complexity, syntactic complexity, coherence, cohesion, and overall writing quality. In line with these aims, following research questions were raised:

1- What are the effects of task complexity and rhetorical mode on syntactic complexity of EFL learners' writing production?
2- What are the effects of task complexity and rhetorical mode on lexical complexity of EFL learners' writing production?
3- What are the effects of task complexity and rhetorical mode on cohesion of EFL learners' writing production?
4- What are the effects of task complexity and rhetorical mode on coherence of EFL learners' writing production?
5- What are the effects of task complexity and rhetorical mode on general writing achievement of EFL learners' writing production?

1.3. Terminology

This section gives brief definitions of basic terms commonly used in the present study**.**

Task-based Language Teaching

A strong version of communicative approach, Task-based Language Teaching (TBLT) is an approach that depends on tasks as the core unit of instruction while teaching a language (Larsen-Freeman & Anderson, 2011). The main concern of TBLT is to provide learners with a meaningful and natural atmosphere in which they are occupied with real language use (Ellis, 2009; Willis & Willis, 2007).

Task

Before making a description of any task, it is of importance to define its type, whether it is a target task or pedagogical task. A target task should be transformed into the classroom so as to be considered as a pedagogical task that has sense within the scope of TBLT. There are many definitions for task from different perspectives (Ellis, 2003; Nunan, 2004; Skehan, 1998; Van den Branden, 2006; Willis, 1996). For instance, Ellis (2003), in simple terms, describes it as "a workplan that requires learners to process language pragmatically in order to achieve an outcome that can be evaluated in terms of whether the correct or appropriate propositional content has been conveyed" (p.16). On the other hand, Willis (1996) suggests that a task is an activity in which the learner uses the target language with communicative purposes to have an outcome.

Task Complexity

Besides providing basic distinction between two terms, task difficulty and task complexity, which are probably interchangeably used, Robinson (2001) also makes a description of task complexity as follows: "the result of attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learners" (p.29). On the other hand, task difficulty is the learners' perception of difficulty on which as much as affective factors, ability variables have influence.

Rhetorical Mode

Rhetoric is described as "the role of discourse toward some end: how language can be used to persuade, convince or elicit support" (Hyland, 2002, p. 208). In simple terms, rhetoric is the study and uses of effective speaking and writing in order to persuade, inform, or entertain target audience. In other words, it is described as the art

of using language at least to persuade at least one person orally or in written form. Accordingly, rhetorical mode of writing refers to the style and purpose of writing.

Syntactic Complexity

Also called as "syntactic maturity" or "linguistic complexity", syntactic complexity refers to "the range of forms that surface in language production and the degree of sophistication of such forms" (Ortega, 2003, p. 492). In line with the automated program of syntactic analyzer by Lu, the present study used 9 of 14 syntactic complexity indices under four main categories as: length of production, sentence complexity, subordination, and coordination.

Lexical Complexity

Involving such components as lexical density, lexical diversity, and lexical sophistication, lexical complexity refers to the proportion of advanced, infrequent, and different words in a text.

Cohesion

Cohesion, a complex term with intertwining treads with another component of writing, coherence (Parsons, 1991), refers to the explicit links (Todd, Khongput, & Darasawang, 2007). In other words, cohesion is "the connectivity of ideas in discourse and sentences to one another in text, thus creating the flow of information in a unified way" (Hinkel, 2004, p. 279). Due to its inherent features such as being objective and existing directly in the text itself, cohesion can be directly measured from the text (McNamara, Graesser, McCarthy, & Cai, 2014).

Coherence

Unlike cohesion, coherence refers to the implicit links that build connection between ideas to create meaning (Lee, 2002). The interaction between cohesion of the text and the reader leads to coherence (McNamara, Louwerse, & Graesser, 2002). In other words, coherence, connecting ideas at discourse level, is what the reader grabs from the text (McNamara et al., 2014).

Strategic Planning

It is the process in which students are prepared for task performance by considering the content to be encoded and the way of presenting it (Ellis & Yuan, 2005).

On-line Planning

It is "the planning that occurs on-line while learners are actually performing a task" (Ellis, 2009, p. 1). On-line planning in the current study is described as unpressured or careful planning (Ellis & Yuan, 2005) since it is conducted under no time pressure.

1.4. Outline of the Dissertation

The rest of this dissertation pursues the following organization. Chapter 2 presents the relevant literature that provided inspiration for the current research. The chapter begins with a general overview of TBLT that underpins this study, followed by explanation of task complexity. With the purpose of situating our research, the two competing models of task complexity – Robinson's Triadic Componential Framework and Skehan's Cognition Hypothesis were then presented. A review of related studies in literature that investigated the effect of task complexity from different perspectives was presented to provide a rationale for the study. Furthermore, studies on the dependent variables of this study, syntactic complexity, lexical complexity, coherence, cohesion, and general writing achievement, were also presented.

In Chapter 3 starting with the introduction of research design followed in the study, the methodological issues were explained. Following presentation of research design, the pilot study particularly carried out to test the reliability of measurement of dependent variables was explained in detail. Next, the chapter gives information about the participants who participated in the study. Then, the process of data collection is explained with an explicit description of writing tasks the students were required to perform. Through the presentation of data analysis process and tools, this chapter goes end.

Chapter 4 comprises the results presented under the research questions of the study. It is followed by the discussion of the findings for each dimension of writing production considered in this study in the light of the literature.

In the last chapter, Chapter 5, following summarization of the findings, a discussion of those findings in terms of methodological, instructional, and theoretical implications was presented. After describing limitations the current study probably has, the chapter ends with possible suggestions for future research.

CHAPTER TWO

2. LITERATURE REVIEW

2.1. Introduction

This chapter presents firstly a general overview of task-based language teaching approach, then the underpinning cognitive theories – Robinson's Cognition Hypothesis and Skehan's Limited Attentional Capacity - of task complexity research, and lastly related literature about studies investigating the effects of cognitive task complexity on L2 performance, particularly writing, in terms of linguistic complexity -syntactic and lexical -, coherence, and cohesion as well as measures applied to assess these dimensions of writing.

2.2. Task-based Language Teaching

Second language learning/teaching has recently paid a great deal of attention to task-based language teaching that depends on the use of tasks in all processes of language learning and teaching such as planning, instruction, and assessment (Richard & Rodgers, 2001) from both researchers of second language acquisition and applied linguistics. Based on the assumption that "the most effective way to teach a language is by engaging learners in real language use in the classroom" (Willis & Willis, 2007, p. 1), Task-Based Language Teaching (TBLT) is an approach where language is acquired through the use of tasks (Larsen-Freeman & Anderson, 2011). Since 1980s, TBLT, one of the examples of 'strong version' of communicative approach, has received great attention of not only second language acquisition researchers but also researchers of second language teaching as it is primarily motivated by a theory of learning (Richards & Rodgers, 2001) and poses several advantages over PPP (present-practice-produce) paradigm that is claimed to be an over-simplified approach (Kırkgöz, 2014).

Long and Norris (2000) characterizes TBLT the main focus of which is on meaning as follows:

> … an attempt to harness the benefits of a focus on meaning via adoption of an analytic syllabus, while simultaneously, trough use of focus on form

> (not forms), to deal with its shortcomings, particularly rate of development and incompleteness where grammatical accuracy is concerned (p. 599).

Task-based instruction provides learners with opportunities to learn through authentic scenarios involving meaningful, intentional, pragmatic, and surely communicative activities in which they rely on their own linguistic resources to complete the task (Arslanyilmaz, 2013; Willis, 1996). In other words, learners in a task-based lesson are provided with the opportunity to make practice of language in a meaningful and natural atmosphere inside the classroom (Ellis, 2009) where the aim is not to produce language but use it as a "vehicle for attending task goals' (Willis, 1996, p. 25).

From the clarification of task-based instruction, it is clearly understood that 'task' is a primary unit for both designing and planning a lesson or research based on TBLT. However, there are some requirements TBLT researchers need to fulfill in order to design a task. For instance, before presenting a definition of task, it is of significance to make distinction between the sorts of task – whether it is a target (or real-world task) taking place in daily life or a pedagogical task, which is, as its name implies, seen in the class. In this sense, though non-technical and non-linguistic, Long (1985) states precisely target tasks as:

> … a piece of work undertaken for oneself or for others, freely or for some reward. Thus examples of tasks include painting a fence, dressing a child, filling out a form, buying a pair of shoes, making an airline reservation, borrowing a library book, taking a driving test, typing a letter, weighing a patient, sorting letters, making a hotel reservation, writing a cheque, finding a street destination and helping someone across a road. In other words, by 'task' is meant the hundred and one things people do in everyday life, at work, at play and in between (p. 89).

To be able to define a target task as pedagogical, it is required to transform it to the classroom atmosphere.

There are several definitions of the term 'pedagogical task' since researchers approach to description of task from different perspectives and for different purposes (Breen, 1987; Ellis, 2000, 2003; Nunan, 2004; Shehadeh, 2005; Skehan, 1998; Van den Branden, 2006; Willis, 1996). Among the descriptions presented by various researchers, that of Nunan (2004) appears to be one of the most widely used:

> … a piece of classroom work that involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is focused on mobilizing their grammatical knowledge in order to express meaning, and in which the intention is to convey meaning rather than to manipulate form (p.4).

On the other hand, Ellis (2000, 2003) simply describes task as a 'workplan' that provides a learning process in which learners achieve an outcome by focusing on meaning and their own linguistic resources. Similarly, Willis (1996) and Van den Branden (2006) state that it is an activity in which learners are required to use the language for productive or receptive skills – writing and speaking – mostly in its real or real-like atmosphere in order to have an outcome. Furthermore, according to Skehan (1998), a task is "an activity in which

- meaning is primary,
- there is a goal which needs to be worked towards,
- the activity is outcome-evaluated,
- there is a real-world relationship (p.268).

Ellis (2003) asserts that these definitions particularly emphasize following dimensions: (1) the scope of a task; (2) the perspective from which a task is viewed; (3) the authenticity of a task; (4) the linguistic skills required to perform a task; (5) the psychological processes involved in task performance; and (6) the outcome of a task (p.2).

It is clearly seen in each description that the main focus of task is on meaning and the primary aim is to achieve an outcome as a result of the process using language. Although it seems to be defined as an activity or used interchangeably with activity, it is of necessity to make distinction between task and activity. Whereas the former is mainly based on communicative meaning, the main focus of the latter is on form rather than outcome. In this sense, Ellis (2009) claims that an activity to be a task should meet some criteria such as (1) it should primarily focus on meaning, (2) some kind of 'gap' should exist, (3) learners' own resources both linguistic and non-linguistic should be basic resources on which they are to rely on, and lastly (4) the outcome should be clearly defined to be distinguished from language use. In this sense, since it does not satisfy the first criterion as the focus of exercise is to produce correct language rather than to convey message based on meaning and also the last criterion as what is essential

is similarly to produce correct language instead outcome, ‘a situational grammar exercise’ shows differences from task although it seems to satisfy the second and third criteria.

In addition, a task can be also ‘focused’ or ‘unfocused. Whilst unfocused tasks are designed to enable learners to use language for general communicative purposes, focused tasks are those designed for using some specific language structures (Ellis, 2009). It can be inferred from the definition that a focused task satisfying four criteria stated above can be also described as a pedagogical task. Another classification by Van den Branden (2006) describes the task at three levels in terms of its goals (syllabus), educational activities (methodology), and assessment. In this sense, the tasks analyzed through their goals – for what purposes people learn a second language and for what functions they need to use tasks – are characterized as ‘target tasks’, the tasks which require the learners to acquire language proficiency in order to perform such tasks are ‘pedagogical tasks’, and lastly, the intermediate tasks providing opportunity to evaluate the learners’ language proficiency and to what extent the task is successfully performed by learners are described as ‘assessment tasks’. For instance, as to the goal of a task, it can be an ‘input-providing’ task which provides learners with opportunity to engage in comprehensive skills, listening or reading, or an ‘output-providing’ task in which learners engage in productive skills, writing or speaking (Ellis, 2003).

Another pivotal issue is to select or design a task. In this sense, there are a variety of elements to be considered. For instance, like Shavelson and Stern (1981) pointing out that a task should be designed in the light of six elements involving goals, activities, materials, content, students, and social community, Nunan (2004) proposes similar components presented in the following model diagrammatically to be considered in designing a task (p.41):

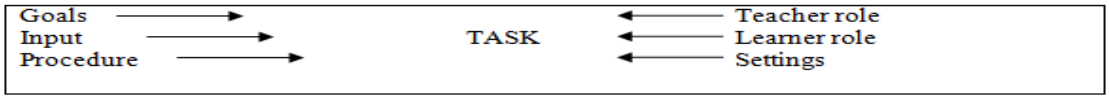


Figure 2.1. Model of Task Designing

All these elements direct the task in terms of both process and outcomes and give it meaning. For instance, goals illustrate general intentions behind any kind of task;

input simply involves all kind of oral, written or visual data to be processed by learners while performing a task; and the other component, procedure, determines what learners will do with the input to get an outcome.

Similarly, researchers differ in their perspectives of components of a task-based lesson. For instance, Ellis (2009) who states that "there is no single way of doing TBLT" (p. 224) proposes three phases involving pre-task phase, the main phase (the only one to be obligatory), and the post-task phase. Willis (1996), likewise, presents three elements but shows difference in their features.

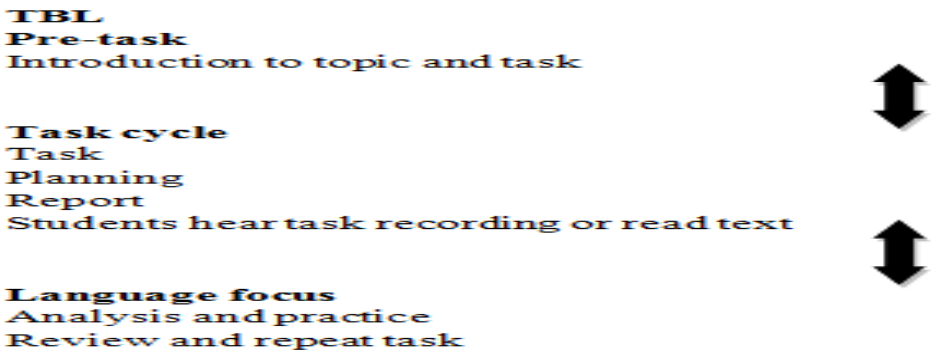


Figure 2.2. A framework for TBLT from Willis (1996, p. 135)

As seen in the figure, teachers introduce topic or task at pre-task phase through clear and insightful instructions; at the phase of task cycle, students become active by conducting task, planning how to report task outcome and task implementation process orally or written, and then reporting to the others; and lastly, in order to focus on language produced as a result of performing task, learners are expected to analyze the recording of their reports and practice the phrases, words or structures.

Among the three main approaches to TBLT – (i) an interactional approach that provides a theoretical account of interaction and language acquisition (Long, 1996); (ii) a sociocultural approach by Vygotsky that is a psychological theory of human consciousness (Lantolf, 2011); and (iii) a cognitive, information-theoretic approach that focuses on internal processes involving attention, working memory, language, and thinking – the last one is the primary approach adopted in most task-based research (Robinson, 2001, 2003, 2005, 2007; Skehan, 1998, 2001, 2003; Skehan & Foster, 1999,

2001). Like many studies in the literature, the current study is concerned with the issue from cognitive approach in the light of Robinson's Cognition Hypothesis and Skehan's Limited Attentional Capacity. Before presenting these competing models, the first step is to define task complexity which is the common feature of the two frameworks.

2.3. Task Complexity

Through the definitions of both task and task type, it is clear that the properties of task, particularly complexity have impact on language production. In this sense, the preliminary issue that those interested in task-based research must concern with is probably to select a task, operationalize task complexity or difficulty, and sequence the tasks through theoretical based, empirically attainable, and pedagogically practicable criteria (Brindley, 1987; Candlin, 1987; Robinson, 2001, 2003, 2005, 2007; Skehan, 1998, 2001, 2003; Skehan & Foster, 1997, 2001). To illustrate, Candlin (1987) provides a scheme of task sequencing components consisting of four factors that influence the level of task difficulty, cognitive load, communicative stress, particularity and generalizability, and code complexity and interpretive density. On the other hand, as seen in Figure 2.2, Brindley (1987) considers three intersecting factors that affect task sequencing: learner, task, and text or input factors. While motivation, confidence, possession of necessary language skills, relevant cultural knowledge, and prior learning experiences, and the ability to learn at the pace required are among the learner-related factors that are all supposed to make tasks easier, the task-related factors involve cognitive complexity, number of steps, and the amount of support and time provided. The features of a text such as the length, clarity, familiarity, and richness in contextual clues have also impact on difficulty of a given task. Likewise, Nunan (1989) proposed similar three dimensions of task input factors, learner factors, and procedural factors that involve a number of criteria for identifying task difficulty.

Easier	More difficult
Learner	
is confident about the task	is not confident
is motivated to carry out the task	is not motivated
has necessary prior learning experiences	has no prior experiences
can learn at pace required	cannot learn at pace required
has necessary language skills	does not have language skills
has relevant cultural knowledge	does not have relevant cultural knowledge
Task	
low cognitive complexity	cognitively complex
has few steps	has many steps
plenty of context	provided no context
plenty of help available	no help available
does not require grammatical accuracy	grammatical accuracy required
has as much time as necessary	has little time
Text / Input	is long and dense (many facts)
is short, not dense (few facts)	presentation not clear
clear presentation	few contextual clues
plenty of contextual clues	unfamiliar content
familiar, everyday content	

Figure 2.3. Three factors for identifying task: learner factors, task factors, and text or input factors (Brindley, 1987)

According to Skehan (1998), there are a number of categories that affect task difficulty as follows: differentiated outcomes, complex and numerous operations, structured tasks, familiar information, complexity of knowledge base. For instance, if a task is based on more familiar information, the performance will be probably more fluent; or, when a clear structure is provided for a task, accuracy and fluency of language performance will increase.

With the goal of demonstrating which criteria are used to sequence tasks in task-based studies, the two most influential models of task-based research will be considered in detail as most of the studies reviewed in this study are based on these models: Robinson's Triadic Componential Framework, also known as the Cognition Hypothesis (Robinson, 2001, 2003, 2005, 2007), and Skehan's Limited Attentional Capacity Model (Skehan & Foster, 1999, 2001). Both models are mainly based on cognitive approach

that primarily focuses on information processing stages, the cognitive processes, and attentional resources used by learners while completing tasks. Although at first glance these two models seemed to be contradictory with each other due to differences in their taxonomies, it can be clearly noticed that they are not so far apart at all in terms of their predictions mainly regarding resource-directing variables.

2.3.1. The trade-off hypothesis - Limited attentional capacity model

In the Limited Attentional Capacity Model, task complexity is distinguished in three dimensions (Skehan 1998, 2001, 2003; Skehan and Foster 1999, 2001): code complexity, cognitive complexity, and communicative stress (see Figure 3). Whereas code complexity concerns the linguistic demands of the task, cognitive complexity involves task content and the structuring of task material under two sub-categories as *cognitive familiarity* and *cognitive processing*. On the other hand, communicative stress, the third area, is mainly concerned with performance conditions regarding participants, presentation, text, and time.

1. Code complexity
   - Vocabulary load and variety
   - Redundancy and density
2. Cognitive complexity
   Cognitive familiarity
   - Familiarity of topic and its predictability
   - Familiarity of discourse genre
   - Familiarity of task

   Cognitive processing
   - Information organisation
   - Amount of computation
   - Clarity and sufficiency of information given
   - Information type
3. Communicative stress
   - Time limits and time pressure
   - Speed of presentation
   - Number of participants
   - Length of texts used
   - Type of response
   - Opportunities to control interaction

Figure 2.4. Skehan's Limited Attentional Capacity Model – Task Analysis Scheme (from Skehan, 1998, p.99)

The basic assumption of the Limited Attentional Capacity Model is that humans have a limited information processing capacity and manipulated task requires learners to use more attentional resources, which, thus, results in trade-off effects among the three aspects of language production: complexity, accuracy and fluency. In this sense, Skehan (2009) has simply described a successful performance in terms of a task-based context as follows:

> - more advanced language, leading to *complexity*;
> - a concern to avoid error, leading to higher *accuracy* if this is achieved; and
> - the capacity to produce speech at normal rate and without interruption, resulting in greater *fluency* (p. 1).

However, due to the limited attentional capacity learners have, they are not capable of paying simultaneous attention to those dimensions of language; that is, while paying attention to one dimension, they fail to pay attention to the others. In other words, attention to complexity probably results in trade-off effect between accuracy and fluency; namely, whereas attention is drawn to complexity, accuracy and fluency decrease and vice versa.

2.3.2. The cognition hypothesis

The main focus of the Cognition Hypothesis is to increase the cognitive demands of tasks that will contribute to their complexity along certain dimensions (Robinson & Gilabert, 2007). It claims that increasing cognitive complexity of such tasks will: "(a) push learners to greater accuracy and complexity of L2 production in order to meet the greater functional and conceptual communicative demands they place on the learner; (b) promote interaction and heightened attention to and memory for input, so increasing learning from the input, and incorporation of forms made salient in the input; as well as (c) longer term retention of input…." (p. 162).

Robinson (2001) first proposes basic distinctions between the two terms, task difficulty and task complexity, and also task conditions, "the interactive demands of tasks" as seen in Figure 2.4. However, these two terms are generally interchangeably used. He describes task complexity as "the result of the attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner" (p.29) whilst describing task difficulty as the L2 learner's "perceptions" of difficulty on which in addition to affective variables (e.g., motivation), ability variables (e.g., aptitude) have also impact.

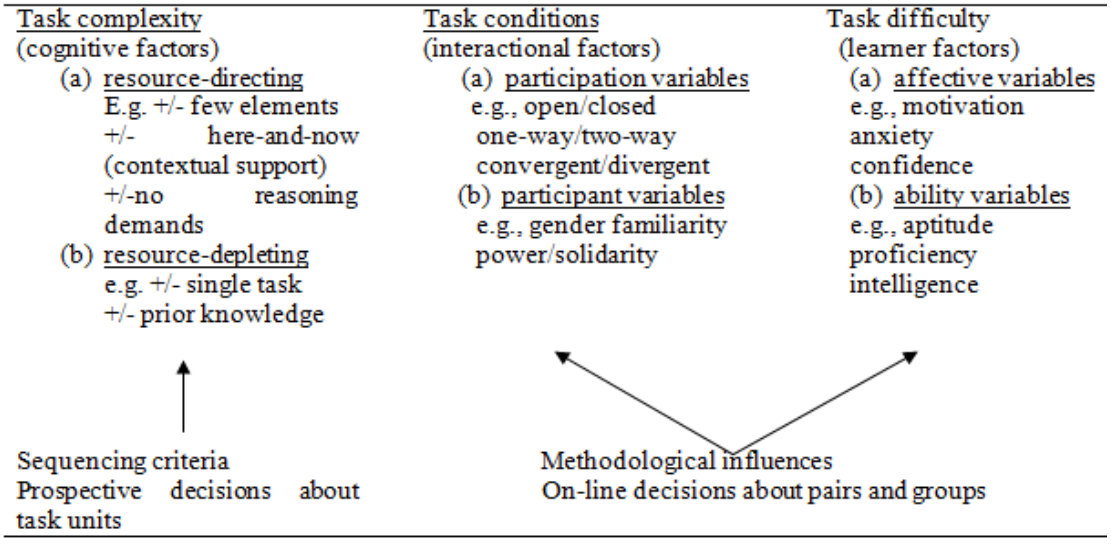


Figure 2.5. Task complexity, condition and difficulty (Robinson, 2001, p. 30)

In Figure 2.5, Robinson characterizes task complexity as involving many dimensions under two main titles 'resource-directing' and 'resource-depleting' (changed

as ‘resource-dispersing’ in the later expanded version of the Cognition Hypothesis). Both dimensions are represented by a +/- component which is described as absent or present or which may be regarded as a continuum along which a component such as here-and-now, single task, or planning may be relatively more versus relatively less. Specifically, whereas resource-directing variables involve the components such as the number of elements to compare or explain (+/- few elements), the availability of contextual support to the learner (+/- here-and-now), and the presence of reasoning demand imposed on learners (+/- reasoning-demands), resource-depleting variables consider the presence or absence of planning (+/- planning time), whether the task is singular or multiple (+/- singular task), or the possession of prior knowledge by learners (+/- prior knowledge).

The second dimension in the model, task conditions involve “participation factors such as the direction of information flow (one-way or two-way) and the communicative goals (one or many solutions) of task performance” rather than task factors or learner factors alone. In addition, participant factors, another component of task condition, include “task goal and task interpretation” besides many features such as “gender, familiarity with each other and with task role, and relative status” (Robinson, 2001, p.32-33).

Robinson also claims that it is of great importance to distinguish *learner factors* (task difficulty) making a task more or less difficult and resulting from differences between learners in terms of attentional, memory, and reasoning resource pools from *cognitive factors* (task complexity) contributing to task complexity and existing as a result of the structure of the task. There are two main kinds of learner factors: (i) affective variables including motivation, confidence, and anxiety, which are more changeable over a course of instruction and (ii) ability variables consisting of aptitude, intelligence, proficiency, and cognitive styles. Moreover, it is also believed that any interaction between all three dimensions – task complexity, task condition, and task difficulty – may have impact on task performance. For instance, complexity along a component such as tasks with no prior knowledge may affect task difficulty (e.g. learner motivation) more than any other component of task complexity; or else, task difficulty may be similarly influenced by a participation factor of task condition such as one-way open-task.

Task Complexity (Cognitive factors)	*Task Condition* (Interactive factors)	*Task Difficulty* (Learner factors)
(Classification criteria: cognitive demands) (Classification procedure: information-theoretic analyses)	(Classification criteria: interactional demands) (Classification procedure: behavior-descriptive analyses)	(Classification criteria: ability requirements) (Classification procedure: ability assessment analyses)
(a) Resource-directing variables making cognitive/conceptual demands	*(a) Participation variables* making interactional demands	*(a) Ability variables* and task-relevant resource differentials
+/− here and now +/− few elements +/− spatial reasoning +/− causal reasoning +/− intentional reasoning +/− perspective-taking	+/− open solution +/− one-way flow +/− convergent solution +/− few participants +/− few contributions needed +/− negotiation not needed	h/l working memory h/l reasoning h/l task-switching h/l aptitude h/l field independence h/l mind/intention-reading
(b) Resource-dispersing variables making performative/procedural demands	*(b) Participant variables* making interactant demands	*(b) Affective variables* and task-relevant state-trait differentials
+/− planning time +/− single task +/− task structure +/− few steps +/− independency of steps +/− prior knowledge	+/− same proficiency +/− same gender +/− familiar +/−shared content knowledge +/− equal status and role +/− shared cultural knowledge	h/l openness to experience h/l control of emotion h/l task motivation h/l processing anxiety h/l willingness to communicate h/l self-efficacy

Figure 2.6. The Triadic Componential Framework (Robinson & Gilabert, 2007, p. 164)

In the expanded version of the model (Robinson & Gilabert, 2007) illustrated in Figure 5, it is clear that whereas some new components not existing in the previous model were added, some components such as +/- reasoning-demands were replaced with new components that represent specific type of reasoning rather than a unitary concept of reasoning (+/- spatial reasoning, +/- causal reasoning, and +/- intentional reasoning). In this sense, the most striking difference between the previous and the later versions of Cognition Hypothesis is no doubt the number of components under all of the three categories. No matter which version is considered, the underpinning message of Robinson is the same: task can be selected and sequenced to facilitate language learning.

2.4. Studies on Task Complexity

In line with the aim of the present study, this section presents a review of the studies that focus on the effect of task complexity on language production, particularly in terms of writing. Although it was pointed out that task difficulty has impact on the

performance and perceptions of learners, neither a clear explanation nor any solid reason was provided about why some tasks are regarded as more difficult than others (Tavakoli & Skehan, 2005). From the perspective of French learners of English as L2, Préfontaine and Kormos (2015) dealt with learner appraisals of task difficulty and their relationships with four fluency measures under three different tasks –unrelated picture narration, story retelling, and related picture comic strips. Although the participants' perceptions of lexical and fluency difficulty showed difference in each task, these seemed to be effective variables in identifying the overall task difficulty. Another important result was that students' fluency for the three tasks was different in terms of articulation rate and average pause time, but not for pause frequency or phonation–time ratio.

Tavakoli (2009b) carried out a study with teachers and learners to illustrate their perceptions of task difficulty and criteria for identifying and defining this task difficulty. It was found that although two groups perceived task difficulty in different ways, they shared common criteria for identifying and defining it. Furthermore, when considered the criteria presented by both teachers and learners that are mostly cognitive, affective and linguistic, it was clear that the study seemed to have results more conveniently related to Skehan's model for the reason that it provides a variety of cognitive factors that are likely to have impact on task difficulty.

On the other hand, reviewing assumptions underpinning the approaches to task difficulty in speaking test Fulcher and Reiter (2003) carried out a study with the aim of presenting a new approach to defining task difficulty in speaking tests particularly in terms of the relationship between pragmatic task features and L1 background. The study focusing on politeness in requests had six tasks that were manipulated according to the conditions of social power and imposition in order to discover if independent variables regarding task conditions or L1 background could be predictive of the scores on those six tasks. It was pointed out that unlike the psycholinguistic approach, such a pragmatic approach adopted in the study may be appropriate while predicting task difficulty since it could take cultural variables into consideration. As a general conclusion, it was also found that although some learners with certain language background may regard some tasks more difficult than others, the "ability contributes more to score variance than task conditions where the rating scale is not task specific" (p. 339).

With the goal of filling the gap in the interaction among tasks, focus-on-form techniques, and performance in L2, Révész (2009) carried out a study in order to demonstrate whether +/- contextual support and recasting a focus-on-form technique had any effects on L2 morphosyntactic development. Confirming the assumptions underpinning the study, it was observed that receiving recasts without contextual support (not viewing photo) enabled to use the target feature more and more; however, those who received recast in addition to viewing the photos showed less development. As much, another important finding obtained through Rasch measurement was that the participants receiving no recasts but just viewing the photos also had better performance in using L2 target structures than those who neither saw photos nor received recasts.

The studies investigating whether increasing complexity of the task had any effect on language production show difference from many perspectives such as in terms of both dependent and independent variables although complexity, accuracy, and fluency are commonly preferred as dependent variables. For instance, Arslanyilmaz (2013) contrasted two different teaching tools – the computer-assisted task-based instruction (CATBI) and computer-assisted form-focused language instruction (CAFFI) in order to investigate the role of teaching approach in second language development in terms of accuracy, lexical complexity, and fluency. According to the results of the study, the students taught through CATBI produced better language than those taught through CAFFI; in particular, although no significant difference for lexical complexity was seen, the language of production of task-based instruction was more fluent and accurate.

Tavakoli and Foster (2008) examined how oral second language performance is affected by narrative structure (tight/loose) and narrative complexity (with or without background information) in terms of the most common measures of task complexity, complexity, accuracy, and fluency (CAF). In support of previous studies, they concluded that accuracy appeared to increase through tight task structure and also that narrative tasks with background information seemed to result in higher syntactic complexity. In another study investigating the effects of task design on L2 task performance in terms of accuracy, fluency, syntactic complexity, and lexical diversity, Tavakoli (2009a) pointed out that syntactic complexity could be enhanced through more structured tasks – narratives with both foreground and background storylines and also

that second language performance in more structured tasks seemed to be more accurate and fluent compared to those in the less structured ones. However, no clear result has been obtained for the effect of task structure on lexical diversity.

The studies reviewed by the researcher mostly depend on the two influential frameworks, Robinson's Cognition Hypothesis and Skehan's Limited Attentional Capacity Model, in terms of either dependent or independent variables, or both of them. Impressed by the cognitive frameworks for TBLT, Révész (2011) conducted a study with the goal of exploring whether there is a relationship between task complexity and learners' use of form-meaning mappings in oral tasks and also whether individual differences have impact on such a relationship. Speech production of the participants performing two versions of the same argumentative task – complex or simple - manipulated along the +/- reasoning and the +/- few elements dimensions was analyzed through some global and specific measures of oral performance. It was illustrated that although participants' speech in complex task was more accurate and lexically diverse but lower syntactically complex speech, no significant effects of learners' individual differences were observed.

In their study, Kuiken and Vedder (2007b) firstly aimed to compare the two most influential models of task complexity – Robinson's Cognition Hypothesis and Skehan's Limited Attentional Capacity Model – regarding the effect of task complexity on L2 writing performance in terms of three measures of linguistic complexity and accuracy. The learners of Italian and French were assigned two writing tasks manipualted along cognitive complexity as non-complex condition in which they were required to write a letter taking three requirements into consideration and complex condition in which six requirements would be considered. Although in previous studies Kuiken et al. (2005); Kuiken and Vedder (2008) revealed an effect of task complexity on accuracy evaluating it through general measures, the study by Kuiken and Vedder (2007) utilised more spesific measures of accuracy and lexical variation regarding error type and the most frequent words used so as to illustrate their role for such an effect. The results of the study confirmed that fewer errors were seen in the complex task which might explain the accuracy case in complex task; in other words, the fact that complex tasks yield more accurate texts probably results from a decrease of lexical errors in such tasks. As for the frequency of words, while French participants used less

frequent words in complex task, the case for the Italian participants was the opposite. In the light of the results, it was also pointed out that it seemed not possible to establish a relationship between task complexity and language proficiency level.

Operationalizing task difficulty as the storyline structure – loose or tight - Ahmadian et al. (2012) investigated the effect of task difficulty on self-repair behavior in L2 oral performance. Based on the assumption that whereas a task with clear instruction will probably establish a 'focus-on-form context', an unstructured task is expected to set up a 'communication context' in which speakers devote their much of attentional capacity to convey the message and is therefore much easier to complete, the study pointed out that task difficulty is clearly interacted with self-repair behavior in the speech of L2 learners. In other words, while performing the structured task the participants were observed to mainly focus on producing error-free units in terms of lexicon, grammar, and phonology; on the contrary, in the unstructured task, they were primarily concerned with conceptualizing the oral production producing D- repairs (different information involving alteration of the content of preverbal message) and A-repairs (appropriacy that includes changes in the content of message in terms of inaccuracy, incoherence, ambiguity, and inappropriacy) regularly.

Similarly, Adams, Nik Mohd Alwi, and Newton (2015) investigated the role of task structure and language support in increasing accuracy and linguistic complexity of writing via text chat. For their four experimental groups, they implemented two task variables – task structure (+/- TS) from Robinson's (2007) Triadic Componential framework and language support (+/- LS) utilizing pre-task in order to raise consciousness. Whereas learners in the +TS case were provided with detailed written instruction about task performance and also a worksheet guiding them, those in the condition of low task structure (-TS) were given just basic instructions but no worksheet. Similarly, learners of +LS condition were provided pre-task language support activities, but others did take no language support in their task that is therefore expected to be more complex. Analysis of the chat texts on engineering simulation task revealed that although the learners performing more complex tasks (-TS and –LS) produced less accurate texts, making tasks more complex had no impact on the linguistic complexity.

In order to investigate whether cognitive task complexity influences lexical and syntactic complexity, Frear and Bitchener (2015) utilized resource-directing variables (Robinson, 2007) by manipulating the amount of reasoning demands (+/- reasoning) and numbers of elements (+/- few elements). As a result of their analysis of letters by L2 writers of English in terms of lexical variety through a mean segmental type-token ratio and syntactic complexity by the ratio of dependent clauses to T-units, it was pointed out that an increase appeared on the lexical complexity as a result of increasing complexity of cognitive task. However, in contrast to the expectation of the Cognition Hypothesis (Robinson, 2001, 2007), in which it is assumed that increases in task complexity will lead to language development resulting in complex language performance, no significant change was seen on syntactic complexity among tasks.

Like many studies based on the assumptions of Robinson's Cognition Hypothesis, Salimi, Dadaspour, and Asadollahfam (2011) investigated the effects of tasks manipulated along resource-directing factors on accuracy, fluency and syntactic complexity. Using two versions of the same decision-making task, complex and simple, their findings on fluency and complexity confirmed the predictions of Cognition Hypothesis that complex tasks will lead to more fluent and syntactically complex texts; nevertheless, the case for accuracy was different. No significant difference was obtained between complex and simple tasks in accuracy.

Another task variable commonly used in task-based research like current study is planning that also takes place among resource-directing dimensions of Robinson's triadic framework (Ellis, 2005; Ellis & Yuan, 2005; Kawauchi, 2005; Skehan & Foster, 2005; Tavakoli & Skehan, 2005). For instance, Skehan and Foster (1997) carried out a study with 40 EFL students with diverse L1 background to examine the effect of planning and post-task activity on three tasks (personal information exchange, narrative, and decision making tasks) in terms of CAF. Planning was manipulated along whether learners were given 10 minutes to make plan or no-planning time, and similarly post-task activity took two versions (+/- knowledge of post-task activity). As an independent variable, planning was found to have clear impact on the three measures of task performance.

Ellis (2005) makes a basic classification of planning into two principal types according to when planning takes place, before or during the performance: *pre-task planning* done before performing the task and *within-task planning* occuring on-line while task is performed. Pre-task planning is also divided into rehersal planning (in which learners get opportunity to perform the task before 'main performance') and strategic planning (in which learners plan about what content to produce and how to express it without rehearsing before the task). Likewise, two main distinctions are made for within-task planning as pressured (in which learners have limited time to make plan and perform the task) and unpressured planning (in which learners are alloted limitless time to complete the task and its planning).

Ellis (2009) presented a review of studies investigating whether three types of planning (rehearsal, strategic planning, and within-task planning) influence three measures of L2 oral performance, accuracy, fluency, and complexity. The studies reviewed demontrated that although three types of planning had clear impact on fluency, its effect on accuracy and complexity was a bit varying according to the type of planning and other variables such as language proficiency, individual differences, and particularly task design.

Likewise, operationalizing planning at three levels as pre-task planning, on-line planning (unpressured performance), and no planning, Ellis and Yuan (2004) examined the effect of planning conditions on fluency, complexity, and accuracy of Chinese learners' written narrative performances. In no-planning condition were 42 undergraduate students required to complete narrating a story through pictures in written production in 17 minutes, for pre-task planning they were similarly given 17 minutes to complete writing but given also extra 10 minutes to plan before starting writing, and they had no time pressure to complete their last written task in on-line planning. It was pointed out that, whereas pre-task planning had greater impact on fluency and syntactic complexity in written texts, on-line planning resulted in greater accuracy. In additon, the results also illustrated that both sorts of planning, pre-taks planning and on-line planning, had effects on different aspects of writing process; for instance, whereas on-line planning promotes monitoring, pre-task planning provides better opportunities for formulation. As for no-planning condition, since the writers

were cognitively demanded to formulate and monitor under great pressure, it had no impact on fluency, accuracy, and complexity.

In another study investigating the effects of within-task planning (pressured vs. on-line planning) on both oral and written narrative performance in terms of complexity, accuracy, and fluency, Ellis and Yuan (2005) had consistent results that learners had syntactically complex written and oral productions in both planning conditions although any effect of the two planning was seen in neither oral nor written performance. On the other hand, after careful on-line planning learners had more accurate production than the pressured group did. Furthermore, another significant finding of the study on was that learners were more fluent in speaking tasks but more accurate and syntactically complex in their written task performance. Similarly, in their study regarding the effects of strategic pre-task planning and task complexity manipulated as complex or simple on written performance of L2 learners in terms of accuracy Salimi, Alavinia, and Hosseini (2012) found that although there was a slight relationship between strategic pre-task planning and accuracy in complex tasks performed by learners, strategic planning in simple tasks led to more accurate written texts.

Similarly, Kawauchi (2005) examining the effects of strategic planning and language proficiency on oral narrative performance of L2 learners in terms of complexity, accuracy, and fluency demonstrated that strategic planning had positive effects on the three measures. It was clear from the increased number of words produced and the decreased number of the repetitions that planned task performance of three groups – the Low-EFL, the High-EFL, and the Advanced-ESL learners - resulted in higher fluency; and similarly, greater complex narrative performance was seen in the planned performance of the three groups. Examined the use of past tense forms by the three gorups, it was clearly seen that planning had more limited impact on accuracy. As for the effect of L2 proficiency on the measures of complexity, accuracy, and fluency, it was illustrated that language proficiency is a determining factor for the effects of strategic planning on oral task performance. That is, though non-significantly different, the results showed that whereas high EFL learners gained most benefits of planning in fluency and complexity, low-EFL learners seemed to gain most in accuracy. It is generally supposed that students in higher proficiency levels produce better language

than those in lower-level proficiency, which results from the fact that the students with high level proficiency have more linguistic resources, words, phrases and structures in their working memory. In line with this assumption, Arslanyilmaz (2012) also illustrated that advanced level students provided better language production in general than the students at intermediate level of proficiency. Although advanced-level learners produced more accurate language, no significant difference was seen between their productions in terms of fluency and lexical complexity.

Like other researchers focusing on the effect of task complexity on written performance, Ong and Zhang (2010) firstly situated their study on comparison of two models - Skehan and Foster's Limited Attentional Capacity and Robinson's Cognition Hypothesis and later on filling the gap in studies investigating task complexity in terms of L2 writing. For that purpose, manipulating task coplexity using the three among resource-directing and one resource-dispersing factors, +/- planning time at four levels (extended pre-task, pre-task, free-writing, and control), +/- ideas and macro-structure at three levels (topic, ideas and macro-structure given; topic and ideas given; and topic given), and +/- draft at two levels (draft available and draft unavailable), they evaluated argumentative writing of 108 EFL learners in terms of fluency and lexical complexity. As a measure for lexical complexity, they applied the formula "WT^2/W (word types squared divided by the total number of words)" and evaluated fluency through two measures: Fluency I ("the mean number of words produced per minute of transcription") and Fluency II ("the mean number of words produced per minute of the total time spent on the task") (p. 233). It was found that increasing complexity of the task in planning condition promoted Fluency II and lexical complexity and similarly manipulating task complexity along the provision of ideas and macro-structure produced lexically complex text but had no effect on either Fluency I or Fluency II; however, no difference was seen in either fluency or lexical complexity in the case of draft availability.

Using various linguistic and discourse variables, Kormos (2011) who examined the effect of task demands on narrative writing performance of learners besides describing the linguistic and discourse characteristics of narrative writing of English learners at upper-intermediate level. The task condition varying according to whether students were required to produce a narrative story along the content they were given or

make a free plan of the content had influence on just a measure of lexical variety and had a salient effect on temporal cohesion. In addition, among measures of lexical variety and complexity, syntactic complexity, cohesion, accuracy, and the indices such as lexical competence, clausal complexity, and causal cohesion appeared to have the most power in predicting the relationship between task and writing proficiency. Moreover, it was clear from the results of the study that L1 narration showed great difference from L2 narrative writing in terms of lexical variety, sophistication, and range.

Another researcher investigating task complexity in terms of L2 learners' narrative writing performance was Ishikawa (2006). Based on the resource-directing variables from Robinson's triadic framework, Ishikawa (2006) operationalized task complexity along the dimension of Here-and-Now (HN) versus There-and-Then (TT). In the condition of HN were writers alloted 30 minutes to write a narrative in present tense after viewing a cartoon for five minutes and they were also free to see the cartoon strip whenever they needed. On the other hand, for TT condition, after seeing the cartoon for five minutes they were similarly given 30 minutes to write a narrative but in past tense and without seeing the cartoon strip. Besides manipulated task complexity, he was also concerned with the effect of language proficiency based on the scores of MEPT (Michigan English Placement Test) on the measures of accuracy, structural and lexical complexity, and fluency. The most significant result obtained was that language proficiency and task complexity had largely independent effects on dependent variables. For instance, lower-level students appeared to have greater benefits from manipulation of task complexity in terms of lexical variety particularly in target-like use of English articles although no significant difference was seen in four of the measures between two groups in both task conditions.

Furthermore, unlike numerous studies handling task complexity from English learners' angle, Way et al. (2000) investigated the effects of three writing prompts such as bare, vocabulary, and prose model and similarly three different writing mode as descriptive, narrative, and expository in order to assess the writing quality, syntactic complexity, accuracy, and fluency through four evaluation methods respectively-holistic scoring, mean length of T-units, percentage of error-free T-units, and length of product in terms of learners of French. It was shown that task difficulty had significant

impact on writing particularly for novice learners. Among the writing tasks, whereas descriptive writing was found as the easiest, the expository writing was the most difficult. In addition to the writing task, prompts also significantly influenced the writing of French learners at beginning level. For instance, whereas the highest mean scores in all four measures were seen in prose model, the lowest ones were in the bare prompts.

Similarly, Ruiz-Funes (2015) conducted a study with foreign language learners of Spanish to investigate the manipualtion of task complexity and learner-related variables in terms of syntactic complexity, accuracy, and fluency (CAF) of L2 writing. For this purpose, less complex task of advanced learners was to write an analytical essay in which they compare and contrast one or more significant development or tradition in their own culture. For the complex task, they were required to write an argumentative essay discussing the main thesis of a text they had not read before but read just before writing about it. That task was identified as complex because of the elements it involved such as selecting the text to be read, identifying its main thesis and developing their own argument. The other group with intermediate level of proficiency was asked to write a personal essay about themselves as the less complex task and an expository essay on the benefits and challenges of studying abroad as the more complex task. Although the study did not reach a statistical significant difference for the three measures, there was descriptively higher values for syntactic complexity but a decrease in accuracy and fluency for more complex tasks; that is, though being not statistically significant, it was found that task complexity determined by four variables, topic familiarity, genre, and/or task type, and reasoning demands influenced L2 writing production in terms of CAF.

On the other hand, using objective measures such as production units and complexity measures, Gan (2012) investigated the relationship among grammatical complexity measures, task type, and analytic assessment of students' speaking proficiency through an in-depth analysis of oral performance on two different tasks – presentation (monologic) task and discussion (interactive) task. As to the effect of task type on complexity of oral performance, the results demonstrated that although a closer relationship appeared in all of the four production units (T-unit, clause, verb phrase, and word) and the scores on the presentation task, among five complexity measures just the

three (length of T-units, the verb–phrase ratio, and mean length of utterance) were found higher. In other words, compared to discussion task, learners produced more complex oral performance in terms of language production units. Furthermore, it was also pointed out that mean length of units (MLU) in both tasks seemed to be a more reliable measure for predicting learners' oral language production.

Although most studies in the literature seem to deal with the skills of speaking and writing, Révész and Brunfaut (2013) investigated task difficulty in terms of listening. With the aim of examining the effects of the speed, linguistic complexity, and explicitness of the input text on the difficulty of L2 listening text, they analyzed 18 versions of a listening task in terms of a variety of measures such as the speed of delivery, phonological complexity, lexical complexity, syntactic complexity, discourse complexity, and explicitness using various automated tools. All these six measures were found, in general, to be significant predictors of task difficulty, although lexical complexity appeared to be its most critical determinant and none of the measures of syntactic complexity, neither structural complexity nor incidence of negative expressions, had direct impact on task difficulty. Likewise, Brindley and Slatyer (2002) focused on the effects of task conditions and characteristics of listening tasks on learners' listening performance. Conducting totally four listening tasks, two of which are manipulated, they investigated whether main task variables such as speech rated, text type, number of hearings, input source, and item format affected the difficulty of competency-based listening task assessment. The results of the study illustrated that the interaction between these variables and task difficulty seems to be relatively complex in that an item considered as "difficult" in a task may remain "easy" in another one.

The number of studies in literature that investigate task complexity in writing seems to be far and few between compared to the number of studies on oral performance. For instance, in their study reviewing the studies on task complexity, Salimi and Dadaspour (2012) revealed that many of the studies regarding the effects of task complexity mainly focus on L2 oral performance but just few on written performance of L2 learners. Furthermore, Ellis and Yuan (2004), Kormos (2011), Salimi et al. (2011); Yang et al. (2015) drew attention to limited number of studies on the effect of task complexity on L2 writing performance. Therefore, this study focuses on whether writing performance of EFL learners are affected by task complexity

manipulated along resource-dispersing variables (planning) by Robinson's Cognition Hypothesis and rhetorical task in which students produced their writing.

The studies on task complexity reviewed in this section are summarized in Table 2.1.

Table 2.1.

Summary of Studies on Task Complexity

Study	**Aim**	**Participants**	**Skill**	**Tasks**	**Dependent Variables**	**Results**
Skeahan and Foster (1997)	to investigate the effects of planning and post-task activity on task-based performance	40 EFL students with different L1 background	Speaking	personal information exchange task, narrative task, and decision-making task	complexity, accuracy, and fluency	Planning clearly influenced complexity, accuracy, and fluency.
Way et al. (2000)	to examine the effects of different writing tasks and writing prompts on writing quality, fluency, syntactic complexity, and accuracy of learners' production	330 French learners of English	Writing	writing task involved three modes of discourse	writing quality, fluency, syntactic complexity, and accuracy	Descriptive task was found as the easiest; however, the expository task was the most challenging. Furthermore, whereas the highest mean scores were seen in the prose model prompts, the lowest mean scores were in the bare prompts.
Brindley and Slatyer (2002)	to examine the comparability of listening assessment tasks focusing on the effects of task characteristics and task conditions	284 adult ESL learners	Listening	assessment task	listening competency	The interaction between the variables (speech rate, text type, number of hearings, input source, and item format) and task difficulty is not clear.

Table 2.1. *Continued*

Fulcher and Reiter (2003)	to provide a review of the literature relating to investigations of task difficulty and its relationship with scores awarded to students on speaking assessments, and present a new approach to the problem from the perspective	23 Spanish and 32 English-speaking students	speaking	asking to: 1)borrow book; 2) cover telephone calls; 3) help with moving; 4) swap bus seats; 5) ask for pay advance; and 6) borrow laptop		In contrast to psycholinguistic approach, a pragmatic approach may be more appropriate in predicting task difficulty; different L1 background also lead to difference in perception of task difficulty.
Ellis and Yuan (2004)	to examine the effects of three types of planning conditions on written performance of learners	42 Chinese learners of English	writing	written narrative task	fluency, complexity, and accuracy	Pre-task planning promotedr fluency and syntactic complexity; un-pressured on-line planning lead to higher accuracy; however, accuracy, fluency, and complexity were negativelly affected by no-planning condition.
Ellis and Yuan (2005)	to analyse the effects of within-task planning on both oral and written narrative performance of English learners	42 undergraduate students in China	writing and speaking	written task, oral task	complexity, accuracy, and fluency	Whereas strategic planning had influence on fluency and syntactical and lexical complexity but only sometimes on accuracy, careful on-line planning had just limited negative effect on fluency and lexical complexity but positive effect on syntactical complexity and accuracy.

Table 2.1. *Continued*

Kawauchi (2005)	to study whether strategic planning and language proficiency had any effects on oral performance of L2 learners	39 Japanese learners of English	speaking	library task, hiking task, and jogging task	complexity, accuracy, and fluency	Strategic planning had clear impact on the three measures of language production.
Ishikawa (2006)	to see the effect of task complexity and language proficiency on written performance	52 Japanese high school students	writing	narrative writing task manipulated along the dimension of here-and-now	accuracy, structural complexity, lexical complexity and fluency	Task complexity and language proficiency had independent effects on language production but some signs of their interaction were seen in accuracy and lexical complexity.
Rahimpour and Hazar (2007)	to examine the effect of topic familiarity on L2 learners' oral output	20 upper-intermediate level L2 learners	speaking	two speaking tasks: one on a familiar topic and the other on an unfamiliar topic	accuracy, complexity, and fluency	In the familiar task a more accurate and more fluent but less complex discourse was produced.
Kuiken and Vedder (2007b)	to analyse the effect of cognitive task complexity on written performance of second language learners	84 learners of Italian and 75 learners of French	writing	writing tasks manipulated along cognitive task complexity (+/-few elements)	accuracy and lexical variation	More complex tasks had higher accuracy but task complexity had varying effect on lexical variety.
Tavakoli and Foster (2008)	to examine how oral performance of second language learners was affected by narrative structure and narrative complexity	100 learners of English	speaking	tightly structured narratives, loosely structured narratives, narratives with background, narratives without background	complexity, accuracy, and fluency	Tight task structure lead to higher accuracy and syntactic complexity was increased in narrative tasks with background information.

Table 2.1. *Continued*

Revesz (2009)	to investigate how the task variable +/−contextual support combined with the focus-on-form technique affected morphosyntactic development in L2	90 adult learners of English as a foreign language	writing and speaking	written picture description task; oral photo description tasks; exist questionnaires	any improvement in using the linguistic target (past progressive form)	Receiving recast without contextual support had great effect on improvement of using the target feature
Tavakoli (2009a)	to investigate how language production was affected by task structure and storyline complexity	60 Iranian learners of English	speaking	narrative oral tasks	syntactic complexity, lexical diversity, accuracy, and fluency	More structured tasks enhanced accuracy and fluency, and syntactic complexity was related to story-line complexity
Tavakoli, (2009b)	to investigate learners and teachers' perceptions of and criteria for task difficulty	10 second language learners and 10 EFL/ESOL teachers		narrative tasks		Criteria were mostly cognitive, affective, and linguistic; two groups considered difficulty of task in different ways
Ong and Zhang (2010)	to examine the effects of task complexity on the fluency and lexical complexity of EFL learners' argumentative writing	108 EFL learners	writing	argumentative writing task manipulated along three factors: planning time; provision of ideas and macro-structure, and availability of draft	fluency and lexical complexity	Task complexity in terms of planning time had significant effect on fluency and lexical complexity; increasing complexity of task along the provision of ideas and macro-structure had effect on lexical complexity but no impact on fluency; and task complexity manipulated along draft availability had impact on neither fluency nor lexical complexity.

Table 2.1. *Continued*

Ahangari and Abdi (2011)	to assess the effect of pre-task planning on EFL learners' oral performance	40 Iranian EFL learners	speaking	decision-making task	linguistic complexity and accuracy	Whereas complexity of learners' oral performance was positively affected by pre-task planning, accuracy was not.
Kormos (2011)	to illustrate the effects of task demands on narrative writing performance	44 bilingual students	writing	cartoon description task	lexical competence, syntactic complexity, accuracy, and cohesion	Lexical sophistication and temporal cohesion were influenced by the condition in which writing is produced; furthermore, a great difference was seen between L1 and foreign language in lexical variety, sophistication and range.
Revesz (2011)	to investigate the effects of task complexity on form–meaning connections and whether individual factors had effect on those results	43 English as a second language learners	speaking	argumentative tasks manipulated along the +/- reasoning and the +/- few elements dimensions	speech production measures: syntactic complexity, lexical diversity, and accuracy interactional measures	Complex tasks resulted in more accurate and lexically diverse but lower syntactically complex language production; no effect of individual differences was observed.

Table 2.1. *Continued*

Salimi, Dadaspour, and Asadollahfam (2011)	to assess the effects of task complexity manipulated along resources- directing factors on L2 learners' writing performance	29 senior college students with Turkish background	writing	writing an essay on two versions (simple and complex) of decision making fire tasks	accuracy, fluency, and syntactic complexity	Although complex tasks resulted in more fluent and higher syntactically complex writing, accuracy of their writing was not affected by the complexity of task.
Ahmadian et al. (2012)	to investigate the effect of task difficulty on self-repair behaviour	30 Iranian EFL learners	speaking	oral narrative tasks with loose or task story-line structure	self-repair behaviours	A relationship was found between task difficulty and self-repair behaviour.
Genç (2012)	to study the effects of strategic planning on accuracy of EFL learners' written and oral production	60 Turkish EFL learners	speaking and writing	oral task with strategic planning oral task with no planning written task with strategic planning written task with no planning	accuracy	Whereas accuracy of learners' speech production was not affected by strategic planning, their written production was adversely influenced.

Table 2.1. *Continued*

Malicka and Levkina (2012)	to explore the learners' perceptions of task difficulty and the effect of task complexity on their language production	37 EFL learners	speaking	two oral instruction-giving tasks manipulated along +/- reasoning demands and +/- few elements	complexity, accuracy, and fluency	There was no difference between low and high level students in terms of perception; however, high proficiency group produced more syntactically and lexically complex and accurate but less fluent speech in the complex task. As for the less proficiency group, fluency was increased in complex task but complexity and accuracy were not affected by the complexity of task.
Salimi, Alavinia, and Hosseini (2012)	to investigate the effects of strategic pre-task planning and task complexity	50 English language learners	writing	two versions of the same decision-making task	accuracy	Strategic planning lead to more accuracy in simple task but less difference in accuracy of complex task.
Adams, Nik Mohd Alwi, and Newton (2015)	to investigate the role of task structure and language support on writing performance	96 undergraduates in Malaysia	writing	a 45-minute interactive problem solving task manipulated along task structure (+/-ts) and language support (+/-ls)	accuracy and syntactic complexity	Task complexity had influence on accuracy of students' writing but no effect on linguistic complexity.

Table 2.1. *Continued*

Frear and Bitchener (2015)	to study the effects of cognitive task complexity on lexical and syntactic complexity	34 nonnative speakers of English	writing	writing tasks manipulated along reasoning demands and number of elements	syntactic and lexical complexity	Whereas increasing cognitive complexity of task lead to an increase in lexical complexity of learners' writing, it had no influence on syntactic complexity of their writing.
Gan (2012)	to investigate the relationship among grammatical complexity measures, task type, and analytic evaluations of students' speaking proficiency	30 ESL students	speaking	presentation and discussion tasks	grammatical complexity	Individual presentation task promoted grammatical complexity; no significant correlation was seen between analytic ratings of learner speaking proficiency and complexity measures.
Arslanyılmaz (2013)	to compare two instructional methods (a computer-assisted task-based language instruction and computer-assisted form-focused language instruction) on improving the quality and quantity of language production	38 high school students learning Turkish as a foreign language	speaking	cruise-trip task, family tree task	accuracy, fluency, and lexical complexity	A computer- assisted task-based teaching approach resulted in better language compared to a computer-assisted form-focused instruction.
Revesz and Brunfaut (2013)	to examine the effects of the speed, linguistic complexity, and explicitness of the listening text on learners' listening comprehension and their perceptions of task difficulty for listening	77 university students	listening	listening task	listening comprehension	Speed, phonological complexity, lexical complexity, syntactic complexity, discourse complexity, and explicitness were found as significant predictors of task difficulty.

Table 2.1. *Continued*

Prefontaine and Kormos (2015)	to examine learner perceptions of task difficulty and to see the relationship between their perceptions and fluency measures	40 adult learners of French in a university		unrelated picture narration, story-telling, related picture comic strips	fluency	Although fluency in terms of articulation rate and average pause time was different for the three tasks, phonation-time ratio was not found different.
Ruiz-Funes (2015)	to examine the effect of task complexity and learner-related variables on written task performance of L2/FL learners	32 undergraduate learners of Spanish	writing	analytical essay, personal essay, argumentative essay, expository essay	syntactic complexity, linguistic accuracy, and fluency	Task complexity had an impact on CAF measures in two ways: first, language proficiency seemed to have a trade-off effect among measures of linguistic production; second, language proficiency had influence on the relationship between task complexity and its effects on CAF measures.

2.5. Measures for Task-based Performance

Many studies in the literature assessed L2 performance in terms of complexity, accuracy, and fluency, only one of which, complexity from two dimensions as lexical and syntactical complexity, is issued in this study. Besides relevant literature for measures of the current study (linguistic complexity, coherence, and cohesion), we will also present a brief clarification of two other dimensions, fluency and accuracy seen in most of the task-based research. Furthermore, the studies and the measures used for CAF were also presented in a table.

2.5.1. CAF triad

One significant dimension of task complexity studies is how to measure the outcome of task performance. There are numerous ways of accounting for L2 performance. However, the constructs of CAF (Skehan, 1996) are fundamental measures in several domains of SLA, particularly in TBLT and L2 writing. Whereas studies on TBLT predominantly focus on the effects of task complexity on these measures of L2 performance, L2 writing study may also investigate the relation between CAF and writing proficiency in L2. Similarly, whilst some studies examine all three or two of CAF performance dimensions, some studies examined just one of them as in our study. In this sense, we will provide a brief presentation of the definitions and measures for fluency and accuracy, but complexity – both syntactic and lexical – will be issued in detail as it is among dependent measures of this study. In this section, besides the definitions of three measures, the studies investigating the effects of task complexity on CAF reviewed by the researcher will be also presented in a table that demonstrates which study examined which dimension/s and through which measures these dimensions were evaluated.

Based on the definition by Skehan and Foster (1999), Ellis (2009) provides basic and clear definitions of CAF as follows:

> *Fluency:* the capacity to use language in real time, to emphasize meanings, possibly drawing on more lexicalized systems.
>
> *Complexity/Range:* the capacity to use more advanced language, with the possibility that such language may not be controlled so effectively.

> *Accuracy:* the ability to avoid error in performance, possibly reflecting higher levels of control in the language as well as a conservative orientation, that is, avoidance of challenging structures that might provoke error *(p. 96-97).*

As seen from the definitions and measures used to account for these dimensions illustrated in Table 2.1, it is clear that accuracy (or by the other name used in some studies "correctness") probably is the most transparent construct. In line with its definition, the absence of errors (Polio, 2001), accuracy is the ratios, frequencies or numbers of error-free units or correct forms. On the other hand, fluency basically refers to general language proficiency in productive skills – speaking and writing- particularly in terms of length of product in allotted time and the number of words produced per minute. Although it can be measured in both writing and oral performance, fluency with 'multi-componential construct' is mostly measured in speech under various sub-dimensions such as speed fluency, breakdown fluency, and repair fluency (Tavakoli & Skehan, 2005); or articulation time, pause time, pause frequency, or phonation time (Préfontaine & Kormos, 2015).

However, being a property of both language task (task complexity) and L2 performance (L2 complexity), complexity is the most complex, ambiguous, and challenging dimension of CAF (Robinson, 2001; Skehan, 2001). As demonstrated in Figure 2.7, like Skehan (2003) interpreting L2 complexity under two categories – cognitive complexity and linguistic complexity -, Bulté and Housen (2012) put L2 complexity into two groups as relative complexity (or difficulty) and absolute complexity (or just complexity). Whereas relative complexity refers to learning or processing complexity that may be influenced by learner-related and objective factors (working memory, motivation, aptitude), absolute complexity refers to objective components of L2 systems such as input saliency and linguistic complexity. Furthermore, L2 complexity is divided into three components as "propositional complexity, discourse- interactional complexity, and linguistic complexity". Among these components, linguistic complexity, which is the main concern of the present study, has received the greatest attention of both task-based teaching and L2 writing researchers. As much as being investigated at the level of language system (such as lexicon), linguistic complexity can be also examined at structure level across some

domains of language such as morphology (in formal complexity), syntax, and phonology (in functional complexity).

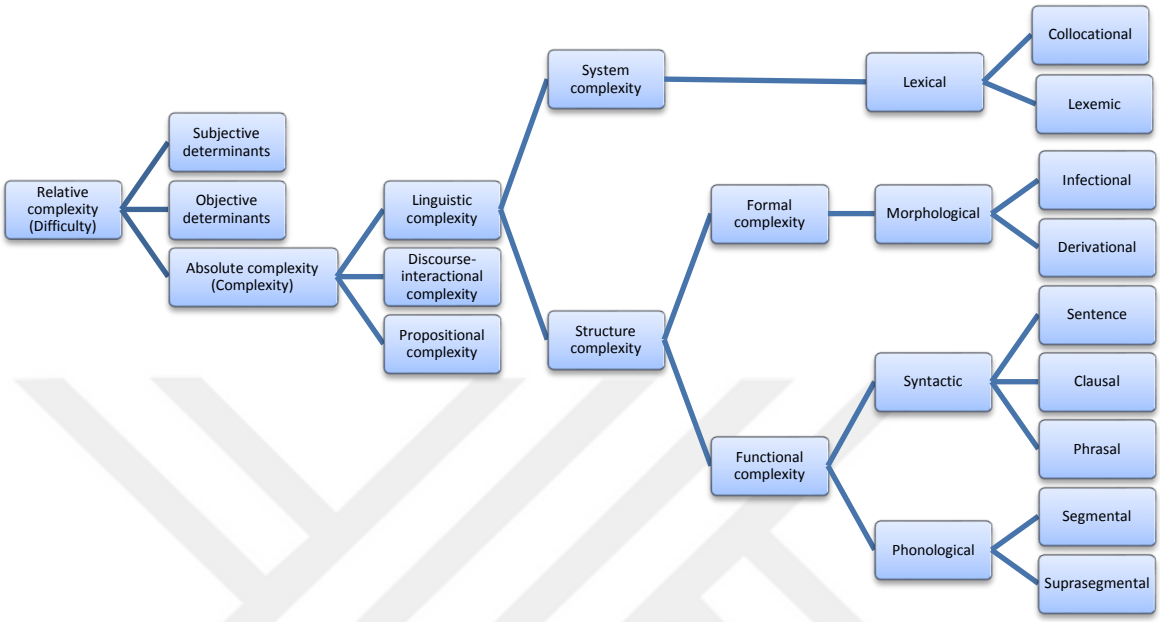


Figure 2.7. Complexity Taxonomy by Bulte and Houssen (2012)

Furthermore, as in our study, whereas cognitive complexity exists as independent variable in task-based research, linguistic complexity to be observed, measured and applied as an indicator of performance exists as a dependent variable. In addition to linguistic complexity to be explained under the name of syntactic complexity and lexical complexity, the other two aspects of language production will be also presented in terms of both their definitions and measures used in the studies reviewed (also presented in Table 2.2).

Table 2.2.

Summary of Measures Used to Assess Complexity, Accuracy, and Fluency

Study	Measures			
	Complexity		Accuracy	Fluency
	Syntactic Complexity	Lexical Complexity		
Adams et al. (2015)	embeddings, using clauses per AS-unit, words/turn as a more global measure of utterance complexity	lexical frequency and the Guiraud index	mean errors per AS-unit (general) and target-like use of auxiliary verbs and modal verbs (specific).	
Frear and Bitchener (2015)	the ratio of dependent clauses to T-units	a mean segmental type-token ratio		
Préfontaine and Kormos (2015)				articulation rate (AR); phonation–time ratio (PTR); pause frequency (PF); and average pause time (APT) - measured by Praat a speech analysis software program
Ruiz-Funes (2015)	length of production—mean length of T-unit (MLTU); amount of coordination—mean number of T-units per sentence (TUS); and amount of subordination—mean number of dependent clauses per T-unit (CTU)		total number of errors, Total number of errors per T-unit (Etot/T = total number of errors divided by the total number of T-units), and Number of errors per 100 words	by length of text produced in timed writing

Table 2.2. *Continued*

Arslanyilmaz (2013)		TTR (calculated as the total number of different words divided by the total number of words)	by the ratio of error-free terminable units	by the number of words per minute
Révész and Brunfaut (2013)	structural complexity, incidence of negative expressions.The syntactic complexity of the listening texts was assessed in terms of four types of indices: complexity by subordination, phrasal complexity, overall complexity (Norris & Ortega, 2009), and incidence of negative expressions.	lexical frequency- assessed with the help of Web VocabProfiler v3 lexical density assessed using the program Web VocabProfiler v3 lexical diversity-measured by D-formula.		
Arslanyilmaz, 2012	the number of subordinate clauses per 'C-unit'	mean segmental type token ratio	by the ratio of 'error-free T-units' to total 'T-units'	by the number of words per minute (WPM)
Salimi et al. (2012)	length of T-units in terms of number of words the number of clauses per T-unit (the T-unit complexity ratio; the ratio of dependent clauses to the total number of clauses (the dependent clause ratio) the number of verb phrases per T-unit (the verb–phrase ratio) mean length of utterance (MLU)		the number of error-free T-units per T-units	

Table 2.2. *Continued*

Kormos (2011)	clause length ratio of subordinate clauses words before main verbs through the Coh-Metrix 2.0 program modifiers per NP through the Coh-Metrix 2.0 program	lexical Variety: D-value based on the VocD software of CHILDES lexical complexity: lexical range with the help of Nation's Range program	ratio of error-free clauses	
Révész (2011)	by dividing the number of syntactic clauses by the number of AS-units	values of D by using the computer program vocd in CLAN	the ratio of errors to AS-units; the ratio of error-free AS-units to AS units	
Salimi et al. (2011)	a measure of S-nodes per T-units		the number error-free T-units per T-units	the fluency of the written production of the learners was measured by words per T-units
Ong and Zhang (2010)		WT2/W (word types squared divided by the total number of words)		two measures of fluency - fluency I and fluency II. Fluency I: the mean number of words produced per minute of transcription, Fluency II: the mean number of words produced per minute out of the total time spent on the task.

Table 2.2. *Continued*

Tavakoli and Foster (2008)	analysis of speech units (AS-units), clauses, mean length of unit (MLU)	D using VocD analysis program on CHST-formatted transcripts	error-free clauses	four repair fluency: reformulations, false starts, word replacement, two breakdown fluency: repetition mid-clause pauses greater than 0.4s, end-clause pauses greater than 0.4s
Folkert Kuiken and Vedder (2007b)	the number of clauses per T-unit; the number of dependent clauses per clause	by means of a type-token ratio	the total number of errors per T-unit was calculated with respect to grammar, lexicon, orthography and appropriateness	
Ellis and Yuan (2005)	syntactic complexity: the ratio of clauses to T-units in the participants' production syntactic variety: the total number of different grammatical verb forms mean segmental type/token ratio (MSTTR)		error-free clauses correct verb forms	production rate disfluencies

Table 2.2. *Continued*

Kawauchi (2005)	clauses per T-unit words per T-unit subordination number of types		number of words repetitions
Ellis and Yuan (2004)	syntactic complexity—the ratio of clauses to T-units in the participants' production syntactic variety—the total number of different grammatical verb forms mean Segmental Type-Token Ratio (MSTTR).	error-free clauses correct verb forms	syllables per minute number of dysfluencies
Way et al. (2000)	mean length of T-unit	percentage of correct T-units- error-free T-units	length of product

2.5.1.1. Syntactic complexity

Syntactic complexity that "(also called syntactic maturity or linguistic complexity) refers to the range of forms that surface in language production and the degree of sophistication of such forms" (Ortega, 2003, p. 492) is of great importance particularly for SLA researchers since it may be regarded as an indicator of language development. In this sense, different researchers applied miscellaneous measures with the purpose of investigating the effect of some interventions on the development of writing skill, identifying differences in writing of students at various levels of L2 proficiency and over time, and examining task-related disparities in writing (Polio, 2000). In his review of literature, Polio (2000) concluded that average length of production unit, frequency of some structural types such as "passive sentences or dependent clauses usually within a certain period of time", and complexity ratios such as clauses per T-unit seemed to be global measures of syntactic complexity.

In his study based on a synthesis of twenty-three L2 writing studies on college-level students, Ortega (2003) reached three main results: first, compared to FL learners, L2 learners probably produced writing with higher syntactic complexity; second difference in proficiency level was related to whether it was measured by a program or holistic scale; and lastly, for four of six complexity measures applied in the study – mean length of sentence (MLS), mean length of T-unit (MLTU), mean length of clause (MLC), and mean number of clauses per T-unit (C/TU), critical magnitudes were proposed. In the light of these results, he also suggested that a longitudinal instruction (e.g., one-year period) at college-level may result in substantial changes in writing of students in terms of syntactic complexity, particularly in MLTU.

On the other hand, Iwashita (2006) who carried out a study to find the most reliable and valid syntactic complexity measure for oral language pointed out that not smaller units but T-unit better reflected the differences between proficiency levels; that is, the best way to predict oral proficiency in second language was the length of T-unit. Beers and Nagy (2009) investigating the interaction between text quality and syntactic complexity through two measures of syntactic complexity- words per clause and clauses per T-unit - pointed out that syntactic complexity was correlated with the quality of text produced by adolescent writers, but such relationship depended on both genre of text

and measures of syntactic complexity. In other words, their results indicated that whereas the measure of syntactic complexity, words per clause, was positively related to essay quality but negatively to the quality for narratives, the case in clauses per T-unit was vice versa.

Makinen, Loukusa, Nieminen, Leinonen, and Kunnari (2013) examined picture-elicited narration of Finnish young learners for "narrative productivity, syntactic complexity, referential cohesion and event content" in order to reflect how their narrative skills develop through a story generation task and what the relationship is between narrative productivity and event content. In line with their aim, they transcribed their data using "the CHAT format of the Child Language Data Exchange System (CHILDES)", chose three measures for productivity – the number of T-units, the number of different word tokens, and the total number of word tokens-, two of them for syntactic complexity – the mean length of communication unit in words that was automatically analyzed by CLAN (Computer Language Analysis) and clausal density, and manually analyzed referential cohesion and event content. At the end of the study, a subtle development was noticed in each measure for all students; however, productivity and event content of narration of five-year-old students significantly differed from those of four-year-old students. Another outstanding finding demonstrated that although older students produced longer texts than did younger students, there was no significant difference in syntactic complexity of texts produced by both groups; that is, it can be concluded that, contrary to the assumption, the longer text does not mean that they are "syntactically complex".

In his study, Lu (2008) applied the revised version of Developmental Level (D-Level) Scale (Covington, He, Brown, Naci, & Brown, 2006) originally developed by Rosenberg and Abbeduto (1987). Whereas the original scale does not rate every sentence and is not used for all levels and in the studies of children and impaired adults, Covington et al. (2006) extended the scale to every sentence and rearranged levels based on psycholinguistic basis. The revised D-Level Scale consists of eight levels from 0 to 7:

- Level 0: simple sentences involving questions, sentences with auxiliaries and semi-auxiliaries, and simple elliptical (incomplete) sentences;

- Level 1: infinitive or *-ing* complement with same subject as main clause;
- Level 2: conjoined noun phrases in subject position, sentences conjoined with a coordinating conjunction, conjoined verbal, adjectival, or adverbial constructions;
- Level 3: relative (or appositional) clause modifying object of main verb, nominalization in object position, finite clause as object of main verb, subject extraposition, raising;
- Level 4: non-finite complement with its own understood subject, comparative with object of comparison;
- Level 5: sentences joined by a subordinating conjunction, nonfinite clauses in adjunct (not complement) positions;
- Level 6: relative (or appositional) clause modifying subject of main verb, embedded clause serving as subject of main verb, nominalization serving as subject of main verb; and
- Level 7: more than one level of embedding in a single sentence (p. 16-17)

In addition to describing a heuristic-based system that takes a sentence and rates it according to an appropriate level identified above, Lu (2008) conducted an experiment using child language acquisition data as the system design for these data. The results of the experiment indicated that an accuracy of 93.2% on unseen spoken data of child language was achieved by the system.

In another study describing a computational system that automatically analyzes syntactic complexity of written sample in English, Lu (2010) constructed a set of 14 syntactic complexity measures in line with two studies- six from Ortega (2003) and five from Wolfe-Quintero, Inagaki and Kim (1998 cited in Lu, 2011) and three new measures in the light of the suggestions by Wolf-Quintero et al. (1998 cited in Lu, 2011). He put this set into five categories as follows (Lu, 2010, 2011, p.43-44):

(1) *Length of production* that consists of three measures at the clausal, sentential, and T-unit level,
   - mean length of clause (MLC)
   - mean length of sentence (MLS)
   - mean length of T-unit (MLT)

(2) *sentence complexity* involves a sentence complexity ratio,
   - clauses per sentence (C/S)

(3) *subordination* that comprises four ratios reflecting the amount of subordination
   - clauses per T-unit (C/T)
   - complex T-units per T-unit (CT/T)
   - dependent clauses per clause (DC/C)

- dependent clauses per T-unit (DC/T)

(4) *coordination* that contains three ratios measuring the amount of coordination
- coordinate phrases per clause (CP/C)
- coordinate phrases per T-unit (CP/T)
- T-units per sentence (T/S)

(5) *particular structures* that consist of three ratios related to larger production units
- complex nominals per clause (CN/C)
- complex nominals per T-unit (CN/T)
- verb phrases per T-unit (VP/T)

Since the system is developed for high proficiency in second language, Lu (2010) used the data from college-level students so as to evaluate the system and illustrate "how the system is used in an example application to investigate whether and to what extent each of these measures significantly differentiate between different proficiency levels" (p.474). The results of the study demonstrated that the system achieved a high reliability in computing those indices of syntactic complexity in addition to identifying language production units and syntactic structures such as word count, sentence, verb phrase, clause, T-unit, dependent clause, complex T-unit, coordinate phrase, and complex nominal.

Similarly, in his study Lu (2011) used the same computational system and evaluated 14 syntactic complexity measures on the basis of corpus in order to reflect how these complexity measures perform as objective indices of language development of ESL writers at college-level, what the relationship among them is, and how external factors such as institution, genre and sampling condition have impact on these relationships to proficiency. His results consistent with those of Beers and Nagy (2009) in terms of genre effect and Ellis and Yuan (2004) in terms of time planning revealed that the relationship between syntactic complexity and language development was significantly affected by external factors such as institution, genre, and timing condition.

The same system "L2 Syntactic Complexity Analyzer" was also used by Ai and Lu (2013) in their study aiming to investigate whether and to what extent there are differences between NNS and NS university students' writing in terms of syntactic complexity. However, unlike Lu (2010, 2011), they subtracted four measures of the fourteen measures- complex T-units per clause, complex T-units per T-unit, clauses per

sentence, and verb phrases per T-unit on account of the fact that they were found to be poor indices for syntactic complexity by Lu (2011). In addition, just six of the nine syntactic structures the system automatically analyzes were involved in this study. The results illustrated that NNS and NS students' writings significantly differ in four aspects of syntactic complexity involving "length of production unit, amount of subordination, amount of coordination, and degree of phrasal sophistication" (p. 257). Similarly, using the fourteen syntactic measures of the same program, Lu and Ai (2015) analyzed syntactic complexity of argumentative writing produced by native speakers and non-native speakers of English with different L1 background so as to find the differences in their essays in terms of syntactic complexity. Their study revealed that significant differences in all of 14 measures were observed between NSs and NSSs put into groups according to their L1, but when they were not grouped by their L1, significant differences were seen only in three measures of them.

In order to illustrate the relationship between syntactic complexity features of native speakers and their speaking proficiency levels, Chen and Zechner (2011) used "the Stanford Parser", "Tregex package", and Lu's "L2 Syntactic Complexity Analyzer" to compute syntactic complexity features put into two categories: "(1) Clause and sentence Boundary based features (CB features)" that involve mean length of sentences, mean length of T-units, mean number of dependent clauses per clause, mean length of simple sentences, frequency of simple sentences per 1000 words, frequency of adjective clauses per 1000 words, frequency of fragments per 1000 words, mean length of coordinate clauses; and (2) "Parse Tree based features (PT features)" including mean number of complex T-units per T-unit, mean number of linguistically meaningful prepositional phrases (PP) per sentence, mean number of noun phrases (NP) per sentence, mean number of complex nominal per sentence, mean number of linguistically meaningful units, mean number of passives per sentence, verb phrases per T-unit, mean number of dependent infinitives per T-unit, mean number of parsing tree levels per sentence, and mean P-based Sampson per sentence (p.726- 727). In line with the results of the study, it was concluded that speaking proficiency scores of the speakers can be predicted through the features of syntactic complexity.

Mazgutova and Kormos (2015) who intended to examine the effect of one-year long English course on syntactic and lexical structures of L2 learners' written

productions used various software packages such as "Coh-Metrix 2.0 and Coh-Metrix 3.0, Lu's L2 Syntactic Complexity Analyzer and Lexical Complexity Analyzer, and Vocabprofiler BNC" (p. 9) so as to analyze syntactic complexity and lexical diversity of those texts. As supposed, some changes were observed in the lexical characteristics of academic writing of students both at lower and higher level of proficiency. However, opposed to the expectation that the more their writing skills developed, the more various syntactic structures they would use, syntactic variety was observed to decrease in their writing. Likewise, with the purpose of investigating the impact of L2 syntactic development on writing quality judged by raters, Crossley and McNamara (2014) calculated eleven indices of syntactic complexity that measure varying features of language at both sentence and phrase level including sentence variety, syntactic transformations, syntactic embeddings, phrase types and phrase length through the use of a computational tool, Coh-Metrix. It was found out that although students' writing developed in syntactic complexity, most of the syntactic structures did not predict the ratings of writing quality.

For the similar purpose - to investigate the relationship between such objective measures of linguistic complexity and subjective evaluation of writing quality by raters and also identify which complexity measures are the best predictors of rater judgments for L2 writing quality -, Bulté and Housen (2014) also evaluated language development of ESL learners in a short-terms EAP program in terms of writing proficiency by means of various measures for syntactic and lexical complexity. Whereas they applied ten measures based either on the *average length* or on a *ratio* of a specific linguistic unit to manually analyze syntactic complexity, three measures of lexical complexity were applied for three related but different aspects of lexical complexity – diversity index *D* for lexical diversity, Guiraud index *G* for lexical richness, and advanced Guiraud *AG* for lexical sophistication – and their analysis was accomplished through automated tools. Their results revealed that a sufficient development was achieved in L2 writing particularly in terms of all levels of syntactic construct over time; however, no significant difference was seen in the most popular complexity diagnostics in the field - subordination ratios and lexical richness. Furthermore, it was also pointed out that dimensions (e.g., lexical richness) were correlated well with subjective ratings of writing quality.

With the goal of pointing out the relationship among writing topic, writing quality, and syntactic complexity, Yang et al. (2015) analyzed totally 380 argumentative essays written by 190 nonnative English speaking students on two different topics. They evaluated general writing proficiency through a five-point scale by human raters and assessed the syntactic complexity by eight different measures including mean length of sentence (MLS), T-units per sentence (TU/S), mean length of T-unit (MLTU), mean length of clause (MLC), dependent clauses per T-unit (DC/TU), coordinate phrases per clause (CP/C), complex noun phrases per clause (CNP/C), and non-finite elements per clause (NFE/C). Using Lu's L2 syntactic complexity analyzer (L2SCA), they both defined most of their linguistic units and computed their six measures - MLS, MLTU, MLC, TU/S, DC/TU, and CP/C. Though at global level, a significant relationship was found between syntactic complexity features such as sentence complexity and amount of subordination and writing quality, the relationship at local level, "clausal coordination, finite subordination, overall elaboration at the finite clause level, non-finite subordination, phrasal coordination, and noun-phrase complexity" (p.64), showed difference across topics. Furthermore, it was also revealed that global sentence complexity measured through mean length of sentence and T-unit complexity seemed to be significant predictors of higher writing quality as scored by human raters though complexity features at local level appeared to vary across two topics in predicting scores.

2.5.1.2. Lexical complexity

Another dimension that is proposed to be affected by task complexity and may be also an indicator of successful L2 performance is lexical complexity. However, as seen in Table 2. 2, although it is a sort of complexity to be assessed in second language performance and "vital in performance models" by Skehan, lexical complexity seems to rarely take place in task-based studies (Skehan, 2009). Before identifying the measure to assess it, it is primarily required to decide which dimension of lexical complexity will be measured since each measure refers to different dimension. For instance, whilst type-token ratio is applied to calculate lexical variety, Lexical Frequency Profile (Laufer & Nation, 1999 cited in Skehan 2009) aids assessing lexical sophistication. Like many studies investigating the effect of task-complexity on CAF triad, while some studies

focus on just one dimension (e.g., Arslanyilmaz, 2012, 2013; Frear & Bitchener, 2015; Kuiken & Vedder, 2007b used lexical variety) others may measure two or three of the dimensions of lexical complexity (e.g., Révész & Brunfaut, 2013 used lexical frequency, lexical density, and lexical diversity). Because the present study depends on three dimensions –lexical sophistication, lexical density, and lexical diversity by Lu's Lexical Complexity Analyzer-, this section primarily proposes lexical complexity measures used in the studies in literature to assess these dimensions.

In general, Révész and Brunfaut (2013) divide lexical complexity into four sub-categories: lexical frequency, lexical density, lexical variety, and lexical complexity while Skehan (2009) proposes three main dimensions such as lexical sophistication, lexical diversity, and lexical richness. Lexical frequency, as its name implies, is usage frequency of a word. Lexical density refers to "the proportion of content words to the total number of words" in a text (Révész & Brunfaut, 2013, p. 38). Since a text rich in content words probably conveys more conceptual information than a text with a higher proportion of function words, it is also dense in information; that is, lexical density can be regarded as an indicator of information density. Lexical variety (also called as 'lexical diversity') that is a shared category in all distinctions refers to the variety and range of words in a text. Another commonly measured and shared dimension in both distinctions above is lexical complexity or, in its other name, lexical sophistication that refers to "the proportion of relatively unusual or advanced words" in a text (Read, 2000, p. 203).

Like other measures dealt with beforehand, lexical complexity also involves the basic issue concerning how to measure. While some studies prefer automated tools to assess it quantitatively, the others focus on manual evaluation. For instance, Lu (2012) developed and used a tool "Lexical Complexity Analyzer" to assess three main dimensions of lexical complexity– lexical density, lexical sophistication, lexical variation involving also many sub-components (number of different words, type-token ratio, verb diversity, and lexical word diversity), but Arslanyılmaz (2012, 2013) manually calculated type-token ratio by dividing the total number of words into the total number of different words.

As seen in Table 2.2 that demonstrates measures applied in previous studies, most of the studies depend on either lexical variety or lexical complexity; or both of them. Furthermore, type-token ratio is the most common measure applied in both automated tools and manual assessment to calculate lexical variety (Arslanyilmaz, 2012, 2013; Frear & Bitchener, 2015; Kuiken & Vedder, 2007b). In simple terms, token refers to the words in a text and type is the word group that a word belongs to. The number of words calculated in a text refers to the number of tokens and thus the number of type is expected to be less since more than one token may belong to the same type. However, the greater the number of types is in a text, the more various words there are and the more lexical variety is in the text. The relationship between the number of types and that of tokens is described as type-token ratio (TTR). Type-token ratio that is easily measured by many automated or web-based tools such as Lu's Lexical Complexity Analyzer and Coh-Metrix is basically calculated as follows:

*Type-token ratio= (number of types/number of tokens)*100*

Another commonly used measure for lexical diversity or variety is D (Diversity) index based on a mathematical formula and developed by Malvern and Richards (2002) in order to overcome the disadvantage of TTR with sample size; that is, since "larger samples of words will give a lower TTR" measures using TTR independent of sample size will be problematic. They also produced a VocD analysis program to automatically calculate D. Many studies on task complexity, as seen in the table, applied it to measure lexical variety (Kormos, 2011; Revesz, 2013; Révész, 2011; Tavakoli & Foster, 2008). Similarly, some studies used G index (the index of Guiraud) to calculate lexical richness or variety as it is advantageous over other measures in terms of text length (Adams et al., 2015). On the other hand, in their study investigating the effects of task complexity manipulated at three levels (+/- planning time, +/- draft, and provision of ideas and structure) on fluency and lexical complexity, Ong and Zhang (2010) calculated lexical complexity by using "the formula WT2/W (word types squared divided by the total number of words)".

In the light of the studies, we aimed to measure both lexical complexity and syntactic complexity through automated tools in order to investigate whether complex

tasks result in complexity in lexis and syntax since it will be assumed to produce more objective and thus more reliable results.

2.5.2. Coherence and cohesion

Other two dependent variables of the present study to be evaluated to see the effect of task complexity on L2 writing performance are coherence and cohesion. In this section, it is primarily aimed to make distinction between coherence and cohesion – two most commonly interchangeably used terms – and then to provide a basis for how they are assessed in writing studies.

Though being an important characteristic of effective writing in terms of connectedness that "refers to all of the links, both explicit and implicit, in a text that make it a unified whole" (Watson Todd et al., 2007), coherence is generally thought to be an abstract and fuzzy term to define exactly and make distinction from other concepts in writing such as cohesion, unity etc. Lee (2002) describes coherence as "the relationships that link the ideas in a text to create meaning for the readers" (p. 135). It is commonly misused with the term of cohesion: whereas cohesion, in simple terms, regards implicit links, coherence refers to the opposite, explicit links (Watson Todd et al., 2007). In other words, whereas cohesion is described as the connection of ideas at sentence level or "the connectivity of ideas in discourse and sentences to one another in text, thus creating the flow of information in a unified way" (Hinkel, 2004, p. 279), being a more broad term coherence is the organization of ideas at discourse level with all elements.

Coherence is, in simple terms, what the reader grabs from the text while cohesion provides the reader with linguistic elements-cohesive devices- to make connection between ideas (Crossley, Kyle, & McNamara, 2016). As stated in their seminal work "Cohesion in English", regarded as a theoretical framework on textual cohesion, Halliday and Hassan (1976) describes cohesion as a semantic concept that illustrates "relations of meaning that exist within a text" (p. 4). They divide cohesion into two main categories- grammatical and lexical cohesion- since like all other semantic systems, it is built through vocabulary and grammar. The devices such as *substitution, reference and ellipsis* are grammatical and *lexical cohesion*, as its name

suggests, is lexical; however, *conjunction*, the fifth kind of cohesive device, is on the borderline of the two. Reference involves three types of ties, personal and demonstrative pronouns, and comparatives. Substitutions exist as nominal (substituting a word), verbal (the verb *"do"),* and clausal (the words *"so"* and *"not"*). Like substitutions, ellipsis also has the same three types – nominal, verbal and clausal ellipsis - and refers simply to "substitution by zero" (p. 142). Lexical cohesion is built through reiteration (repetition, synonym, superordinate, general word) and collocations. The last type of cohesive devices, conjunctions are not cohesive in themselves but indirectly through their meanings.

Similarly, according to Harmer (2004), writers use two main elements to build cohesion in a text- linguistic techniques and grammar structures; in other words, like Halliday and Hassan (1976), he also describes cohesion in two headings- lexical cohesion and grammatical cohesion. Whereas lexical cohesion is achieved through the use of two main devices, repetition of words (repetition of several content words throughout the text) and lexical set 'chains' (words in the same topic interrelating with each other), grammatical cohesion is achieved by the means of "pronoun and possessive reference, article reference, tense agreement, linkers, and substitution and ellipsis". Coherence that enables the reader to catch both "the writer's purpose" and "the writer's line of thought" is far beyond the sentence level and achieved through sequencing information in order to meet the expectations of the discourse community that it is written for (Harmer, 2004, p. 22-25).

However, since the construct of cohesion represents specific features of a text, existence of a number of cohesive devices in a text may not mean anything; that is, a text rich in these devices may still be incoherent and not effective. On the other hand, coherence is challenging to measure in that it is subjective by nature and in contrast to cohesion measured directly from the text it does not exist in the text itself but in people who read and interpret the text (Yule, 2010). Besides, coherence combining all elements of cohesion is related to writing quality. In other words, whilst cohesion can be measured directly and quantified due to the textual elements it posses, coherence can be only measured indirectly through how the reader grabs the text (McNamara, Graeser, McCarthy & Chai, 2014).

As in definitions of coherence and cohesion, the research shows also differences in the ways or measures to assess them. For example, one of the scales applied to assess coherence both in spoken and written discourse is topic-based analysis which depends on identifying key terms in a text, finding the relationships between these terms, ranking these relationships, and then mapping the text along the hieararcy identified through the relationships (Todd, Thienpermpool, & Keyuravong, 2004). In their study, Todd et al. (2004) applied topic-based analysis because it meets the three criteria defined by the researchers to select an appropriate scale to evaluate coherence: it (1) is objective, (2) unequivocally measures coherence, and (3) focuses on propositional coherence that is predominant in written discourse rather than interactional coherence seen in informal spoken language. As a result of their study, they drew a conclusion that although it is easily obtainable and can be thus used to assess coherence, topic-based analysis may be more appropriate for researchers to measure coherence of texts rather than for teachers evaluating coherence in students' essays. Furthermore, Todd et al. (2007) investigated the relationships between connectedness in discourse and comments of tutors on Thais postgraduate students' academic essays. Although they used Hoey's lexical analysis to measure cohesion and genre analysis to assess interactional coherence, they chose topical structure analysis to evaluate propositional coherence.

Similarly, Knoch (2007) reported that the previous scales developed to assess coherence are either too time-consuming or complicated. Therefore, in his study undertaken in three phases as (1) analysis of writing samples, (2) rating scale design and (3) rating scale validation, he chose and adapted a topical structure analysis (TSA) scale with the aim of investigating whether the use of a TSA scale-an empirically-based scale- to evaluate coherence in written production of students is more reliable and has greater discrimination compared to the more traditional measures. However, the results revealed that although raters using the TSA scale scored more accurately, the TSA scale was not less time-consuming than the previous scale; rather, it might require more labor to analyze a large number of written texts and thus not practical in some cases.

McNamara, Crossley, and McCarthy (2009) used Coh-Metrix-an automated tool- to examine whether the quality of the essays- low or high- can be predicted through the three indices as syntactic complexity, lexical diversity and word frequency. In contrast to the general notion that more cohesive and thus more coherent essays are

produced by more proficient writers, their study using linguistic indices of cohesion from Coh-Metrix could not provide any evidence that there is difference between high- and low-proficiency essays in terms of coherence; that is, the essays scored highly were not more coherent than those rated low (McNamara et al., 2009).

According to McNamara et al. (2009), the Coh-Metrix cohesion indices validated by a number of studies are confidential to assess cohesion. In the light of their literature they reached a conclusion that Coh-Metrix is "an extremely powerful text analysis tool, capable of assessing and differentiating an enormous variety of text types from the genre level to the sentence level" (p. 59). Therefore, in their study investigating the degree to which these indices have a role in predicting the quality of essays, they used 26 linguistic indices of cohesion from Coh-Metrix. Similarly, McNamara et al. (2010) point out that Coh-Metrix which is a tool presenting a great variety of linguistic indices for the automatic analysis of text comprehension uses lexicons, latent semantic analysis (LSA), and many other linguistic components and thus meets the needs of researchers who seek a computational linguistic analysis of texts to measure text cohesion and text difficulty in terms of various linguistic features such as word, sentence, paragraph, and discourse dimensions. Futhermore, their study comparing the outcomes of Coh-Metrix indices with two commonly used readability indices – Flesch Kincaid Grade Level and Flesch Reading Ease added evidence on validation of Coh-Metrix as a tool to assess cohesion.

Through the studies on cohesion, we reached the conclusion that we can measure cohesion through the automated tool Coh-Metrix 3.0 that is less time-consuming and more practical. On the other hand, coherence that is more subjective and exists in mind of the reader will be best assessed using an analytical rubric that involves specific dimension of coherence.

2.6. Conclusion

This chapter presented an overall discussion of theoretical background that paves the way for the current study. Besides the discussion of TBLT and two competing models – Robinson's Triadic Framework and Skehan's Limited Attentional Capacity Model -, the studies pertinent to task complexity were briefly presented. Whereas

among the three common measures used in literature, accuracy, and fluency were explained in short, linguistic complexity, also one of the dependent variables of this study, was elaborately handled. Furthermore, the other variables such as coherence and cohesion were also presented both technically and in terms of studies regarding them. Based on these studies, the research design, research questions, and dependent variables to be measured were determined. In this sense, this chapter provides a basis for the next section presenting the methodology of the current study.

CHAPTER THREE

3. METHODOLOGY

3.1. Introduction

This chapter outlines the methodology followed in this study. Beginning with the description of research design, the chapter goes on introducing the participants of the study. Following the presentation of data collection tool, the process in which data were collected is presented in detail. The chapter goes end with the description of both the tools and procedures regarding data analysis.

3.2. Research Design

All issues regarding this study such as data collection procedure, data analysis, and the nature of the research questions clearly build evidence that it is a quantitative study. Our research questions basically addressed the causal effects of the intervention, which is the basis of an experimental study (McMillan & Schumacher, 2006; Muijs, 2004; Tabachnick & Fidell, 2007). Among experimental research designs, we followed a repeated-measures design, "also known as a within-subject design" (Lix & Keselman, 2010, p. 15). In simple terms, in a repeated-measures design all conditions of an experiment are carried out on the same participants (Creswell, 2005; Field, 2012; Johnson & Christensen, 2012). In other words, "all participants are repeatedly measured under each treatment condition" (Johnson & Christensen, 2012, p. 291).

Repeated-measures designs may be dealt with under the title of factorial designs by some researchers (e.g. Field, 2012; Fraenkel, Wallen, & Hyun, 2012) probably due to the fact that in both designs, multiple treatments were administered (Creswell, 2005) and the researcher has the opportunity to study the interaction between two or more other variables (Fraenkel et al., 2012); however, the feature of repeated measures that all

the participants in each experimental condition are the same clearly distinguishes them from factorial designs in which different participants are equally assigned to the groups.

A repeated measure design has some advantages to carry out (Creswell, 2005; Field, 2012; Raykov & Marcoudiles, 2008). For instance, in repeated-measures designs the number of participants in each treatment is equal and there are thus fewer participants compared to factorial designs as the same participants participate in each condition. Furthermore, a repeated-measures design is likely to "result in greater precision of parameter estimates and more efficient inferential analyses" (Lix & Keselman, 2010, p. 15). Another advantage of repeated-measures designs is that they enable the researcher to investigate the growth or maturation of participants through the time effectively. What is more, repeated-measures designs make it possible to keep unsystematic variation which is commonly seen in independent designs at minimum by using the same participants (Field, 2012). Another way of stating this, as long as all other threats such as subject characteristics, history, regression etc. are controlled, repeated-measures designs may be more effective in detecting the effects (Fraenkel et al., 2012). All these features rationalized the repeated-measures design for the current study.

Following the selection of participants generally done through the way of purposive sampling in which the samples are chosen in accordance with a specific purpose (Cohen, Manion, & Morrison, 2005), the researcher in a repeated-measures design defines the treatments, having more than two-levels, to administer separately to the same group. In line with that, we first chose our participants from ELT department of a state university to complete the tasks for our study and performed the treatments in their writing courses run by the researcher herself. Then, the conditions in which the students would perform the tasks were designed. In this sense, based on the fact that in repeated-measures designs there is one or more than one independent variable with at least two levels to be modified to investigate their effects on the dependent variable/s (Creswell, 2005), two independent variables, task complexity and rhetorical task, were used whereas the former was modified at two levels, simple and complex, the later had three levels without any modification. At the end of each treatment, data were obtained from the participants and independently analyzed by the researcher in terms of

dependent variables, which is also a distinct characteristic of repeated-measures designs (Creswell, 2005).

In a repeated-measures design, differences in dependent variables probably result from two main factors: "(1) the manipulation that was carried out on the participants, or (2) any other factor that might affect the way in which a person performs from one time to the next" (Field, 2012, p. 16). However, the former, experimental manipulation is believed to have more clear impact compared to the later. Accordingly, task complexity, the experimental factor of this study, is expected to have greater impact than rhetorical mode of writing on task performance of EFL writers.

Validity of design

Repeated-measures designs give the researcher strength in terms of internal validity since they are not influenced by the threats to internal validity regarding differences between groups such as selection, interactions between groups, maturity, treatments, etc. since the same participants participate in each experimental condition (Creswell, 2005). In addition, error variation may be kept at minimum thanks to unsystematic variation; conversely, if different participants were used in each experimental condition, that will be surely larger (Field, 2012).

Although all these features make a repeated-measures design a powerful experimental design (Johnson & Christensen, 2012; Lawal, 2014), it does not mean it is irreproachable. On the contrary, there are some serious issues to be taken into consideration. For instance, whereas systematic variation makes repeated-measures designs superior to independent designs, it may also lead to some problems such as practice effects and boredom effects (Field, 2012). When the participants participate in each experimental condition, they will thus get familiar with the dependent variable and overcome the problems led by being naive particularly in the first treatment. In addition, after completing the first task, the participants may get bored or tired, which naturally affects the outcome. Another way of stating this, although some distinct features that belong to experimental research such as use of pretest, posttest, control or experimental group separately do not pose a threat for repeated-measures designs, history may be a serious problem since one treatment may affect the following treatment (Creswell, 2005). Therefore, in order to minimize these effects, at least boredom effect, different

activities were carried out between the tasks in two rhetorical modes. That is, after two tasks for one rhetorical mode were performed, new activities generally related to reading and speaking were carried out next two or three weeks before starting to perform tasks in the other rhetorical mode.

3.3. Participants

The participants of our study consisted of forty-one freshmen studying on the ELT department of a state university in Turkey. They were in a context where English is taught as a foreign language and thus have almost no access to produce something outside the classroom. Although the study started with 79 students, the data of only 41 were used. Since a repeated-measures design was applied, it was required to use the writing of the same students for both tasks in each rhetorical modes of writing. Therefore, the data collected from those students (28 students) who did not produce any writing for one of those tasks had to be excluded. In addition, we had to count the writing of 6 students out of our data which were defined as "very poor" according to the results of raters' scores. Furthermore, we could not use 4 of them as automated programs used to analyze the writing in terms of complexity and cohesion gave a warning about not using the data. For instance, those essays for which Coh-Metrix 3.0 provided the following note "You have entered fewer than 200 words. For more meaningful results, we advise entering more than 200 words of the text…" failed to be included among the data of our study. In conclusion, our participants were 41 ELT students whose ages ranged between 19-28 years and sharing the same L1-Turkish. The number of female students (33 students) is almost forth times more than that of male students (8 students).

The proficiency levels of students were generally intermediate whereas there were also a few students at advanced level, as also understood particularly from their scores of general writing achievement. Moreover, although 35 students were at their second year as they had prep-class (included writing course) in the previous year, 6 of the participants were new comers who did not take writing course before. Before collecting data, all of the students were provided basic training for advanced writing and the essays written by these students before main tasks were analyzed by the researcher to see their levels and proficiency in writing. Through these results, it was regarded that

those new comers were at the same knowledge level with other students who were more experienced in terms of writing.

Each participant wrote for both types of task –simple and complex- in each rhetorical task, descriptive, narrative, and cause-and-effect essays; that is, totally six essays were obtained from each student. In all, we had 246 essays to analyze and use in our study. Furthermore, all participants signed a consent form (see Appendix F) showing that they do not mind if we use their written production for research purposes.

3.4. Operationalizations of Task Complexity and Rhetorical Mode

Our first independent variable, task complexity was operationalized at two levels as simple and complex. Based on Robinson's the Triadic Componential Framework, the first writing task of each rhetorical mode was identified as simple as a consequence of availability of time specifically allotted for planning, the second task carried out following the first task in each rhetorical mode was described as complex since it had no specific time for planning and thus writers had more cognitive load in on-line process of writing. In other words, whereas simple tasks of the current study were the writing tasks conducted through strategic planning, complex tasks were those conducted through careful on-line planning that has no time pressure.

As for the second independent variable of this study, rhetorical mode was presented at three levels: descriptive, narrative, and cause-and-effect. Descriptive essay is a kind of writing mode in which writers try to create a vivid image of something – an object, a person, or a feeling – through words. In this sense, in both tasks of descriptive essays the students were asked to describe a city. For the second rhetorical mode, narrative essay in which the writer tells or narrates a story, the students wrote a story in the light of pictures they were given. In a cause-and-effect essay, the writer presents causes (reasons) for something, effects (results), or both causes and effects at the same time (Oshima & Hogue, 2006). Accordingly, while the students were required to propose the causes in complex task of the third rhetorical mode, they were asked to write about just the effects in the simple task. Writing process and details of writing instructions for each rhetorical mode and task are explicitly presented in the following section.

3.5. Data Collection Procedure

Data were collected in the first term of 2015-2016 academic year during Advanced Reading and Writing I course the researcher herself conducted. After one-month writing training that involved basic instructions for essay writing and during which the students also wrote sample paragraphs and one essay, the tasks were carried out and data collection thus began. Two kinds of writing tasks described as simple and complex were given to participants for each rhetorical mode. Whereas simple tasks were applied as in-class writing through strategic planning, complex tasks were conducted through on-line planning without having time pressure. In this respect, the students were required to do pre-task planning for the first task in each rhetorical mode and hand in their outlines or first drafts indicating that they did pre-task planning; on the other hand, while conducting the complex task, they needed to present just their final drafts they completed by having neither time pressure nor obligation to make planning.

Since they had more opportunities to reach sources – both online and written- and thus obtain information about the topics while carrying out the simple tasks, particularly for the first and third rhetorical tasks, we strictly warned them about plagiarism both orally before each task and in written through the syllabus issued at the first week. Accordingly, first three written paper of each students were analyzed by a plagiarism detection software to see whether they used the sources without giving citation and to show that the plagiarizing students would not be tolerated but presented with the result clearly stated in the syllabus "…you will be given a 0 with no chance to redo the assignment or test".

Writing prompts for each writing task was explicitly presented in Table 3.2. As seen in the table, similar topics for both simple and complex tasks in each rhetorical task were chosen in order to avoid the effect of topic familiarity which is believed to have impact on writing performance of students (Stapleton, 2001; Tedick, 1988) as much as other skills of second language, speaking (Rahimpour & Hazar, 2007), listening (Salahshuri, 2011; Schmidy-Rinehart, 1994), and reading (Leeser, 2004; Shimoda, 1993).

For the first task, simple task of descriptive writing, the students were asked to describe their favorite cities in the light of the writing prompts provided in Table 3.2.

Just after the topic was given, the researcher and participants discussed and talked about it for a few minutes with the purpose of providing a general idea before writing. Following the discussion and think-aloud protocols, the students were first given 15 minutes to make planning and prepare an outline of what they would write and then 45 minutes to complete their writing production. After the expiration of 60 minutes, they submitted in their writing paper with the outline paper they prepared during strategic planning. Completing their first task –simple task- carried out through strategic planning, the students were to perform the second task –complex task- for the same rhetorical task. They were required to produce a descriptive essay in which they would describe a city they visited before or liked the most in accordance with writing prompts presented in Table 3.1. They did have no time limitation to complete their writing and were also free to submit their writing productions in five days.

Similarly, the simple task of the second rhetorical mode –narrative writing- was carried out in class through strategic planning. In this sense, a picture (in Appendix E) was illustrated and the students were asked to look at it for five minutes. After discussing for a few minutes, they had again 15 minutes to make planning of the content and then 45 minutes to narrate the picture and complete their writing task. In the simple task, the students were given 16 related pictures involving the same characters and also the same scene as the picture illustrated in the simple task to create their own story with no time limitation nor any obligation to do pre-task planning.

As for the third rhetorical task, the students were asked to perform a timed cause-and-effect writing in class on the causes of overuse of internet. Following 15-minute strategic planning, the students were similarly given 45 minutes to complete the simple task of the last rhetorical mode. Without making a great change on topic, the students were asked to write an essay on the effects of overuse of internet and submit it when they completed it in five days.

Table 3.1.

Writing Prompts for Each Rhetorical Task

Planning/ Rhetorical Modes of Writing	Simple (Pre-task planning)	Complex (Unpressured/Careful On-line planning)
Descriptive	Describe one of your favourite cities. -Describe the main feature of that city. -Why do you regard it to be your favourite city?	Describe a city you have visited before/ you like the most. - Describe the city in detail - When did you visit / Why do you like it? - What are the attractive places of the city? - What can you do in this city?
Narrative	Look at the picture on the board for five minutes and then create your own story for that picture	Look at the 16 related-pictures in paper you were given. Create a story that reflects the events and characters you see in pictures.
Cause-and-Effect	Write an essay on the reasons of overuse of internet	Write an essay on the effects/results of overuse of internet.

3.6. Data Analysis

Although it is believed to be an important aspect of communication and predictor of academic success (Crossley & McNamara, 2010; Crowhurst, 1990), writing seems to be more complex skill among four basic language skills (Liu & Braine, 2005) particularly in terms of assessment since it may lead to reliability problems inherent in writing rather than related to rating (Hamp-Lyons, 1993, 2003). As much as deciding which aspect of writing to be assessed, making a decision on the way of assessment – whether to use an automated program or evaluate manually – is also another challenge for writing teachers and researchers.

In line with the literature, it was first decided to evaluate five main dimensions of L2 writing – syntactic and lexical complexity, coherence, cohesion, and general writing achievement. Different automated tools for the analysis of linguistic complexity and cohesion were applied since they seemed to make the evaluation process less

laborious but more time-saving in addition to being more objective (Crossley & McNamara, 2009; Lu, 2008). The most striking point of the automated tools used in this study is that they all are freely available and easy to use. On the other hand, we manually evaluated the writing of our participants in terms of coherence and overall quality using an analytical rubric that consists of five dimensions since they are more subjective and refer to what the reader grabs from the text.

3.6.1. Syntactic complexity analyzer

In order to evaluate syntactic complexity of students' writing, Lu's "Web-based L2 Syntactical Complexity Analyzer" (L2SCA) was used. The system is a useful tool that enables both second language teachers and researchers to analyze the syntactic complexity of samples written in English language under two main titles- syntactic structure and syntactic complexity indices. The former deals with nine structures, word count (WC), sentence (S), verb phrase (VP), clause (C), T-unit (T), dependent clause (DC), complex T-unit (CT), coordinate phrase (CP), and complex nominal (CN). The later, syntactic complexity, involves 14 different indices, mean length of clause (MLC), mean length of sentence (MLS), mean length of T-unit (MLT), and clauses per sentence (C/S) that cover length of production units and sentence complexity, coordinate phrases per clause (CP/C), coordinate phrases per T-unit (CP/T), T-units per sentence (T/S) showing amounts of coordination, clauses per T-unit (C/T), complex T-units per T-unit (CT/T), dependent clauses per clause (DC/C) reflecting amounts of subordination, complex nominals per clause (CN/C), complex nominals per T-unit (CN/T), verb phrases per T-unit (VP/T) illustrating degree of phrasal sophistication and overall sentence complexity. However, as illustrated in Table 3. 2, just 9 of 14 indices were used since written productions of students were not advanced enough to be measured by all indices.

Table 3.2.

Nine of the Fourteen Syntactic Complexity Measures Automated (Lu, 2010, p. 6)

Measure	Code	Definition
Type 1: Length of production		
Mean length of clause	MLC	# of words / # of clauses
Mean length of sentence	MLS	# of words / # of sentences
Mean length of T-unit	MLT	# of words / # of T-units
Type 2: Sentence complexity		
Clauses per sentence	C/S	# of clauses / # of sentences
Type 3: Subordination		
Clauses per T-unit	C/T	# of clauses / # of T-units
Complex T-units per T-unit	CT/T	# of complex T-units / # of T-units
Dependent clauses per clause	DC/C	# of dependent clauses / # of clauses
Dependent clauses per T-unit	DC/T	# of dependent clauses / # of T-units
Type 4: Coordination		
T-units per sentence	T/S	# of T-units / # of sentences

Technically, the system runs on UNIX-like (LINUX, MAC OS, or UNIX) systems with hardware involving a 750 MHz Pentium III processor or better and 2 GB or more RAM and consists of three components – Stanford parser, Tregex, and the syntactic complexity analyzer. It takes a written sample in English in plain text format and analyzes it using the measures explained above. Analysis process takes place in two stages: (1) reprocessing stage at which syntactic structures of a sentence is analyzed and (2) the syntactic complexity analysis stage at which nine production units and syntactic structures are first retrieved and counted, and then the counts of those incidences are used to calculate the syntactic complexity of the written sample.

In order to give insight for the indices used in this study, basic constructs are explained briefly below.

Sentence: A sentence, constructed in line with “language-dependent rules” in terms of grammar, content, and intonation (Bussmann, 2006, p. 1059) is basically a complete expression of a thought - an assertion, a question, a command, a wish, an exclamation, or the performance of an action ("Merriam-Webster - Online Dictionary").

Clause: Described as “a structure with a subject and a finite verb” (Hunt, 1965, p. 15), a clause involves two major types – independent clauses and dependent clauses.

Dependent clause: In contrast to independent clauses, a dependent clause, "another name for subordinate clause" (Trask, 2014, p. 66) cannot stand by itself but depends on a main clause to construct a sentence. Dependent clauses include three types such as adjective clauses, noun clauses, and adverbial clauses (Tercanlıoğlu, 2000).

T-unit: Also known as "Minimal Terminable Unit" or "the Terminable Unit", T-unit is a measure of syntactic complexity and consists of "one independent clause together with whatever dependent clauses are attached to it" (Richards & Schmidt, 2010, p. 613). Two or more T-units exist in a compound sentence (complex sentence).

Complex T-unit: If a T-unit involves a dependent clause, it is called as complex T-unit (Casanave 1994 cited in Lu, 2010)).

Coordinate phrase: It is a word or a group of words that connect words or other constructions (Crystal, 1992).

Coordination: It is a term that refers to the linking linguistic units usually with equivalent syntactic status, such as clauses, phrases or words (Crystal, 1992, 2003). Its feature of linking the things with equal grammatical status makes it distinct from subordination in which "one clause functions as part of another" (Crystal, 2010, p. 99). Those markers illustrating linkage are called as coordinating-conjunctions or coordinator.

Subordination: It links the linguistic units with syntactically different status by being dependent upon the other or a being a component of the other (Crystal, 1992). For instance, a subordinating clause may depend on a main verb in terms of word order, tense agreement, and modality besides illocution (Bussmann, 2006).

3.6.2. Lexical complexity analyzer

We analyzed lexical complexity of the texts written by the students through Lu's Lexical Complexity Analyzer (LCA) designed for that purpose using 25 different measures of lexical density, variation and sophistication. The system technically runs on like L2 Syntactical Complexity Analyzer (L2SCA). We used just four main indices, lexical density (LD), lexical sophistication (LS1), type-token ratio (TTR), and lexical

word variation (LDV). The mean of TTR and LDV was submitted as lexical variety. All indices were explained in detail below.

Lexical Density: A measure of the difficulty of a text, lexical density measures the proportion of lexical words (content words such as nouns, adverbs etc.) to the total number of words (Johansson, 2009, p. 65).

Lexical density is normally expressed as a percentage and is calculated by the following formula (Williamson, 2014)

$$\text{Lexical Density}=\frac{\text{number of lexical words}}{\text{total number of words}} \times 100$$

Lexical Sophistication: Also called as 'lexical rareness', lexical sophistication is "a measure of the proportion of relatively unusual or advanced words in the learner's text" (Read, 2000, p. 203).

Lexical Variety: Lexical variety, though used interchangeably with lexical density that is a measure of the proportion of lexical items, is "a measure of how many different words are used in a text" (Johansson, 2009, p. 61). A text with high lexical density does not mean that its lexical variety is also high. That is, a text may contain various words that show its lexical variety, however, it may include less lexical words such as nouns or adjectives but more pronouns, which makes that text less lexically dense. Besides lexical word variation, we also used the indices of TTR (type-token ratio) to measure lexical variety.

Type-token ratio: A traditional measure of lexical variety, TTR is the ratio between word types and tokens which refer to the number of words. If the results of TTR is high, the ratio of repetitive words is low (Richards, 1987). In addition, a text with large numbers of tokens results in lower value for TTR or vice versa, which implies that longer texts may have lower values than shorter texts have (Johansson, 2009). In this sense, it is more logical to use TTR when analyzing the texts with equal length.

Table 3.3.

Automated Lexical Complexity Measures Used in This Study (Lu, 2012, pp. 193,195)

Measure	Code	Definition
Lexical Density	LD	*Nlex/N*
Lexical Sophistication	LS	*Nslex/Nlex*
Lexical Variety		
Lexical Word Variation	LV	*Tlex/Nle x*
Type–Token Ratio	TTR	*T/N*

3.6.3. Cohesion

As clearly stated in Chapter 2, it is possible to measure cohesion directly due to its basic characteristic that is to be in existence in the text itself and quantitatively through textual elements in the text; however, coherence that exists in the mind of the reader is measured indirectly through the comprehension of the reader (McNamara et al., 2014). Although there are many ways of measuring cohesion, some of which are automated and some of which depend on manually analysis of cohesive devices, it is quite challenging to decide which one to apply to assess cohesion of the students' writing. Like measurement of other dependent variables, syntactic and lexical complexity, cohesion was assessed through an automated evaluation program, Coh-Metrix 3.0 which was developed based on the assumption that "(1) cohesion is in the text and (2) cohesion can be computationally measured" (McNamara et al., 2014, p. 20). Furthermore, although Coh-Metrix is originally developed to provide measures for a variety of cohesive devices, it seems to be used in a great deal of research as it is a comprehensive tool that enables to analyze texts at various language and discourse levels.

As in complexity analyzer software programmes, we used the term "index" rather than "measure" in accordance with the programme. To show the difference between these two terms, McNamara et al. (2014) propose that the terms "index" or "indices" is used "to describe any one of the ways Coh-Metrix assesses" (p.61) one measure which is described as a theoretical construct (e.g., lexical diversity, deep cohesion, word concreteness). Among five components of text easability, just two directly related to cohesion, referential and deep cohesion, were used in this study (McNamara et al., 2014, p. 85). These terms were briefly explained below:

Referential Cohesion: Referential cohesion in a text illustrates the extent of connections that link the ideas together to help the reader process easily.

Deep Cohesion: A text with high deep cohesion contains causal and intentional cohesive devices where they are required to establish relationships among ideas, events, or actions since these devices enable the reader to understand the relationships more deeply and coherently.

3.6.4. Coherence

Since coherence is a subjective feature of writing and simply what the reader grabs from the text (Crossley et al., 2016), we analyzed it through the same analytic rubric also used for the analysis of the students' general writing achievement. The rubric we used included a separate section to evaluate coherence under the name of organization and coherence. As in other sections of the rubric, the scores ranged between 1 and 5. If a text takes the maximum score of 5 for coherence, it means that the text uses a logical structure regarding the purpose, audience, subject of the paper, utilizes true and enough transitions to build clear connection between sentences, and lead the reader to comprehend the chain of reasoning or progression of ideas. On the contrary, if it is given the minimum score of 1, the text is lack of an organization, coherence, and transitions.

3.6.5. General writing achievement

Our last dependent variable to be measured was the general writing achievement of the participants. In line with that purpose, we preferred an analytic rubric since use of analytic rubrics provides more reliable scores through exemplars and/or rater training (Jonsson & Svingby, 2007) and provide more-detailed information about writers' performance (Weigle, 2002). We reached the rubric titled "Example of a Grading Rubric For a Term Paper in Any Discipline" and designed by "the UC Davis English Department Composition Program" online. The rubric consisted of five sections as ideas, organization and coherence, support, style, and mechanics. The scores for each section ranged between a minimum score of 1 and a maximum score of 5. With the

purpose of providing inter-rater reliability, all texts produced by participants were evaluated by two different raters, one of which is the researcher herself.

3.6.6. Statistical analysis

The results of automated tools –Lu's Lexical and Syntactic Analyzers, and Coh-Metrix 3.0- and analytical rubrics were computed into a statistical program, IBM-SPSS 21. We had only a group, each subject of whom was exposed to all combinations of two independent qualitative (categorical) variables, task complexity and rhetorical task. Therefore, it was required to use a two-way repeated measures design which arises when the same subjects are exposed to all conditions of an experiment or when they are observed at multiple time points (Field, 2012; Lix & Keselman, 2010) to see the effects of these independent variables "referred to as repeated-measures factors or within-subjects factors" (Green & Salkind, 2004, p. 238). Repeated measure designs have obvious advantages such as studying with fewer participants per experiment or not being affected by "between-subjects differences for experimental error" (Ho, 2006, p. 117). Though violating the basic assumption of multivariate analysis –independence-, they still test for differences among participants in terms of various dependent variables. Furthermore, in a repeated-measure design in which the same entities participate in all experimental conditions, "the test for treatment effects is not on experimental error, but on within subject error" (Bonate, 2000, p. 119).

While two-way repeated measures ANOVA was conducted for the two dependent variables, general writing achievement and coherence, separately, two-way repeated measures MANOVA was carried for the other three dependent variables, lexical complexity, syntactic complexity, and cohesion, as they involved subcomponents, each of which was considered and assessed independently. Furthermore, paired-samples t-test was carried out for the variables found to have statistically significant effect to see the differences between the variables, controlling familywise error rate using Holm's sequential Bonferonni approach (Green & Salkind, 2004).

3.7. Reliability and Validity

Manual analysis of so many texts (totally 241 essays) would be more time-consuming and unreliable without having rater agreement. However, it was so difficult to find another trained rater to spare time to assess so greater number of written texts in terms of cohesion and linguistic complexity. Furthermore, an automated tool would provide more reliable and fast available results. In this respect, using an automated-tool designed for the evaluation of these dimensions stood to reason. Following analysis of each essay separately by the programs, 30 randomly chosen essays were re-analyzed by all three of analyzer program, Lu's Lexical Complexity Analyzer and L2 Syntactic Complexity Analyzer, and Coh-Metrix 3.0 for cohesion, to see whether they provide the same results. It was seen that there was no difference between the previous and the later results.

In the light of the literature, it was agreed to use Coh-Metrix 3.0 that provides automated evaluation of text. Furthermore, the validation of indices in this program was verified by McNamara, Louwerse, McCarthy, and Graesser (2010) that those indices measure what is expected to be measured and can be compatible with all types of data regarding human performance.

For the evaluation of syntactic complexity, an automated tool was used for aforementioned reasons. In the light of studies in the literature investigating measures of syntactic complexity, Lu's L2 Syntactic Complexity Analyzer was considered as more appropriate for the current study. Firstly, it was easy and free to use. Besides, its reliability was experimentally tested in the study of Lu (2010). Two trained annotators first independently labeled the units and structures in 10 of the 40 essays and had a high inter-annotator agreement on the identification of those structures and units. Moreover, according to the results provided by annotators, the system also achieved a high degree of reliability ranging from .830 to 1.000 in identifying relevant units and structures.

The other two dimensions in the current study, coherence and general writing achievement not rated by the computer analysis were evaluated by another rater to ensure the reliability of coding data. Firstly, 30 essays randomly chosen were rated by three raters, one of whom is the researcher herself, in order to test the reliability of rubrics. The raters were trained about what the dimensions involved in the rubric ask

and how to score those dimensions. After reaching a .87 inter-rater reliability, with a Cronbach's alpha coefficient reported of a high reliability level, the two raters went on analyzing the rest.

Following the evaluation of all essays, the results for coherence and overall writing quality to see whether there was an inter-rater reliability were retested. Although the level of inter-rater reliability (Cronbach's alpha coefficient=.88) was not low, 34 essays in different rhetorical modes having more than 2-point difference in coherence or 5-point difference in overall writing quality were reread and rerated to avoid extreme scores. As a result, we reached a .89 inter-rater reliability, with a Cronbach's alpha coefficient reported of a high reliability level.

CHAPTER FOUR

4. RESULTS AND DISCUSSION

4.1. Introduction

This chapter presents the results for the effect of rhetorical mode and task complexity on linguistic complexity, coherence, cohesion, and overall writing quality respectively, answering the research questions of this study. Following presentation of statistical results, the main findings are discussed in the light of relevant literature. In each figure and table in this section, whereas task 1 refers to the simple task that was carried out through strategic planning, task 2 points out the complex task conducted under careful (unpressured) on-line planning. As to the rhetorical mode, 1 refers to descriptive writing, 2 to narrative writing, and 3 to cause-and-effect writing.

In order to see the effects of two independent variables, task complexity and rhetorical mode, two-way repeated measure MANOVA was carried out for the three dependent variables, syntactic complexity, lexical complexity, and cohesion, since they have more than one dimension. For the other two dependent variables, two-way repeated measure ANOVA was conducted. While presenting statistical results, besides main effects of independent variables, results for the interaction effect between these variables are also presented. Furthermore, when statistical results showed a significant effect of any variable, paired-samples t-test was also conducted to assess the differences in results, controlling familywise error rate using Holm's sequential Bonferonni approach (Green & Salkind, 2004).

4.2. Results for Syntactic Complexity

4.2.1. Effects of rhetorical mode and task complexity on syntactic complexity

Students' essays were analyzed through Lu's L2 Syntactic Complexity Analyzer (L2SCA) in terms of two dimensions: syntactic structure and syntactic complexity.

Results of univariate and multivariate tests for descriptive statistics were separately displayed for the two dimensions.

4.2.1.1. Effects of rhetorical mode and task complexity on syntactic structure

Syntactic structure for which analysis program provided such linguistic constructs as word count, sentence, verb phrase, clauses, t-unit, dependent clause, complex t-unit was obtained through the means of those constructs. In order to see the effects of task complexity and rhetorical task on syntactic structure of students' essays, we conducted two-way repeated measures ANOVA. In this sense, multivariate test results indicated a significant main effect of rhetorical mode, Wilk's Λ=.30, F (2, 39) = 45.61, p = .00, η^2_p=.70, and task complexity, Wilk's Λ=.73, F (1, 40) = 15.07, p = .00, η^2_p=.27. However, there was no interaction effect between task complexity and rhetorical modes of writing, Wilk's Λ= .93, F (2, 39) = 1.41, p = .26, η^2_p=.06. In other words, although the results of syntactic structure showed significant difference among rhetorical tasks and between simple and complex tasks, syntactic structure of essays written by students was not influenced by interaction of rhetorical task and task complexity. Furthermore, main effects of both rhetorical task and task complexity had a large effect size. The results were displayed in the table below.

Table 4.1.

Results of Repeated-Measures ANOVA for Syntactic Structure

Effect	Wilks' Lambda Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
rhetorical mode	.299	45.61	2	39	.000*	.701
task complexity	.726	15.07	1	40	.000*	.274
rhetorical mode X task complexity	.933	1.41	2	39	.257	.067

*Significant effect is reached

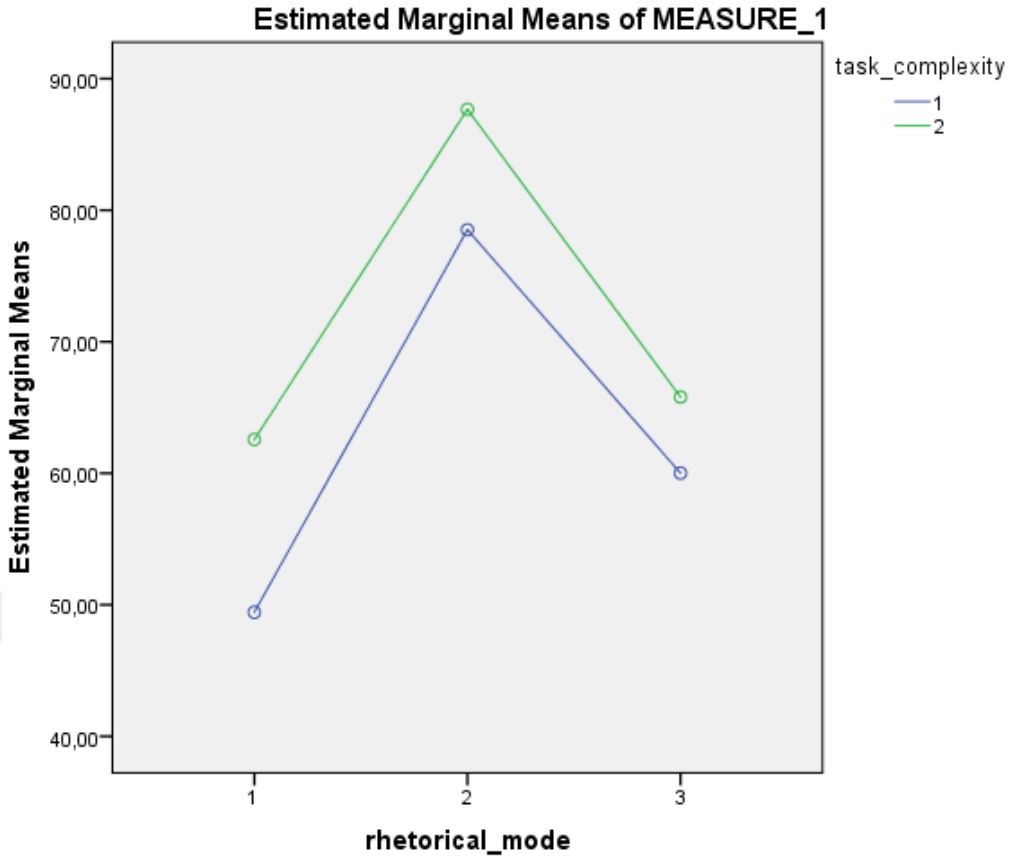


Figure 4.1. Interaction graph for syntactic structure

According to Figure 4.1, there was no interaction between rhetorical modes of writing and task complexity in terms of syntactic structure. As seen in both Figure 4.1 and Table 4.1 while complex task in narrative writing (M= 87.67, SD= 28.9) took the highest score for syntactic structure, the lowest score of syntactic structure belonged to the simple task of descriptive essay (M= 49.44, SD= 15.03). Furthermore, it is clear that complex tasks in all rhetorical modes of writing – descriptive, narrative; and cause-and effect - had significantly higher scores for syntactic structure than simple tasks in the same rhetorical modes did, $t(40)= 3,88$, $p<.05$.

As statistically stated above, rhetorical mode, whether the text was produced in descriptive, narrative or cause-and-effect writing, also had significant impact on the results of syntactic structure. The results of three-paired samples t test carried out to see the effect of rhetorical mode revealed that the students' narrative writing had higher results for syntactic structure compared to their descriptive, $t(40)= -9.67$, $p=.00$ and cause-and-effect writing, $t(40)=6.34$, $p=.00$. Similarly, the results of syntactic structure

for cause-and-effect writing were found higher than those for their descriptive essays, t(40)=2.76, p=.01; that is, descriptive writing of the students had the lowest mean scores for syntactic structure (M=56, SD=14.05).

4.2.2. Descriptive results for syntactic structure

The descriptive results – means and standard deviations – for syntactic structure that was obtained through the means of structural components such as word count, sentence, verb phrase, clause, T-unit, dependent clause, and complex T-unit were also displayed in Table 4.2.

Table 4.2.

Means and Standard Deviations for Syntactic Structure

	Descriptive Essay		**Narrative Essay**		**Cause-and-Effect Essay**	
	Task 1 Mean (SD)	Task 2 Mean (SD)	Task 1 Mean (SD)	Task 2 Mean (SD)	Task 1 Mean (SD)	Task 2 Mean (SD)
Syntactic Structure	49.44 (15.03)	62.56 (19.77)	78.52 (19.12)	87.67 (28.9)	60.01 (10.98)	65.79 (18.63)

Those components that constructed syntactic structure were also illustrated in Table 4.3 separately for the two tasks in each rhetorical mode of writing.

Table 4.3.

Means and Standard Deviations for Components of Syntactic Structure

		Descriptive Essay		**Narrative Essay**		**Cause-and-Effect Essay**	
		Task 1 Mean (SD)	Task 2 Mean (SD)	Task 1 Mean (SD)	Task 2 Mean (SD)	Task 1 Mean (SD)	Task 2 Mean (SD)
SYNTACTIC STRUCTURE	WORD COUNT	222.44	291.1	337.95	373.22	275.95	305.02
		(70.89)	(93.49)	(83.14)	(122.40)	(51.89)	(87.89)
	SENTENCE	20.61	25.34	32.56	39.4	24.37	26.4
		(8.23)	(11.08)	(10.14)	(14.72)	(6.55)	(8.21)
	VERB PHRASE	35.20	41.85	61.56	71.93	44.10	46.14
		(10.68)	(13.55)	(15.84)	(24.16)	(9.47)	(13.3)
	CLAUSE	30.78	36.44	53.46	58.41	34.66	37.15
		(10.05)	(11.91)	(14.52)	(21.46)	(6.92)	(11.3)
	T-UNIT	22.39	27.78	38.24	44.88	26.32	28.46
		(8.91)	(10.93)	(10.80)	(16.93)	(6.60)	(8.52)
	DEPENDENT CLAUSE	8.32	8.49	14.05	13.59	8.12	8.56
		(4.46)	(4.97)	(7.31)	(6.02)	(3.68)	(4.59)
	COMPLEX T-UNIT	6.32	6.95	11.80	12.22	6.59	8.85
		(3.31)	(3.61)	(5.36)	(4.98)	(2.65)	(6.73)

Table 4.3 pointed out that complex task of narrative essay took the highest score for all components, word count, sentence, verb phrase, clause, T-unit, dependent clause, and complex T-unit, whereas the simple task of descriptive essay got the lowest. In addition, complex tasks in the three rhetorical tasks took higher scores for each structural component than simple tasks in those modes of writing.

4.2.1.2. Effects of rhetorical mode and task complexity on syntactic complexity

As we considered syntactic complexity under four components as length of production, sentence complexity, coordination, and subordination each of which was measured independently, two-way repeated measures MANOVA was conducted to analyze the effect of task complexity and rhetorical mode on those components of syntactic complexity. According to the multivariate test, though rhetorical mode had a significant main effect, Wilk's Λ=.25, F (8, 33) = 12.30, p = .00, η^2_p= .75, there was nonsignificant main effect of task complexity, Wilk's Λ=.89, F (4, 37) = 1.15, p = .35, η^2_p= .11. In addition, no interaction effect between task complexity and rhetorical modes of writing, Wilk's Λ= .81, F (8, 33) = 1.00, p = .45, η^2_p= .20, was observed. In other words, there was just difference among rhetorical modes but not between simple and complex tasks in those rhetorical modes of writing in terms of syntactic complexity; moreover, the interaction between rhetorical task and task complexity had no influence on the results of syntactic complexity of students' essays.

Table 4.4.

Results of Two-Way Repeated-Measures MANOVA for Syntactic Complexity

Effect	Wilks' Lambda Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
rhetorical mode	.251	12.30	8	33	.000*	.749
task complexity	.889	1.15	4	37	.349	.111
rhetorical mode X task complexity	.805	1.00	8	33	.453	.195

*Significant effect is reached

Since Mauchly's test results for each component were violated, Greehouse-Geisser results were submitted. However, results of sphericity assumption were presented for the interaction effect on length of production as they met the results of Mauchly's test. In line with those results, although rhetorical task had significant effect on length of production, F(1.18, 47.31)= 7.46, p= .01, η^2_p =.16, it had no effect on any other components, coordination, F(1.05, 42.07)= .25, p= .63, η^2_p =.01, sentence complexity, F(1.09, 43.6)= 2.22, p= .14, η^2_p =.05, and subordination, F(1.18, 47.32) = 1.26, p = .28, η^2_p =.03. As for the effect of task complexity, simple tasks showed no significant difference from complex tasks in terms of length of production, F(1, 40) = .40, p = .53, η^2_p =.01, coordination, F(1, 40) = .69, p = .41, η^2_p =.08, sentence complexity, F(1, 40) = 1.76, p = .19, η^2_p =.04, and subordination, F(1, 40) = 1.57, p = .22, η^2_p =.04. Similarly, rhetorical task interaction by task complexity did have significant effect on none of the components of syntactic complexity, length of production, F(2,80) = .87, p = .43, η^2_p =.02, sentence complexity, F(1.4, 55.89) = .96, p = .36, η^2_p =.02, subordination, F(1.71, 68.46) =1,77, p = .18, η^2_p =.04, and coordination, F(1.09, 43.63) = .92, p = .35, η^2_p =.02.

Table 4.5.

Univariate Test of Within-Subject Effects

Source	Measure	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
rhetorical mode	length of production	205.297	1.183	173.557	7.464	.006*	.157
	sentence complexity	2.541	1.090	2.331	2.219	.142	.053
	subordination	.176	1.183	.149	1.255	.276	.030
	coordination	.223	1.051	.212	.250	.632	.006
task complexity	length of production	4.075	1	4.075	.395	.533	.010
	sentence complexity	1.136	1	1.136	1.755	.193	.042
	subordination	.215	1	.215	1.568	.218	.038
	coordination	.299	1	.299	.699	.408	.017

Table 4.5. *Continued*

	length of production	4.075	1	4.075	.395	.533	.010
rhetorical mode X task complexity	sentence complexity	1.136	1	1.136	1.755	.193	.042
	subordination	.215	1	.215	1.568	.218	.038
	coordination	.299	1	.299	.699	.408	.017

4.2.1.2.1. Effects of rhetorical mode and task complexity on length of production

It is clear in the interaction graph below indicating the results for length of production that there is no interaction between rhetorical task and task complexity. Furthermore, getting almost similar scores, both simple and complex tasks of descriptive essay had the highest scores for length of production (M=10.92, SD=6.8), though the complex task of narrative essay got the lowest score (M=8.44, SD=1.82). In addition, results of length of production for complex tasks (M=9.84, SD=1.86) in all rhetorical modes of writing were lower than those of complex tasks (M=10.1, SD=3.79) in the same rhetorical tasks; however, the difference between simple and complex task is statistically nonsignificant; t(40)=.63, p=.53. That is, whether writing performances were carried out under strategic planning or careful on-line planning did have no effect on the length of students' writing productions.

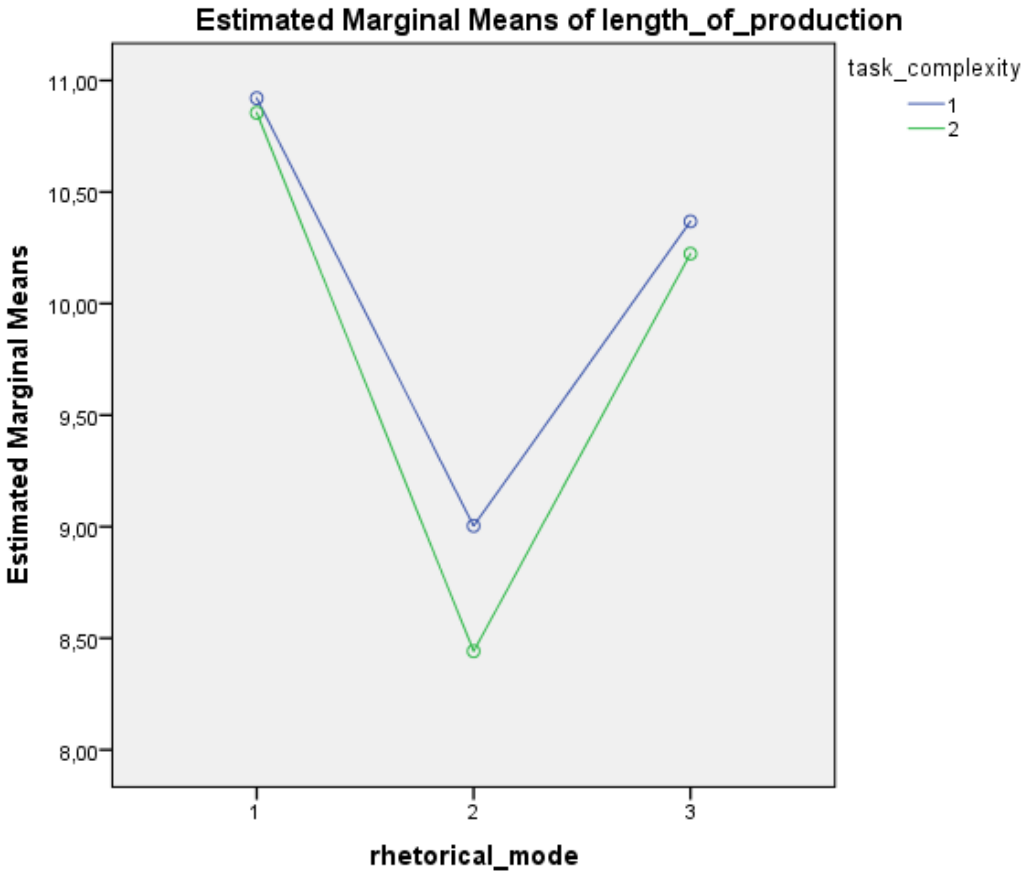


Figure 4.2. Interaction graph for length of production

Rhetorical tasks showed significant difference from one another in terms of length of production with a large effect size. Particularly, there was a significant difference between descriptive and narrative essays, t(40)= 3.15, p=.00, and, likewise, between narrative and cause effect essays, t(40)=-9.06, p=.00, but not between descriptive and cause-and-effect essays, t(40)=.57, p=.57. It was found that narrative essays (M=8.72, SD=1.77) were written shorter in terms of production than both descriptive (M=10.89, SD=5.32) and cause-and-effect essays (M=10.55, SD=2.41). Namely, whereas the shortest production was seen in students' narrative writing, the longest production was in their descriptive writing.

4.2.1.2.2. Effects of rhetorical mode and task complexity on sentence complexity

Figure 4.3 displays the results of sentence complexity for simple and complex tasks in the three rhetorical modes of writing. It is clear that there is no interaction between rhetorical task and task complexity. In addition, as in length of production, simple tasks in all rhetorical tasks (M=1.75, SD=1.49 for descriptive essay; M=1.76, SD=.69 for narrative essay; M=1.48, SD=.37 for cause-and-effect essay) had higher complexity scores in terms of sentence compared to complex tasks in those essays (M=1.64, SD=.92; M=1.51, SD=.28; M=1.43, SD=.23 respectively); however, such difference between complex and simple tasks was not significant, t(40)=1.33, p=.19. Whereas the highest score for sentence complexity belonged to simple task of narrative essay, complex task of cause-and-effect essay had the lowest score in terms of sentence complexity.

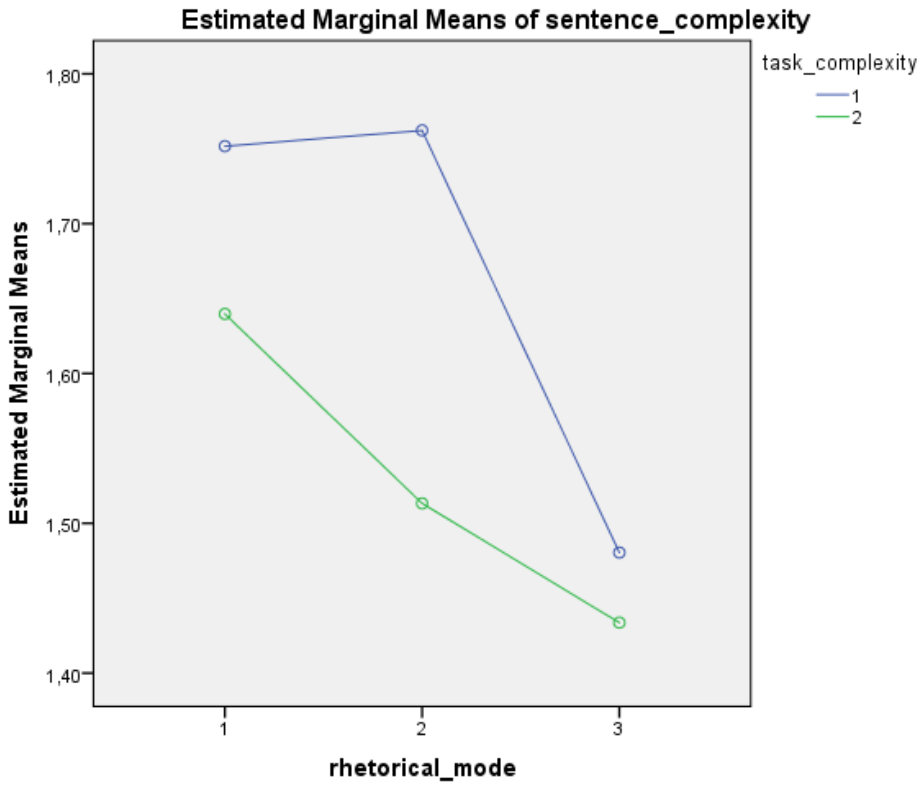


Figure 4.3. Interaction graph for sentence complexity

As statistically stated above, although sentence complexity was not affected by the rhetorical task, there was a significant difference just between narrative and cause-and-effect essays, $t(40)=4.22$, $p<.05$. In other words, compared to cause-and-effect essay ($M=1.51$, $SD=.33$), more complex sentences were found in students' narrative writing ($M=1.63$, $SD=.38$). Furthermore, it is clear in the figure above that there is a linear decrease in sentence complexity of complex tasks among rhetorical modes from the first (descriptive) to the third (cause-and-effect). However, such difference was not significant.

4.2.1.2.3. Effects of rhetorical mode and task complexity on subordination

As seen in Figure 4.2 and Figure 4.3 that display the results of length of production and sentence complexity, Figure 4.4 demonstrates that simple tasks in two of the rhetorical modes, ($M=.68$, $SD=.64$ for descriptive essay; $M=61$, $SD=.23$ for narrative essay) had higher scores for subordination compared to complex tasks in those writing modes ($M=.56$, $SD=.21$ for descriptive essay; $M=.54$, $SD=.11$ for narrative essay taking almost similar scores). However, complex task in cause-and-effect essay ($M=.56$, $SD=.20$) had higher score than its simple task ($M=.55$, $SD=.20$), though the difference was very slight. The differences between simple and complex tasks according to the rhetorical mode were not found significant; thus, there was no need to carry out paired samples t test.

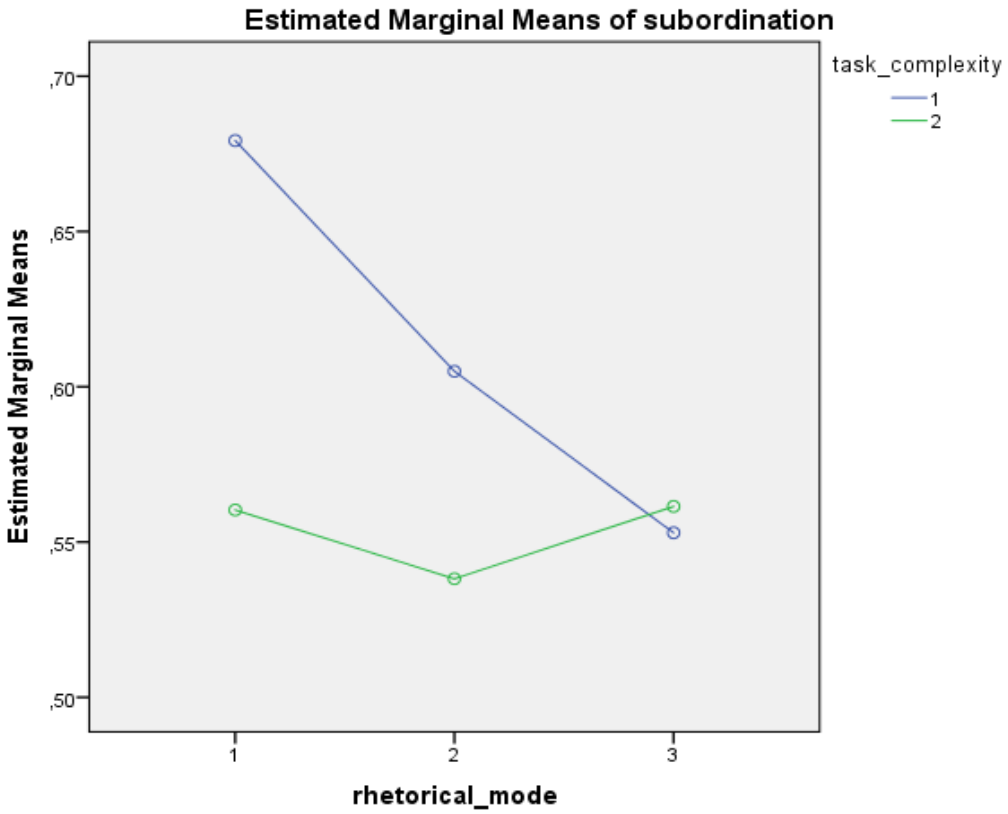


Figure 4.4. Interaction graph for subordination

It is clear in the figure above that there was a clear linear decrease in subordination of simple tasks in the three rhetorical modes. Whereas simple task of the descriptive writing had the highest rate of subordination, that of cause-and-effect had the lowest. However, such difference was not at a significant level, it could not be interpreted that subordination showed difference according to the rhetorical mode of writing.

4.2.1.2.4. Effects of rhetorical mode and task complexity on coordination

Coordination results for simple and complex tasks in the three rhetorical modes of writing were submitted in Figure 4.5. However, there seems to be an inconsistence with the statistic results presented above in terms of the interaction effect of task complexity and rhetorical task. In other words, though there is a seemingly interaction between task complexity and rhetorical task in the figure, the statistical results showed the opposite.

Unlike the results in other components, difference between simple and complex tasks in terms of coordination is not linear. While simple task of descriptive and cause-and-effect essays (M=1.15 , SD=.27; M=1.03, SD=1.57 respectively) had lower score than complex tasks of those rhetorical modes (M=1.1, SD=.11 for descriptive; M=1.09, SD=.09 for cause-and-effect essay) did, the case for narrative was the opposite; that is, coordination results for simple tasks narrative essay (M=1.21, SD=.20) were higher than those of complex task (M=1.15 , SD=.13). Furthermore, it is also clear in the figure that simple tasks in descriptive and narrative essays took similar results for coordination, and complex task of cause-and-effect essays had the highest score. Since the results of coordination did not show significant difference according to the complexity of task, rhetorical modes of writing, and the interaction between these two variables, paired samples t-test was not necessarily conducted.

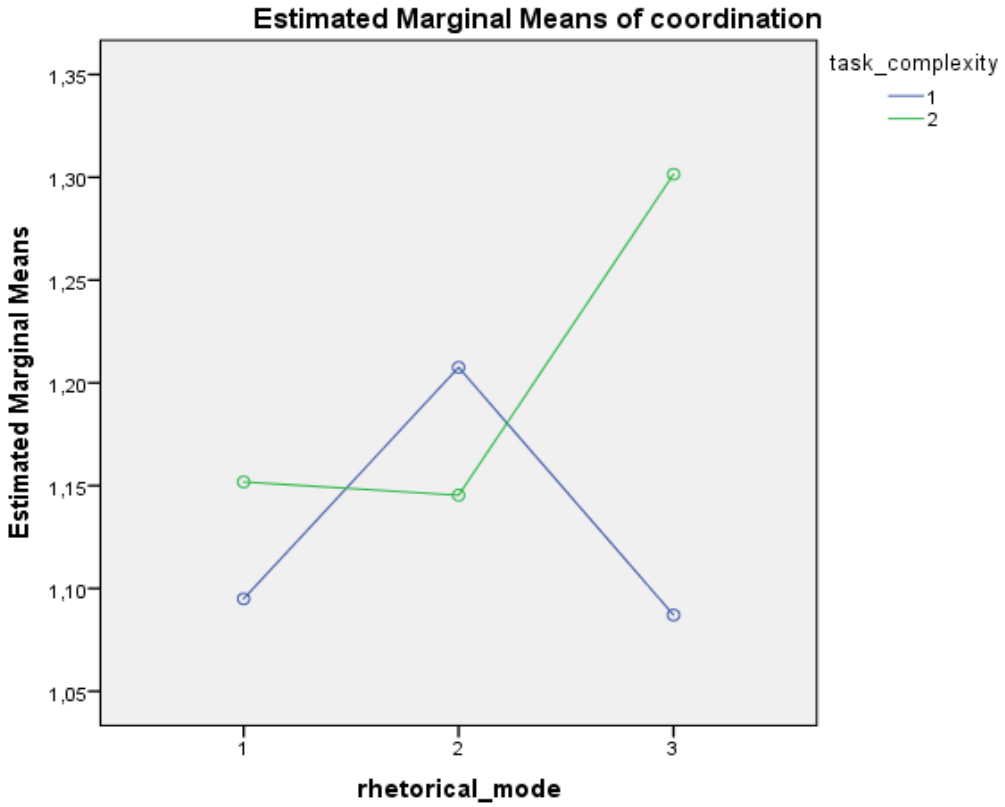


Figure 4.5. Interaction graph for coordination

According to the figure above, the results of coordination in students' writing showed curvilinear increase in rhetorical modes of writing. Although strategic planning

in both descriptive and cause-and-effect writing seemed to result in higher coordination than careful on-line planning, the case for narrative writing was the opposite. In contrast to descriptive results showing differences in coordination, statistical results did show no significant difference between simple and complex tasks, and among the three rhetorical modes of writing.

4.2.3. Descriptive results for components of syntactic complexity

Descriptive statistics are also displayed in the table on the next page for each task in three rhetorical modes of writing.

Table 4.6.

Means and Standard Deviations for Components of Syntactic Complexity

	Descriptive Essay		**Narrative Essay**		**Cause-and-Effect Essay**	
	Task 1 Mean (SD)	Task 2 Mean (SD)	Task 1 Mean (SD)	Task 2 Mean (SD)	Task 1 Mean (SD)	Task 2 Mean (SD)
Length of production	10.92 (6.8)	10.85 (4.1)	9 (2.52)	8.44 (1.82)	10.37 (2.78)	10.22 (1.98)
Sentence complexity	1.75 (1.49)	1.64 (.92)	1.76 (.69)	1.51 (.28)	1.48 (.37)	1.43 (.23)
Subordination	.68 (.64)	.56 (.21)	.61 (.23)	.54 (.11)	.55 (.20)	.56 (.20)
Coordination	1.09 (.11)	1.15 (.27)	1.21 (.20)	1.15 (.13)	1.09 (.09)	1.3 (1.57)

As also statistically stated before, the table above shows that students produced longer, more complex sentences, and higher coordinated writing in descriptive mode than other two types of writing. Similarly, students had better performance in those components while completing simple tasks. However, mixed results were obtained for coordination. Whereas complex task of cause-and-effect had the highest score for coordination, the simple task of the same writing had the lowest. Furthermore, coordination in descriptive essays was lower than that in narrative and cause-and-effect

essays. Except for the length of production, those slight differences in other three components of syntactic complexity were not statistically significant.

Besides the results for components of syntactic complexity, descriptive results for all the indices of these components were also submitted in the table on the next page.

Table 4.7.

Means and Standard Deviations for Indices of Syntactic Complexity

		Descriptive Essay		**Narrative Essay**		**Cause-and-Effect Essay**	
		Task 1 Mean (SD)	Task 2 Mean (SD)	Task 1 Mean (SD)	Task 2 Mean (SD)	Task 1 Mean (SD)	Task 2 Mean (SD)
Length of Production	Mean length of sentence (MLS)	13.16 (10.4)	13.37 (8.41)	11.37 (4.9)	10.02 (2.78)	12.55 (5.08)	12 (3.93)
	Mean length of T-unit (MLT)	12.02 (9.88)	11.1 (3.39)	9.24 (2.39)	8.7 (1.89)	11.56 (4.89)	11.07 (3.88)
	Mean length of clause (MLC)	7.58 (1.32)	8.09 (1.33)	6.44 (0.91)	6.60 (1.11)	8.07 (8.03)	1.27 (0.95)
Sentence Complexity	Clause per sentence (C/S)	1.75 (1.49)	1.63 (0.92)	1.76 (0.69)	1.51 (0.28)	1.55 (0.52)	1.48 (0.37)
Subordination	Clause per T-unit (C/T)	1.61 (1.48)	1.37 (0.31)	1.43 (0.35)	1.31 (0.16)	1.42 (0.49)	1.36 (0.36)
	Dependent clause per clause (DC/C)	0.27 (0.12)	0.24 (0.12)	0.26 (0.11)	0.23 (0.07)	0.25 (0.09)	0.23 (0.09)
	Dependent clause per T-unit (DC/T	0.52 (0.86)	0.36 (0.28)	0.40 (0.31)	0.32 (0.14)	0.39 (0.31)	0.35 (0.25)
	Complex T-unit per T-unit (CT/T)	0.31 (0.2)	0.28 (0.15)	0.32 (0.16)	0.28 (0.09)	0.31 (0.16)	0.27 (0.14)
Coordination	T-unit per sentence (T/S)	1.09 (0.11)	1.15 (0.27)	1.21 (0.20)	1.14 (0.13)	1.09 (0.08)	1.30 (0.09)

As illustrated in Table 4.7, whereas the length of production was obtained through the means of three indices as mean length of sentence, clause, and T-unit, sentence complexity comprised just one index, clause per sentence. Furthermore, subordination in texts was assessed by four indices, clause per T-unit, dependent clause per T-unit, dependent clause per clause, and complex T-unit per T-unit. That is, those indices show the rate of linguistic units having different syntactic status but dependent on others. As for coordination that refers to linguistic units linked to others with equivalent syntactic status, it consisted of three indices, coordinating phrases per clause, coordinating phrases per T-unit, and T-unit per sentence.

It is clear in Table 4.4 that all indices for complex tasks, except for mean length of sentence (M= 13.16, SD=10.4) and T-unit per sentence (M=1.15, SD=.27) in descriptive essays and mean length of clause in cause-and-effect essays (M=6.60, SD=1.11) were higher than simple tasks in the three rhetorical modes of writing. Accordingly, strategic planning conducted before writing task performance in the three rhetorical modes had partially negative effect on the indices of syntactic complexity; however, such effect was not at a significant level. Therefore, it cannot be said that strategic and careful on-line planning resulted in significant difference in any components of syntactic complexity. Furthermore, mean scores for all indices in descriptive essays particularly in complex tasks of that rhetorical task were higher than those of the other two rhetorical tasks.

4.3. Discussion of Results for Syntactic Complexity

In answering the first research question “What are the effects of task complexity and rhetorical mode on syntactic complexity of L2 writers?”, the results were analyzed in terms of two dimensions: syntactic structure and syntactic complexity. It was found that syntactic structure consisting of word count, sentence, verb phrase, clause, T-unit, dependent clause, and complex T-unit was significantly affected by the rhetorical mode in which students produced their writing. When those components were counted, the highest scores for each component were seen in narrative writing of students. On the other hand, compared to descriptive essays, students produced cause-and-effect essays richer in word, sentence, verb phrase, clause, T-unit, dependent clause, and complex T-unit.

With respect to task complexity, when students wrote through strategic planning that was operationalized as simple task in the current study, their performance was found poorer. In contrast, while performing their complex tasks conducted under unpressured on-line planning, students got higher scores for each component of syntactic structure. That difference between simple and complex task was statistically significant. In the light of these results, it can be concluded that careful on-line planning had positive impact on the results of syntactic structure; on the contrary, when the students were required to write while making strategic planning under time pressure, syntactic structure of their writing was poorly affected. In this regard, these findings revealed that besides rhetorical modes of writing, task complexity modified at two levels (strategic planning or unpressured on-line planning) influenced syntactic structure of students' EFL writing performance at a significant level. That is, increasing complexity of task yielded to a decrease in syntactic structures. In other words, special time allotted to make a plan of the content and the way of expressing ideas resulted in an increase in number of syntactic structures.

As for syntactic complexity, rhetorical mode had significant effect just on one component of syntactic complexity, length of production measured through mean length of sentence, mean length of T-unit, and mean length of clause. That is, whether students write a descriptive essay, a narrative or cause-and-effect essay significantly affected mean length of their writing production in terms of sentence, clause, and T-unit. According to the findings, it can be pointed out that students had significantly higher mean scores in their descriptive writing than the two other rhetorical modes, and in cause-and-effect than narrative writing although Beers and Nagy (2009) and Ravid (2004) found out that narratives were produced at greater length. Moreover, contradicting our results that found narrative essays with the least production length, Ravid (2004) attributed his results to the fact that narratives involve mostly personal experience easier to generate and thus have greater length. Although rhetorical modes are different, this inference is also valid for the descriptive writing in this study since the topics used in both tasks of descriptive essays were more personal.

The common rhetorical mode seen in this study and the previous studies (Beers & Nagy, 2009; Ravid, 2004; Yang, 2014) was just narrative writing, but the other rhetorical modes in this study were totally different. Furthermore, the syntactic features

used to measure length of production also showed difference from the measures used in those studies. For instance, Beers and Nagy (2009) applied three measures, words per T-unit, clauses per T-unit, and words per clause but we measured length through the indices such as mean length of sentence, mean length of T-unit, and mean length of clause. In this sense, we needed to take some caution while discussing the effects of rhetorical mode on text length in the light of previous studies. According to Iwashita (2006), the length of writing cannot completely express the syntactic maturity, another name of syntactic complexity, but other structures for range or sophistication like coordination and subordination are also necessary to describe it as syntactically complex.

The other three components of syntactic complexity (sentence complexity, coordination, and subordination) measured in this study did not differ according to the rhetorical mode of writing. However, in contrast to specific results for components, the general results for syntactic complexity displayed that it was significantly affected by the rhetorical mode of writing. That is, though influencing just one component of syntactic complexity – length of production -, writing mode had a significant effect on syntactic complexity as a whole dimension. The results were also supported by Lu (2011) providing a corpus-based evaluation of syntactic complexity measures. He examined the effect of genre in terms of argumentative and narrative writing and revealed that narrative essays were less syntactically complex.

On the other hand, considering the effect of task complexity, we found out that whether students produced their writing under strategic planning or careful on-line planning had no significant effect on syntactic complexity. In contrast to predictions of both hypotheses by Robinson (the Cognition Hypothesis) and Skehan and Foster (the Limited Attentional Capacity), increasing complexity of task did not result in syntactically more complex production. On the contrary, simple tasks of each rhetorical task, particularly regarding length of production, coordination, and sentence complexity, seemed to have higher scores for syntactic complexity than complex tasks; however, the difference was not found significant. In this sense, these results showed similarity with Ellis and Yuan (2004) reporting no significant effect of unpressured on-line planning on syntactic complexity.

Similarly, in another study based on the studies he reviewed to investigate effects of three types of planning, rehearsal, pre-task planning, and on-line (within-task planning), Ellis (2009) had a conclusion that the effects of the three kinds of planning on complexity do not seem to be clear since some other mediating factors such as learner factors or task variables may also affect the results. However, Yuan and Ellis (2003) suggested significant effect of on-line planning on grammatical complexity of students' speech performance. Furthermore, Lu (2011) also revealed that untimed essays were more syntactically complex than timed essays. However, our results did not find any evidence confirming those results for the effect of strategic or on-line planning on syntactic complexity of students' writing.

In contrast to expectations, increasing complexity of task did not result in an increase in syntactic complexity of students' essays. Similar results are also supported by Frear and Bitchener (2015) suggesting that task complexity manipulated according to reasoning demands and number of elements led to a decrease in syntactic complexity. Supporting our study, Adams et al. (2015) also found no significant effect of task complexity on either syntactic or lexical complexity of language production. However, our results show contrast to Salimi et al. (2011) that indicated significant effect of task complexity on syntactic complexity. Furthermore, Ellis and Yuan (2005) also contradict with our study since they revealed a large or medium effect of careful condition (unpressured on-line planning) on syntactic complexity of students' speech and writing. Another striking point suggested by the results is that task complexity did have impact on none dimension of syntactic complexity depending on which rhetorical mode students wrote in although rhetorical mode independently affected syntactic complexity. That is, rhetorical mode interaction by task complexity did not influence the results of syntactic complexity of students' writing.

4.4. Results of Lexical Complexity

4.4.1. Effects of rhetorical mode and task complexity on lexical complexity

Two-way repeated measures MANOVA was conducted to analyze the effect of task complexity and rhetorical modes of writing on lexical complexity that was considered under the three titles – lexical density, lexical sophistication, and lexical

variety. Task complexity, one of the factors, was divided into two levels as simple and complex. Whereas simple task was operationalized as writing through strategic planning, complex task was described as writing through unpressured on-line planning. The other independent variable of this study, rhetorical mode was identified at three levels, descriptive, narrative, and cause-and-effect. According to multivariate test results, although there was no main effect of task complexity, Wilk's Λ=.87, F(3, 38) = 1.91, p = .14, η^2_p=.13, there was a main effect of rhetorical mode, Wilk's Λ=.16, F(6, 156) = 31.26, p = .00, η^2_p=.84, besides there was an interaction effect of rhetorical mode and task complexity, Wilk's Λ=.62, F(6, 35) = 3.51, p = .01, η^2_p=.38 on lexical complexity. The results are presented for each dimension of lexical complexity separately.

Table 4.8.

Results of Two-Way Repeated-Measures MANOVA for Lexical Complexity

Effect	Wilks' Lambda Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
rhetorical mode	.157	31	6	35	.000*	.843
task complexity	.869	.258 1 .910	3	38	.144	.131
rhetorical mode X task complexity	.624	3 .508	6	35	.008*	.376

*Significant effect is reached

Since Mauchly's test indicated that sphericity was met, we looked at the results of sphericity assumption for univariate test. In line with those results, it is clear that although rhetorical task had impact on all of the three components - lexical variety, F(2, 80) = 26.89, p = .00, η^2_p=.40, lexical density, F(2, 80) = 27.27, p = .00, η^2_p=.40, and lexical sophistication, F(2,80) = 77.58, p= .00, η^2_p=.66, task complexity had no significant effect on these components of lexical complexity - lexical variety, F(1,40) = .89, p = .35, η^2_p=.02, lexical sophistication, F(1, 40)= .06, p = .79, η^2_p=.00, and lexical density, F(1, 40) = 3.60, p= .07, η^2_p=.08. On the other hand, whilst lexical variety, F(2, 80) = 7.35, p = .00, η^2_p=.15 was significantly affected by the interaction between rhetorical task and task complexity, lexical density, F(2, 80) = 2.03, p = .14, η^2_p=.05 and

lexical sophistication, F(2, 80) = 1.30, p = .28, η^2_p=.03, were not; that is, task complexity had various effect on the results of lexical variety of students' essays depending on what rhetorical modes of writing they did. In addition, except for main effect of task complexity and interaction effect between rhetorical task and task complexity effect size of which were negligible as they did have no significant impact, main effect of rhetorical task for all dependent variables, and interaction effect between rhetorical task and task complexity on lexical variety have large effect sizes.

Table 4.9.

Univariate Test of Within-Subject Effect

Source	Measure	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
rhetorical mode	lexical density	.094	2	.047	27.265	.000*	.405
	lexical sophistication	.473	2	.237	77.579	.000*	.660
	lexical variety	.191	2	.095	26 885	.000*	.402
task complexity	lexical density	.004	1	.004	3 604	.065	.083
	lexical sophistication	.000	1	.000	.075	.786	.002
	lexical variety	.002	1	.002	.888	.352	.022
rhetorical mode X task complexity	lexical density	.004	1	.004	3.604	.065	.083
	lexical sophistication	.000	1	.000	.075	.786	.002
	lexical variety	.002	1	.002	.888	.352	.022

*Significant effect is reached

4.4.1.1. Effects of rhetorical mode and task complexity on lexical density

The results statistically stated above demonstrated no statistically significant interaction between rhetorical mode and task complexity and no main effect of task complexity, but main effect of rhetorical mode on lexical density. In other words, lexical density scores showed significant difference according to the rhetorical modes of writing but not to the complexity of task, t(40)= -1.90, p=.07. For instance, it was found

that differences between descriptive and cause-and-effect essays, t(40)= -3.68, p=.00 between descriptive and narrative essays, t(40)= -3.93, p=.00, between narrative and cause-and-effect essays, t(40)= -7.02, p=.01, were significant. As seen in Figure 4.6, cause-and-effect essays (M=.53, SD=.04) were more lexically dense than narrative (M=.48, SD=.03) and descriptive essays (M=.50, SD=.04), which had higher scores for lexical density than narrative essays.

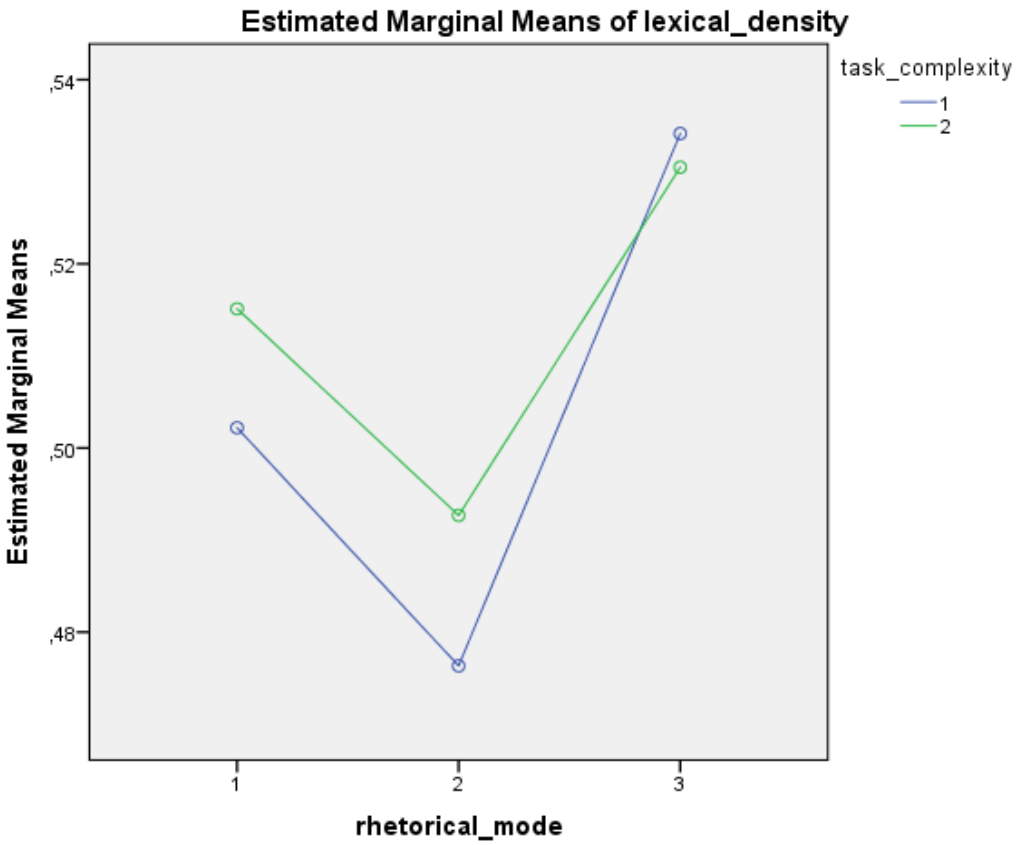


Figure 4.6. Interaction graph for lexical density

Complex tasks in descriptive (M=.52, SD=.03) and narrative writing (M=.49, SD=.04) got slightly but nonsignificantly higher scores than simple tasks in those rhetorical tasks (M=.50, SD=.06 for descriptive; M=.48, SD=.04 for narrative). That is, neither allocating special time to pre-task planning nor increasing the complexity of task along on-line planning without time-pressure had any effect on lexical density of students' writing.

4.4.1.2. Effects of rhetorical mode and task complexity on lexical variety

Unlike lexical density, as clearly seen in interaction graph for lexical variety, there was an interaction between narrative and cause-and-effect essays. The results of three-paired samples t test conducted to investigate the main effect of rhetorical task, found statistically significant by two-way repeated measures MANOVA, revealed that whereas descriptive essays (M=.58, SD=.06) had more lexical variety than narrative (M=.52, SD=.05); t(40)= 6.30, p=.00, and cause-and-effect writing (M=.53, SD=.05); t(40)= 7.23, p=.00, difference between narrative and cause-and-effect essays in lexical variety was nonsignificant, t(40)= -.36, p=.72. That is, lexical variety of students' written productions significantly varied according to the rhetorical mode. Furthermore, difference just between narrative and descriptive writing of students did show difference depending on the task, whether it was simple or complex; t(40)=. -2.11, p=.05.

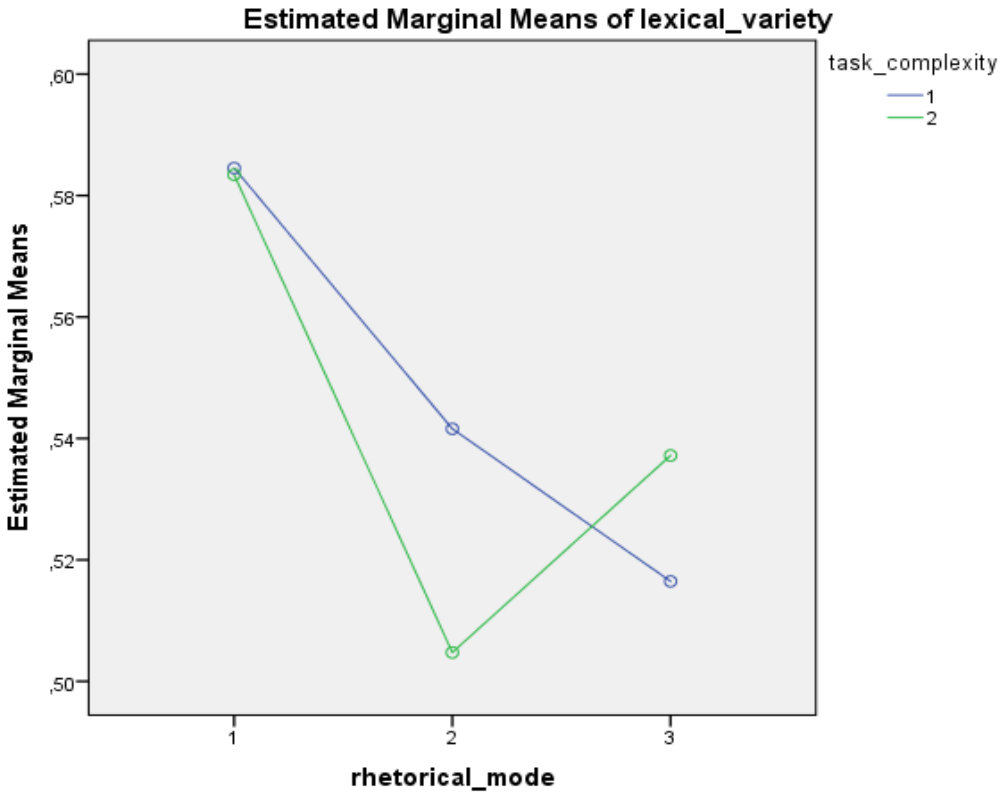


Figure 4.7. Interaction graph for lexical variety

Although both simple and complex tasks in descriptive essay took the same scores for lexical variety (M=.58, SD=.06 for simple and complex tasks), which were also the highest scores, there was difference between simple and complex tasks of narrative and cause-and-effect essays. The biggest difference between complex and simple tasks was seen in narrative essay (M=.54, SD=.07 for simple; M=.50, SD=.06 for complex). Furthermore, while simple task in cause-and-effect essay (M=.52, SD=.06) was less lexically various than complex task of that rhetorical mode (M=.54, SD=.07), simple task of narrative essays (M=.54, SD=.07) had higher scores for lexical variety than complex task (M=.50, SD=.06). However, such difference between complex and simple tasks in terms of lexical variety was not statistically significant, t(40)= .94, p=.35. Accordingly, there was no difference between lexical variety of essays produced under careful on-line planning and those produced through strategic planning. Namely, increasing complexity of the task did not yield any increase in lexical variety.

4.4.1.3. Effects of rhetorical mode and task complexity on lexical sophistication

Figure 4.8 illustrates descriptive results for lexical sophistication of simple and complex tasks in the three rhetorical modes of writing. According to the figure, also stated above with statistical results, there was no rhetorical mode interaction by task complexity in terms of lexical sophistication. Complex tasks in both narrative (M=.24, SD=.06) and descriptive essays (M=.35, SD=.09) got slightly higher scores for lexical sophistication than simple tasks (M=.33, SD=.08 for descriptive; and M=.23, SD=.07 for narrative writing). On the other hand, complex task in cause-and-effect essay (M=.26, SD=.07) was found lexically less sophisticated than simple task in the same rhetorical mode (M=.28, SD=.04). However, there was almost no mean difference between simple and complex task (MD=.001, SD=.05) in terms of lexical variety and it resulted in no significant difference in lexical sophistication, t(40)= -.27, p=79.

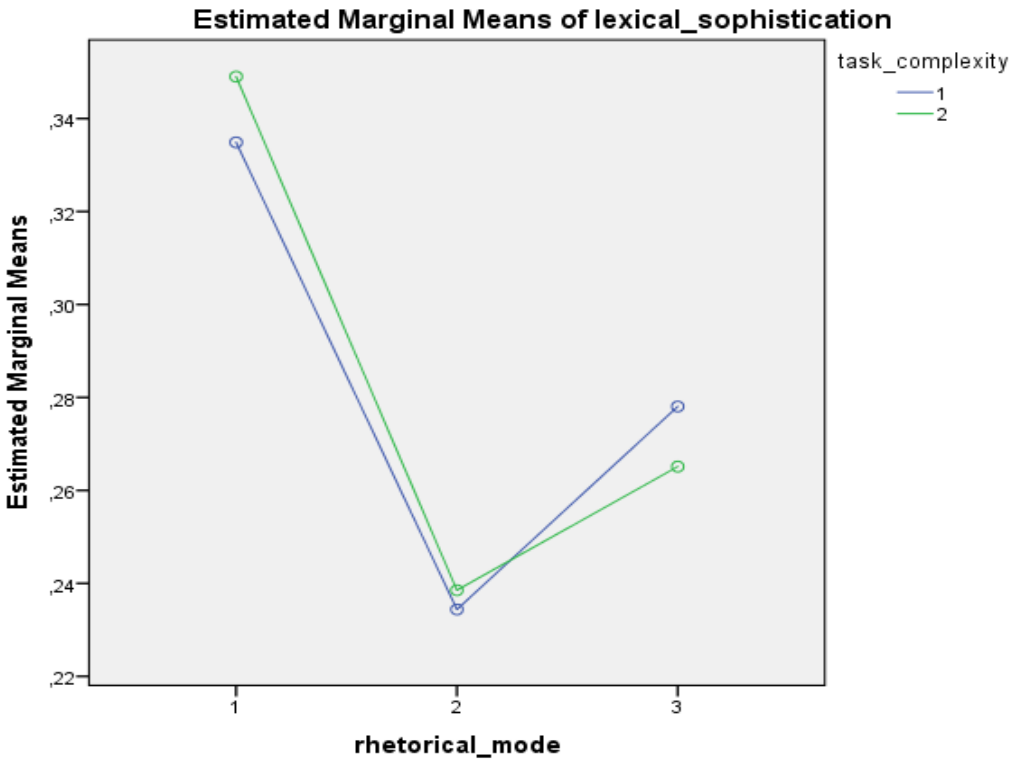


Figure 4.8. Interaction graph for lexical sophistication

As in lexical variety, descriptive writing of students (M=.34, SD=.07) got higher scores for lexical sophistication than narrative (M=.23, SD=.05); t(40)=10.69, p=.00, and cause-and-effect essays (M=.27, SD=.04); t(40)=8.13, p=.00. Moreover, cause-and-effect essays were found more lexically sophisticated than narrative essays, t(40)=4.93, p=.00. That is, descriptive essays having the highest scores of lexical sophistication were followed by cause-and-effect and narrative essays of students respectively.

4.4.2. Descriptive results for the components of lexical complexity

Means and standard deviations were also illustrated in Table 4.5 for simple (Task 1) and complex tasks (Task 2) in three rhetorical tasks – descriptive, narrative, and cause-and-effect – in terms of three dimensions of lexical complexity, lexical sophistication, lexical density, and lexical variety that was obtained through the mean scores of type-token ratio and lexical word variation.

Table 4.10.

Descriptive Results for Independent Measures of Lexical Complexity

		Descriptive Essay		Narrative Essay		Cause-and-Effect Essay	
		Task 1	Task 2	Task 1	Task 2	Task 1	Task 2
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Lexical Density		.50 (.06)	.52 (.03)	.48 (.04)	.49 (.04)	.53 (.05)	.53 (.04)
Lexical Sophistication		.33 (.08)	.35 (.09)	.23 (.07)	.24 (.06)	.28 (.04)	.26 (.07)
Lexical Variety	Type-token ratio	.49 (.06)	.48 (.06)	.44 (.07)	.41 (.06)	.45 (.05)	.47 (.07)
		.58 (.06)	.58 (.06)	.54 (.07)	.50 (.06)	.52 (.06)	.54 (.06)
	Lexical word variation	.67 (.07)	.68 (.07)	.64 (.08)	.60 (.07)	.54 (.07)	.61 (.07)

It is clear in the table that complex tasks had almost the same scores for the components of lexical complexity as the simple tasks. In this regard, increasing task complexity did have no significant effect on lexical variety, sophistication, and density. In other words, lexical variety, sophistication, and density in writings of students did not vary according to the type of planning, whether it was carried out before the task (strategic planning) or during the task having no time-pressure (careful/on-line planning). However, the results of the components of lexical complexity showed significant difference according to the rhetorical mode in which essays were produced. In this sense, whereas descriptive essays had averagely higher results in all components, narrative essays had almost the lowest.

4.5. Discussion of results for lexical complexity

In response to research question 2 "What are the effects of task complexity and rhetorical mode on lexical complexity of L2 writers?", the findings obtained through the analysis of students' written production by Lu's automated tool, Lexical Complexity Analyzer (LCA), were presented. In line with those findings, it was clear that rhetorical mode in which students produced their writing had significant small to medium effect on the results of lexical complexity in terms of the three dimensions, lexical density, lexical variety, and lexical sophistication.

For all three components of lexical complexity, narrative essays were given the lowest scores, but descriptive essays had the highest scores for lexical variety and lexical sophistication and cause-and-effect essays for lexical density. In other words, although descriptive essays which were found more lexically various and sophisticated followed by cause-and-effect and narrative writing of students in that order, higher results of lexical density were given to cause-and-effect essays compared to descriptive essays that were found more lexically dense than narrative essays. Yang (2014) similarly found that lexical density was affected by the rhetorical mode of writing; furthermore, in congruence with the results of this study, narrative essays were found to have the least lexical density among the four rhetorical tasks in that study. That is, as far as the proportion of lexical words to the total words is considered, it is the least in narrative essays of this study.

According to these results, narrative writing performances of students had the lowest results for the three dimensions of lexical complexity. In this respect, it can be concluded that students had a poorer performance in terms of lexical complexity while completing their narrative task. Whereas their descriptive performance was found the most successful in terms of lexical variety and sophistication, their cause-and-effect writing was the richest in lexical density. In other words, there were more unusual or advanced words in descriptive essays but higher proportion of lexical words in cause-and-effect essays. That is, narrative essays of students involved more familiar and less advanced words, which is in line with the results of Graesser, McNamara, and Kulikowich (2011) suggesting that narratives tend to be on familiar topics and thus have more commonly used words.

However, the results for lexical variety were affected not only by the rhetorical mode of writing but also its interaction by task complexity; that is, the effect of rhetorical mode showed difference depending on whether it was produced through strategic planning or unpressured on-line planning. For instance, whereas simple task was more lexically dense and sophisticated than complex task in cause-and-effect writing, the case for descriptive and narrative writing was the opposite. However, lexical variety was greater in complex task of cause-and-effect writing than it was in the simple task, but it had the opposite results in narrative writing.

As for the effect of task complexity, the findings of this study were again contradictory to the prediction of two competing models. Though suggested by previous research that increasing complexity of a task would yield more complex language production (Levkina & Gilabert, 2012), the current study had no supporting results; on the contrary, it was pointed out that there was no significant difference in lexical complexity according to whether learners performed their writing task through strategic planning or careful on-line planning. These results were also supported by Lourdes Ortega (1999) who also found no effect of pre-task planning on lexical complexity. On the other hand, at odds with this study, Abrams and Byrd (2016) evinced the positive effect of pre-task planning on lexical richness of learners' texts.

Contradicting the studies that found important effect of task complexity on syntactic and lexical complexity, this study did find no effect of task complexity

operationalized at two levels. Although in the light of literature it was initially believed that writing tasks conducted through strategic planning which was identified as simple task in this study would result in more syntactically and lexically complex production, the current study had no advocating results. However, these results were confirmed by the studies pointing out that increasing complexity of a task had no effect on linguistic (syntactic and lexical) complexity of written production.

4.6. Results of Cohesion

4.6.1. Effects of rhetorical mode and task complexity on cohesion

Two-way repeated measures MANOVA was conducted to see the effects of rhetorical mode and task complexity on cohesion since there were five dimensions to be measured and two factors believed to have effect on these dimensions. In terms of general effects, multivariate test indicated that besides main effects of task complexity, Wilk's Λ=.79, F (2, 39) = 5.16, p = .01, η^2_p=.21, and rhetorical task, Wilk's Λ=.31, F (4, 37) = 20.94, p = .00, η^2_p=.69, there was also an interaction effect between these variables, Wilk's Λ=.71, F (4, 37) = 3.71, p = .01, η^2_p=.29, on cohesion of students' essays. Furthermore, it is clear that their effect sizes are large.

Table 4.11.

Results of Two-Way Repeated-Measures MANOVA for Cohesion

Effect	Wilks' Lambda Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
rhetorical mode	.306	20.938	4 .000	37	.000*	.694
task complexity	.791	5.162	2 .000	39	.010*	.209
rhetorical mode X task complexity	.714	3.705	4 .000	37	.012*	.286

* Significant effect is reached.

Cohesion was analyzed by Coh-Metrix 3.0 under two dimensions, referential and deep cohesion. Since Mauchly's test revealed that just the results of main effect of rhetorical task and interaction effect on referential cohesion were met, the results of sphericity assumption were presented. However, the results of Greenhouse-Geisser for the main effect of task complexity on referential and deep cohesion, main effect of

rhetorical task and the effect of its interaction by task complexity were naturally presented due to violation seen in Mauchly's test results. In this sense, univariate test results were submitted for each component respectively.

Accordingly, rhetorical modes of writing significantly affected students' essays in terms of referential cohesion, F(1.95, 77.9)= 16.52, p= .00, η^2_p=.29, and deep cohesion, F(1.6, 63.91)= 17.03, p= .00, η^2_p=.30. Whereas deep cohesion, F(1,40)= 9.05, p= .01, η^2_p=.18, was influenced by task complexity, referential cohesion, F(1,40)= .87, p= .36, η^2_p=.02, was not. On the other hand, rhetorical interaction by task complexity had significant effect on referential cohesion, F(2, 80)= 8.2, p= .00, η^2_p=.17, but no impact on deep cohesion, F(1.72, 68.93)= .12, p= .86, η^2_p=.00.

Table 4.12.

Univariate Test of Within-Subject Effect

Source	Measure	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
rhetorical mode	Referential cohesion	11848.057	1.947	6084.098	16.520	.000*	.292
	Deep cohesion	17451.293	1.598	10922.227	17.030	.000*	.299
task complexity	Referential cohesion	380.634	1.000	380.634	.865	.358*	.021
	Deep cohesion	3637.870	1.000	3637.870	9.046	.005*	.184
rhetorical mode x task complexity	Referential cohesion	4549.195	1.915	2375.217	8.200	.001*	.170
	Deep cohesion	90.886	1.723	52.741	.121	.857*	.003

*Significant effect is reached

Besides descriptive results for both components respectively, Figure 4.9 also displays whether there is any interaction between rhetorical modes of writing and task complexity.

4.6.1.1. Effects of rhetorical mode and task complexity on referential cohesion

Descriptive results for the first dimension of cohesion that indicates the proportion of overlap in words and ideas between sentences, were displayed in Figure 4.9. It is clear in the figure that although simple tasks in descriptive (M=42.22, SD=21.95) and cause-and-effect essays (M=56.2, SD=24.56) had more referential cohesion than complex tasks (M=34.54, SD=22.07; M=46.78, SD=23.12 respectively), when it comes to the simple task of narrative essay, it is the exact opposite case. That is, the rate of referential cohesion in complex task of narrative (M=59.12, SD=18) was higher than that in its simple task. (M=49.49, SD=23.64). Nonetheless, besides results of two-way repeated measures MANOCA, the results of two paired-samples t-test revealed simple tasks (M=49.30, SD=17.16), though having higher referential cohesion, showed no significant difference from complex tasks (M=46.81, SD=16.67), t(40)= .93, p=.36.

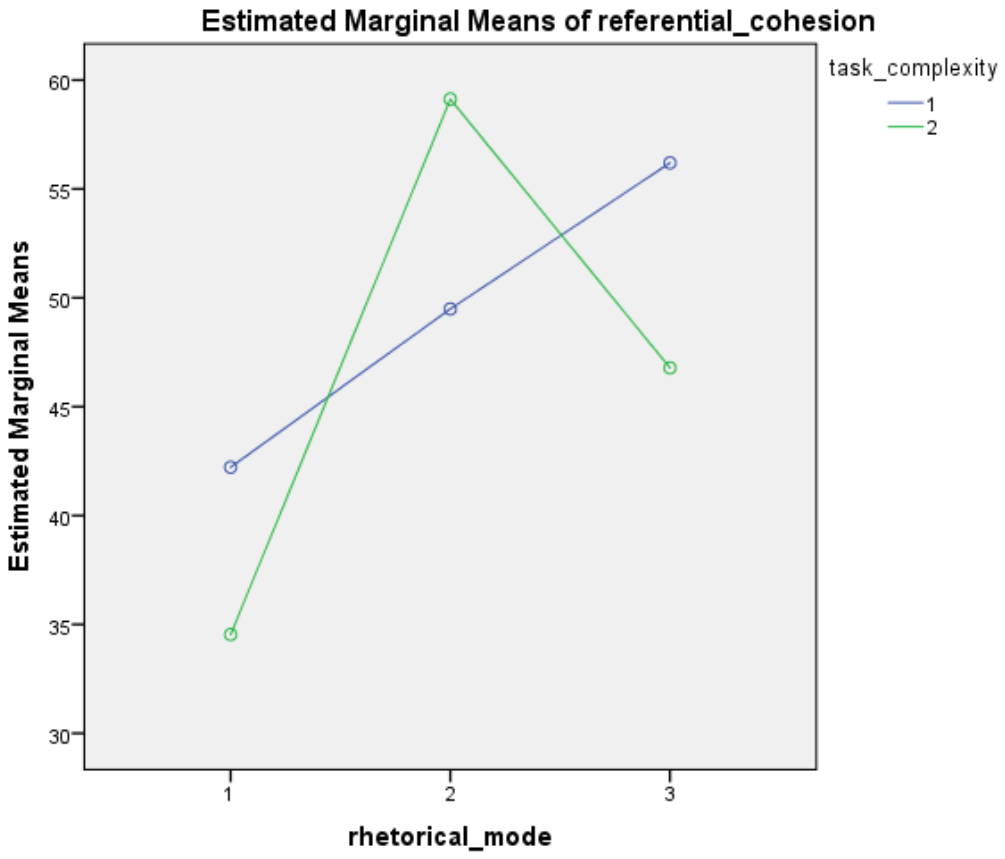


Figure 4.9. Interaction graph for referential cohesion

Descriptive essays (M=38.37, SD=17.97) had significantly lower referential cohesion than both narrative (M=54.30, SD=16.49); t(40)= -5.09, p<.05, and cause-and-effect essays (M=51.48, SD=20.05); t(40)=-4.83, p<.05; however, difference between narrative and cause-and-effect was not significant, t(40)= .94, p>.05. In this respect, the rhetorical mode in which the writing task was performed significantly affected the rate of referential cohesion in the text. According to these results, whereas narrative essays of students were the richest in referential cohesion, descriptive essays were the poorest.

Paired-samples t-tests was carried out to investigate the interaction effect, which was also clear in Figure 4.9, revealed that differences between complex and simple tasks also had significant differences between narrative and cause-and-effect writing, t(40)=-3.83, p=.00, and between descriptive and narrative writing, t(40)= 3.03, p=.00, but had no significant difference between descriptive and cause-and-effect essays, t(40)-.36, p=.72.

4.6.1.2. Effects of rhetorical mode and task complexity on deep cohesion

Figure 4.10 shows a linear increase in the deep cohesion of the three rhetorical tasks. The rate of deep cohesion reflecting causal and intentional cohesive devices showed increase from the first rhetorical mode to the third. That is, cause-and-effect essays (M=88.74, SD=11.74) were more deeply cohesive than both narrative (M=72.10, SD=14.49); t(40)= -6,61, p=.00 and descriptive (M=69. 87, SD=21.97); t(40)=-4.63, p=.00, essays, However, the difference between descriptive and narrative essays which seems very low, in terms of deep cohesion, was nonsignificant, t(40)=-59, p=.56. Since there was no significant interaction between rhetorical modes of writing and complexity of the tasks, paired-samples t test for the interaction effect was not necessarily conducted.

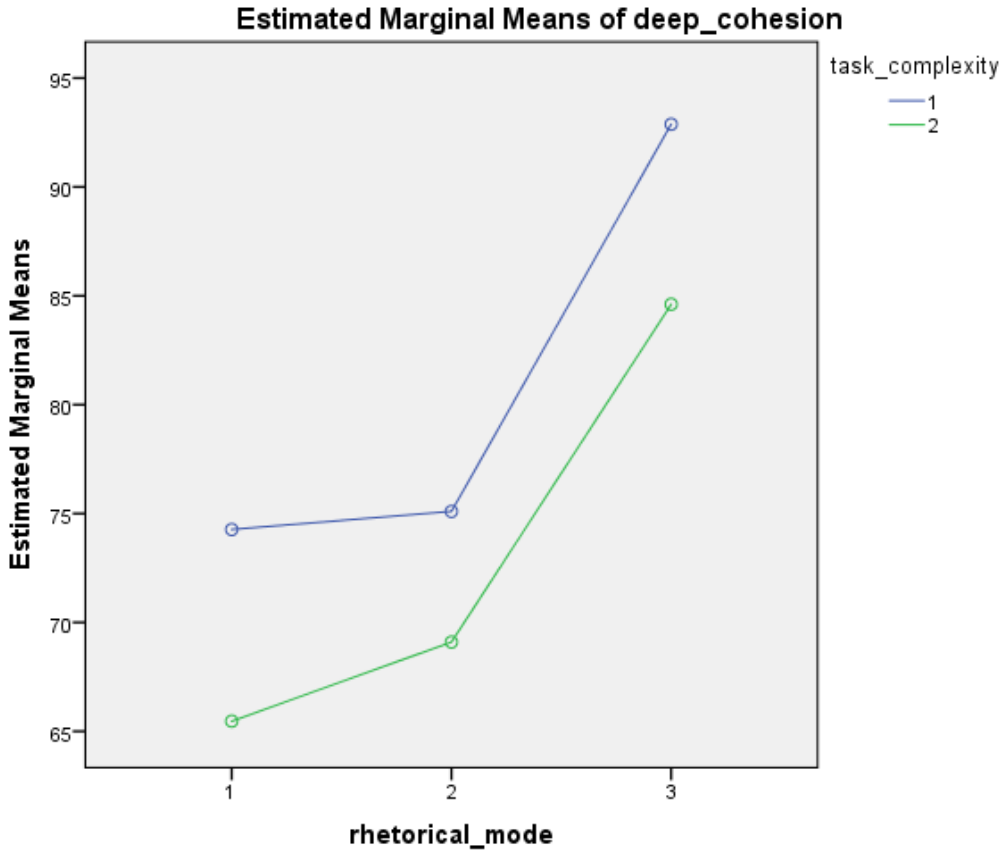


Figure 4.10. Interaction figure for deep cohesion

As clearly seen in the figure above, simple tasks (M=74.27, SD=25.28 for descriptive; M=, SD=75.10 for narrative; and M=92.88, SD=14.15 for cause-and-effect essays) had higher deep cohesion than complex tasks (M=65.46, SD=26.5; M=69.1, SD=21.70; M=84.61, SD= 18.29 respectively). Furthermore, that difference between complex (M=73.06, SD=12.85) and simple tasks (M=80.75, SD=13.45) was found statistically significant, t(40)= 3.01, p=.01. Accordingly, besides the rhetorical mode of writing, task complexity manipulated along planning (strategic or careful on-line planning) had effect on deep cohesion of students' writing.

4.6.2. Descriptive results for components of cohesion

Descriptive results – means and standard deviations - for components of cohesion were displayed in Table 4.13.

Table 4.13.

Means and Standard Deviations for Components of Cohesion

	Descriptive		**Narrative**		**Cause-and-effect**	
	Task 1	**Task 2**	**Task 1**	**Task 2**	**Task 1**	**Task 2**
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**	**Mean (SD)**
Referential Cohesion	42.22 (21.95)	34.54 (22.07)	49.49 (23.64)	59.12 (18)	56.20 (24.56)	46.78 (23.12)
Deep Cohesion	74.27 (25.28)	65.46 (26.5)	75.10 (21.61)	69.10 (21.70)	92.88 (14.15)	84.61 (18.29)

As illustrated in the table above, deep cohesion in all rhetorical modes had higher proportion compared to referential cohesion. That is, students used more causal and intentional cohesive devises explicitly rather than linking the ideas through referential devices in order to build connection between ideas. Furthermore, whereas cause-and-effect writings of students were the richest in explicit use of cohesive devices, their descriptive writings were the poorest. As for the use of referential cohesive devices, while narrative essays had the highest proportion, descriptive essays had the lowest. Accordingly, student descriptive essays were found the least cohesive.

Moreover, according to Table 4.13, simple tasks were more successful in terms of deep cohesion. Apart from the case in narrative writing, the same results were also seen in referential cohesion. In other words, although complex tasks of narrative writing had higher results for referential cohesion, simple tasks of the other two rhetorical modes had higher results than complex tasks.

4.7. Discussion of results for cohesion

Written task performances of students were analyzed by Coh-Metrix 3.0 to provide evidence for the third research question "What are the effects of task complexity and rhetorical mode on cohesion of EFL learners' writing production?". That program evaluated the essays under five measures, narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion. However, just two of these dimensions were involved in this study as it was mainly concerned with cohesion but the other three referred to the easibility of the text. Each component had varying results for the effect of task complexity and rhetorical mode of writing on cohesion of EFL writers' task performance.

First dimension of cohesion measured in this study, referential cohesion was found to be affected by the rhetorical mode of writing at a significant level. In this sense, the results revealed that descriptive essays were weaker in referential cohesion that indicates to what extent ideas are connected in a text (McNamara et al., 2014) than both narrative and cause-and-effect writing of students. Although difference between narrative and cause-and-effect essays in referential cohesion was not significant, narrative essays seemed to have higher referential cohesion than cause-and-effect essays. Low proportion of referential cohesion is an indicator of cohesion gap in a text that challenges the reader to make inferences particularly when they have limited prior knowledge. Accordingly, descriptive essays were the most difficult to infer and build connections between ideas and sentences.

In contrast to rhetorical mode, task complexity did have no significant effect on referential cohesion of a text. Although referential cohesion in simple tasks showed a linear increase according to the rhetorical mode, complex tasks had higher referential cohesion than simple tasks; however, such difference was not found significant. It can be reported that both simple tasks conducted under strategic planning and complex tasks performed by making on-line planning without time pressure had no significant difference in referential cohesion that indicates overlap in words and ideas between sentences.

As for the second dimension of cohesion in this study, the findings displayed that both rhetorical mode of writing and complexity of task had significant effect on the

results of deep cohesion in a text. Deep cohesion reflects the rate of causal and intentional cohesive devices used to build coherence and thus to show links between ideas or events. Because of such a support, a text is easier to comprehend even if the topic is unfamiliar or the reader has no sufficient background. In this respect, compared to the two rhetorical modes in this study, higher proportion of deep cohesion, cause-and-effect writing of students had, made them easier to understand. On the other hand, it was also pointed out that the most difficult texts to comprehend the relationship between ideas or events were produced in descriptive mode since they were found to lack connecting words.

Furthermore, the results also revealed that there was significant difference in deep cohesion between the texts produced through strategic planning and those written through unpressured on-line planning. It was found that simple tasks (writing through strategic planning) had more linking words to clarify the relationship between ideas or events and thus had higher proportion of deep cohesion than complex tasks (writing through careful/unpressured on-line planning). That is, it was suggested that increasing complexity of a task yields texts with deep cohesion. However, since most of the studies on task performance focused on three measures of CAF, no support or contradiction was seen in the literature regarding the effect of task complexity on cohesion.

4.8. Results of Coherence

4.8.1. Effects of rhetorical mode and task complexity on coherence

Students' essays were evaluated by two trained raters through an analytical rubric that contains a specific dimension to measure coherence. The results obtained from rubrics were analyzed by two-way repeated measures ANOVA to see whether rhetorical task and complexity of task had any effect on coherence of students' writing. According to multivariate test results, neither task complexity, Wilk's Λ=.93, F (1, 40) = 3.16, p = .08, η^2_p=.06, nor rhetorical task, Wilk's Λ=.87, F (2, 39) = 3.05, p = .06, η^2_p=.14, had significant effect on the results of coherence. Similarly, no significant effect of rhetorical interaction by task complexity, Wilk's Λ=.96, F (2, 39) = 78, p = .46, η^2_p=.04 was seen.

Table 4.14.

Results of Two-Way Repeated-Measures ANOVA for Coherence

Effect	Wilks' Lambda Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
rhetorical mode	.865	3.046	2	39	.059	.135
task complexity	.927	3.158	1	40	.083	.073
rhetorical mode X task complexity	.961	.784	2	39	.464	.039

*Significant effect is reached

Results for coherence in the two tasks of each rhetorical mode were displayed in the figure below.

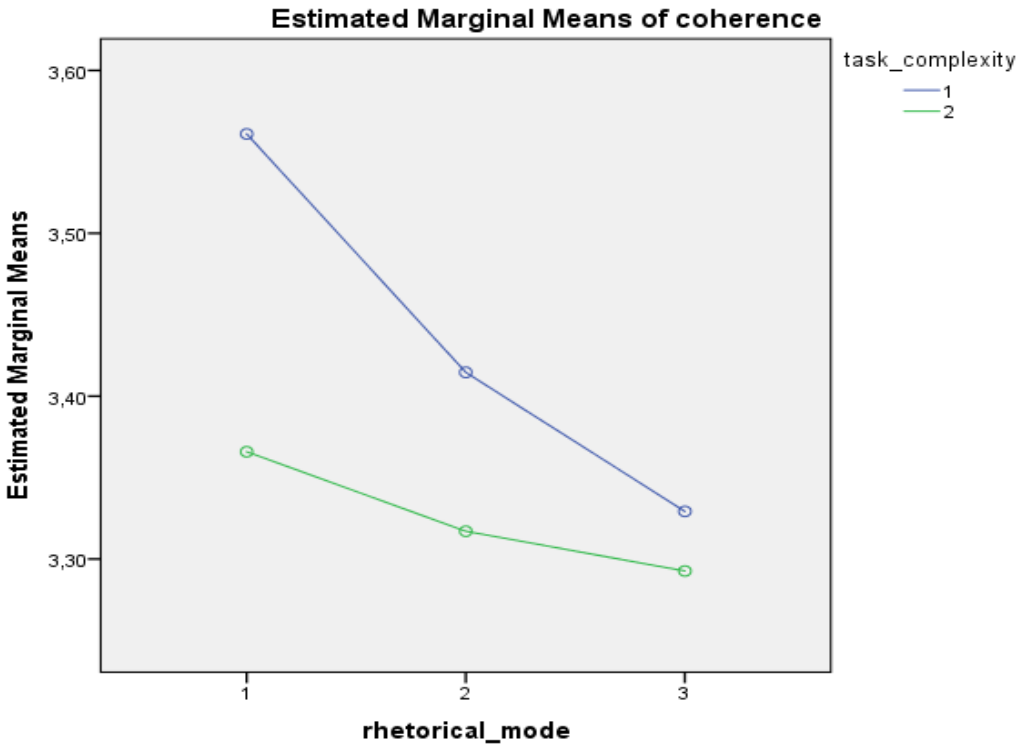


Figure 4.11. Interaction graph for coherence

It is clear in the figure above, complex tasks of the three rhetorical modes (M=3.43, SD=.49) had higher scores for coherence than simple tasks (M=3.32, SD=.38); however, as stated statistically before, such difference was not significant. It can be seen in the figure that whereas complex task of descriptive essays (M=3.56,

SD=.61) had the highest level of coherence, simple task of cause-and-effect essays (M=3.29, SD=.52) had the lowest. Furthermore, there was almost no difference in coherence between complex task and simple task of students' narrative writing. Another striking point observed in the results, the scores of coherence showed a linear decrease according to the rhetorical mode as follows: descriptive essays (M=3.46, SD=.44) were found more coherent than narrative essays (M=3.36, SD=.52) and cause-and-effect essays (M=3.31, SD=.46) respectively. Even so, these results showing difference in coherence among rhetorical tasks were not found significant.

Unlike other dimensions such as linguistic complexity and cohesion in this study, we did have no necessarily to conduct three paired samples t-test to see differences since neither main effects of task complexity and rhetorical task nor interaction effect between these two factors were obtained.

4.8.2. Descriptive results for coherence

The descriptive results – means and standard deviations – for coherence were illustrated in Table 4.15.

Table 4.15.

Means and Standard Deviations for Coherence

	Descriptive Essay		**Narrative Essay**		**Cause-and-Effect Essay**	
	Task 1 Mean (SD)	Task 2 Mean (SD)	Task 1 Mean. (SD)	Task 2 Mean (SD)	Task 1 Mean (SD)	Task 2 Mean (SD)
Coherence	3.37 (.47)	3.56 (.61)	3.32 (.61)	3.41 (.68)	3.29 (.52)	3.32 (.51)

Considered the scores for coherence ranged between 1 and 5 (1=the lowest, 5=the highest) in the direction of instructions of the rubric, the score of 3 can be described as midlevel of coherence. Table 4.6 illustrates that all scores were around 3. Accordingly, it can be concluded that essays, on average, had midlevel of coherence. Although all essays took similar scores, the biggest difference seemed to be between the complex task of descriptive essays (M=3.56, SD=.61) and simple tasks of cause-and-

effect essays (M=3.29, SD=.52); however, as stated before, no difference in students' writing in terms of coherence was significant.

4.9. Discussion of Results for Coherence

In response to research question 4 asking the effect of rhetorical task and task complexity on coherence of students' writing, the results obtained through the evaluation of students' essays were presented. Regarded as a substantial feature of effective writing, coherence shows the relationship between ideas to create meaning for the reader (Lee, 2002). Subjective by its nature, coherence does not exist in text itself but reflects what the reader grabs from the text; therefore, it is more challenging to measure it objectively (Todd et al., 2004; Yule, 2010).

In the light of literature, we measured coherence by an analytical rubric involving a special section for coherence. It states that if an essay "uses a logical structure that is appropriate to paper's subject, purpose, audience, thesis, and disciplinary field, employes sophisticated transitional sentences which often develop one idea from the previous one or identify their logical relations, and clearly guides the reader through the chain of reasoning or progression of ideas", it should be scored as 5. However, if it has no appreciable organization; lacks transitions and coherence, it should be scored as 1. According to descriptions of rubric for coherence, it can be inferred that the students' essays in this study having average score of 3 had an organization without any evident logical structure. That is, although students used some transitions, presented a topic idea in each paragraph, and related paragraphs to the main idea, arrangement in sentence probably lacked coherence and logic.

On the other hand, the results revealed that the factors in this study, rhetorical mode and task complexity, did have no significant effect on coherence. In other words, in which rhetorical mode students produce their writing, descriptive, narrative, and cause-and-effect, did show no difference in coherence. Similarly, results of coherence were not affected by whether they made strategic or unpressured on-line planning while writing. In contrast to descriptive results showing that complex tasks conducted through on-line planning under no time pressure were more coherent than simple tasks carried

out under strategic planning, ANOVA test results revealed that such difference was nonsignificant.

Similarly, rhetorical tasks seemed to have different scores for coherence according to descriptive results; however, differences within rhetorical modes of writing were not found significant. Although descriptive essays appeared to be the most coherent, the case for cause-and-effect essays was the opposite. These results showed contrast to the results of referential cohesion demonstrating the extent of connections that link the ideas in this study. Descriptive results had the lowest scores for referential cohesion but the highest scores for coherence; on the contrary, cause-and-effect essays had the highest results for referential cohesion but the lowest for coherence. Confirming Crossley and McNamara (2012), the results illustrated that a text high in referential cohesion does not mean to have also coherence.

Furthermore, the results of deep cohesion measured in the current study were not supported by the results of coherence. Whereas cause-and-effect writing of students had also the highest scores for deep cohesion that illustrates the proportion of causal and intentional cohesive devices in a text, narrative writing had the lowest. However, cause-and-effect having the highest scores for both referential and deep cohesion was found the least coherent. From these results, it can be inferred that the rates of deep and referential cohesion do not reflect the level of coherence. In other words, a text involving cohesion may be described as incoherent. It probably results from the fact that cohesion exists in the text itself, but coherence is what the reader grabs from the text (Danielle S. McNamara et al., 2014).

As for the effect of rhetorical tasks and task complexity on coherence, we did find no study providing support or contradiction. Unlike the current study, most of task-based research investigated task complexity in terms of complexity, accuracy, and fluency particularly to test their trade-off effects. However, investigating task complexity in the perspective of writing, besides linguistic complexity, we also evaluated students' essays in terms of coherence and cohesion which are the significant components to be involved in an effective writing.

4.10. Results of Overall Writing Quality

4.10.1. Effects of rhetorical mode and task complexity on overall writing quality

After evaluating students' essays for specific features of writing, we evaluated them in terms of general writing achievement through the same analytical rubric used in assessment of coherence. As a statistical test, two-way repeated measures ANOVA was carried out to see whether rhetorical task and complexity of task had any effect on overall quality of students' writing. According to multivariate test results, in addition to main effect of task complexity, Wilk's Λ=.82, F (1, 40) = 9.10, p = .00, η^2_p=.19, rhetorical task, Wilk's Λ=.82, F (2, 39) = 4.39, p = .02, η^2_p=.18, also had significant effect on general writing achievement of students. Furthermore, rhetorical task interaction by task complexity also significantly affected the results of writing quality, Wilk's Λ=.28, F (2, 39) = 5.40, p = .01, η^2_p=.21.

Table 4.16.

Results of Two-Way Repeated-Measures ANOVA for Overall Writing Quality

Effect	Wilks' Lambda Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
rhetorical mode	.816	4.387	2	39	.019*	.184
task complexity	.815	9.099	1	40	.004*	.185
rhetorical mode X task complexity	.783	5.401	2	39	.009*	.217

*Significant effect is reached

Finding significant main effects of rhetorical mode and task complexity, we also carried out paired-samples t-test to assess the differences in results, controlling familywise error rate using Holm's sequential Bonferonni approach (Green & Salkind, 2004). The figure below clearly illustrates the results for overall writing quality of each task.

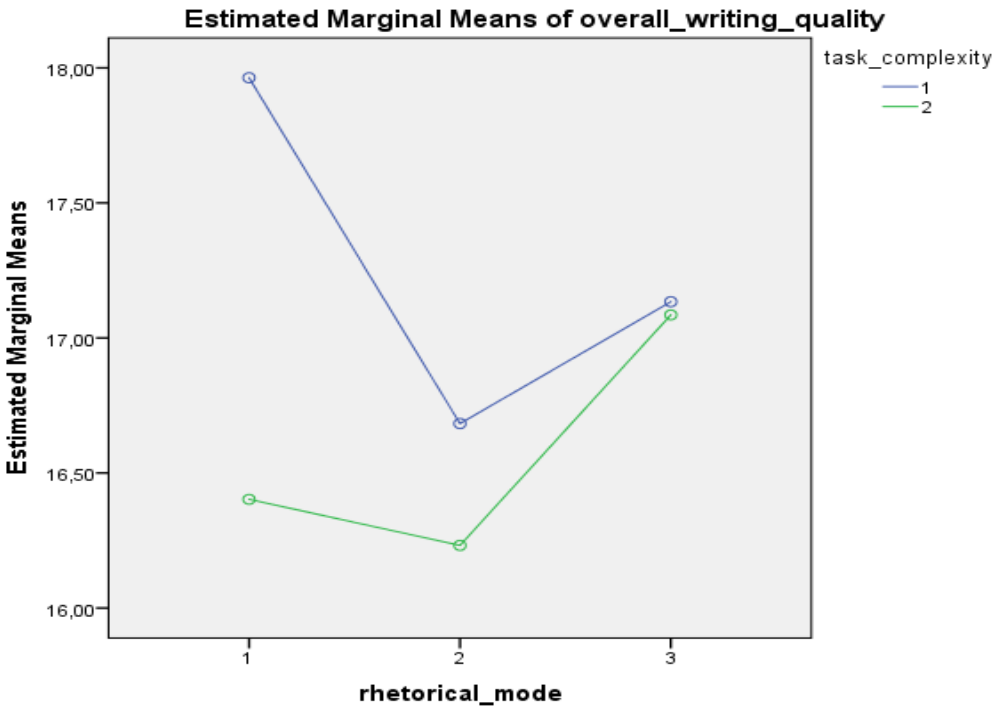


Figure 4.12. Interaction graph for overall writing quality

Figure 4.12 displays the results of overall writing quality for both tasks, simple and complex, in the three rhetorical modes of writing. It is clear that, though having almost the same scores as the simple task in cause-and-effect writing did, complex tasks of the three rhetorical modes (M=17.26, SD=1.94) were found to have higher writing quality than simple tasks (M=16.57, SD=1.47). Although there was less than 1 point-difference in mean scores, difference was found significant; t(40)=3.02, p=.00. In this regard, it can be suggested that when students produced their writing through strategic planning (simple task), their writing was found less qualified than that they produced under unpressured on-line planning (complex task). That is, increasing complexity of task resulted in higher qualified essays.

As seen in the figure above, overall writing quality for descriptive essays (M=17.18, SD=1.81) were found better than narrative (M=16.45, SD=1.87) and cause-and-effect essays (M=17.10, SD=1.82). Although the slight difference between descriptive and cause-and-effect essays was not statistically significant; t(40)=.29, p=.77, difference between descriptive and narrative writing, though low, was found significant; t(40)= 2.90, p=.01. Furthermore, it was also found that students' cause-and-

effect writing (M=17.10, SD=1.82) had significantly higher writing quality than their narrative writing (M=16.45, SD=1.87); t(40)= 2.28, p=.03. As a consequence, the rhetorical task, descriptive, narrative, or cause-and-effect writing, in which students produced their writing had significant effect on the results of their general writing achievement. Furthermore, whereas the complex tasks in descriptive essays (M=17.96, SD=2.60) were found the most qualified, their complex tasks in narrative writing were the least (M=16.23, SD=1.95), which probably results from interaction effect between task complexity and rhetorical mode of writing.

In order to see the interaction of rhetorical task by task complexity, three-paired samples t-test was conducted. The results revealed that differences in writing quality between simple and complex tasks showed also significant difference between descriptive and narrative essays, t(40)=-2.46, p=.02, and between descriptive and cause-and-effect essays, t(40)= -3.15, p=.00; however, such difference between complex and simple tasks was not seen between narrative and cause-and-effect essays, t(40)=-.86, p=.40. In this sense, it is pointed out that the effect of tasks complexity showed difference according to the rhetorical mode in which they were performed. Furthermore, nonsignificant difference between narrative and cause-and-effect essays also resulted in no significant difference between complex and simple tasks.

4.10.2. Descriptive results for overall writing quality

The descriptive results – means and standard deviations – for overall writing quality of students were illustrated in Table 4.17.

Table 4.17.

Means and Standard Deviations for Overall Writing Quality

	Descriptive Essay		**Narrative Essay**		**Cause-and-Effect Essay**	
	Task 1 Mean (SD)	Task 2 Mean (SD)	Task 1 Mean. (SD)	Task 2 Mean (SD)	Task 1 Mean (SD)	Task 2 Mean (SD)
Overall writing quality	16.40 (1.59)	17.96 (2.60)	16.68 (2.52)	16.23 (1.95)	17.08 (1.99)	17.13 (2.10)

As seen in Table 4.17, the scores for writing quality ranged between about 16 and 18. When it is considered that the rubric used in this study gives 25 as the highest score and 5 as the lowest score, these results can be interpreted as being at medium level. Furthermore, although differences in scores between simple and complex tasks and among rhetorical tasks do not seem quite high, they were found statistically significant.

4.11. Discussion of results for overall writing quality

In order to provide response to the last research question of this study asking the effect of rhetorical mode and task complexity on general writing achievement of EFL learners, the students' writing productions were evaluated by an analytical rubric. According to the results obtained through the statistical tests, rhetorical mode had significantly affected students' general writing achievement. In support to Engelhard et al. (1992) Prater (1985); Prater and Padia (1983), the results of this study revealed that writing quality of students showed difference according to the rhetorical mode in which they wrote. In contrast to those studies, the current study described narrative essays as having the lowest writing quality but descriptive essays as the most qualified. Similar results were also presented by Way et al. (2000) who illustrated that students had best writing performance on descriptive mode compared to narrative and expository writing. However, unlike the current study, narrative essays in Engelhard Jr et al. (1992) were found the highest qualified essays but their explanation why narratives were the most qualified also provided evidence for our results. They suggested that the more personal responses a writing task requires, the higher proficiency it receives. That is, since topics in descriptive tasks were more personal compared to those in the other two rhetorical modes, learners had natural higher scores for writing quality.

These results were also supported by the results of length of production showing that descriptive essays were produced at the greatest length. The same results were also seen in dimensions of lexical variety and lexical sophistication. In line with these results, lexical variety, lexical sophistication, and length of production seemed to have predictive power of writing quality. Crossley and McNamara (2010, 2012) similarly described the essays with higher lexical sophistication as high qualified. Furthermore,

according to Crossley and McNamara (2012) lexical variety can be used to significantly predict writing proficiency of L2 learners, which provided support to this study.

In addition, results pertaining to coherence demonstrated that descriptive essays were the most coherent, which also supported the results of general writing achievement. However, the case for narrative and cause-and-effect essays was varying. Although narrative writing of students had the least writing quality but higher coherence than their cause-and-effect writing, the results for coherence and writing quality in cause-and-effect writing were exactly opposite. That is, coherence might be a good indicator of writing quality (McCulley, 1985) in descriptive essays; however, making such an inference for the other two rhetorical modes may not be possible.

As for cohesion results, they do not show any similarity to the results of writing quality. In other ways, whereas the lowest rate of deep and referential cohesion was in descriptive essays but the highest in cause-and-effect essays, descriptive essays were found to have the highest quality but cause-and-effect essays did not have the lowest. It can be inferred that higher cohesion did not result in higher writing achievement. In congruence with this study, Crossley and McNamara (2012) revealed that writing described as higher proficient is not meant to be higher cohesive but higher lexically sophisticated. Moreover, they (2010) suggested that higher coherence found as a significant indicator of writing quality in a text does not necessarily show great proportion of cohesive devices; nevertheless, cohesion was not necessarily a predictor of writing quality results. These results provided support to the current study.

Similar findings for coherence and cohesion were also suggested by Crossley and McNamara (2011), McNamara et al. (2009), and Crossley and McNamara (2012). In contrast to the common sense that high rate of cohesion, and by proxy, coherence in a text shows high writing quality, they pointed out that either higher coherence or cohesion does not result in higher writing proficiency.

Besides rhetorical mode of writing, the other independent variable of this study, task complexity had also significant effect on the quality of students' writing. In this regard, students' essays produced under unpressured on-line planning (described as complex) were more qualified than their essays produced under strategic planning

(described as simple). That is, increasing complexity of task along planning resulted in higher quality of students' writing.

CHAPTER FIVE

5. CONCLUSION

5.1. Introduction

This chapter firstly presents a summary of the study and its main findings. Following to that, implications for EFL writing teachers, instruction, and assessment processes are presented. With the discussion of limitations of the current study and suggestions for further studies, the chapter goes end.

5.2. Summary of the Study

Key to this study was the attempt to investigate whether rhetorical mode of writing and task complexity had effect on the results of EFL students' writing in terms of linguistic complexity, coherence, cohesion, and overall writing quality. First of all, students' writing tasks that are carried out through strategic planning and unpressured on-line planning can be accurately described as a task since they meet all of the four criteria (Ellis, 2009) a task should satisfy on the basis of its definition (Ellis, 2003). In this regard, as students mainly focus on processing information in terms of semantic and pragmatic meaning while doing writing, the first criterion is well satisfied. As for the second criterion stating that "there should be some kind of 'gap'", it is already met since the main purpose of writing is to convey information or express an opinion about a specific topic. While writing an essay, the students initially rely on their own linguistic resource, which is a clear satisfaction of the third criterion. Lastly, a task should end with an outcome rather than just use of language; in this regard, writing production obtained at the end of writing process is a response to that criterion.

Motivated by two competing models in literature on task-based research, Skehan's Trade-off Hypothesis and Robinson's Cognition Hypothesis, the study operationalized writing tasks in each rhetorical mode of writing at two levels as simple and complex. Manipulation of tasks was based on Yuan and Ellis (2003)'s definition of

on-line planning (or within-task planning) that refers to careful and deliberate effort in 'planned language use'. In line with that definition, whereas writing task carried out through strategic planning was identified as simple, the other task conducted through unpressured on-line planning was complex. Like many studies (Way et al., 2000), three different writing modes as descriptive, narrative, and cause-and-effect were used. However, the last mode was completely different from those in literature though narrative writing, the second mode, was commonly used in many studies of task-based research.

Following a repeated-measures design, the study was carried out with 41 junior ELT students at a state university in Turkey. Each student wrote 6 essays in response to the instructions by their writing teacher, the researcher herself. A total of 241 essays were obtained and evaluated in terms of dimensions in this study. Whereas automated tools were applied to analyze syntactic complexity (Lu's L2 Syntactic Complexity Analyzer), lexical complexity (Lu's Lexical Complexity Analyzer), and cohesion (Coh-Metrix 3.0), other two dimensions, coherence and writing quality, were evaluated through an analytic rubric consisting of a specific section for coherence by two trained raters.

Syntactic complexity was presented in two main categories as syntactic structure and syntactic complexity that was also divided into four subcategories as length of production, sentence complexity, subordination, and coordination. The study revealed that rhetorical mode of writing did have no significant effect on the results of sentence complexity, subordination, and coordination but had small effect on the length of production. In this regard, descriptive essays were found significantly greater at length than cause-and-effect and narrative essays respectively. However, the results of the components of syntactic complexity showed no significant difference according to complexity of the task. Namely, there was no significant difference in results of syntactic complexity between students' writing productions conducted through strategic planning and those performed under careful on-line planning.

As to lexical complexity that consists of three components as lexical density, lexical sophistication, and lexical variety in this study, all the components were also affected by the rhetorical mode in which the students produced their writing. In this

sense, their worst performance in using various, advanced, unfamiliar, and lexical words was seen in narrative writing. However, like the components of syntactic complexity, none of the components of lexical complexity was affected by task complexity. That is, increasing complexity of the task did not yield any rise in lexical complexity of EFL writers' task performance.

Another dimension of this study, cohesion involved two components as referential and deep cohesion. The rhetorical mode of writing had small to medium significant effect on the results of these components. Whereas the highest results for deep cohesion were found in their cause-and-effect writing, students' narrative writing had the highest scores for referential cohesion. As for the effect of task complexity on cohesion, the results revealed that writing through unpressured on-line planning had significant negative effect on referential cohesion. That is, students' writing productions under unpressured on-line planning had a significant lower rate of referential cohesion compared to their writing produced following to strategic planning. However, the effect of task complexity on deep cohesion was not significant. In this regard, although simple tasks seemed to have higher deep cohesion than complex tasks, such difference was found nonsignificant. Accordingly, one dimension of cohesion, referential cohesion, was significantly affected by the complexity of task, but the other, deep cohesion, was not.

On the other hand, neither rhetorical mode nor task complexity had any effect on the results of coherence in students' writing. Since coherence is, by its nature, subjective (Yule, 2010) and illustrates what the reader grabs from the text (McNamara et al., 2014), it was analyzed by an analytical rubric in this study. Although descriptive essays were found to have the highest coherence followed by narrative essays and cause-and-effect essays respectively, the difference was not significant. Similar results were also seen in terms of task complexity. Namely, whereas students' writing productions in complex tasks performed under unpressured on-line planning were found less coherent than those in simple tasks, such difference was not at a significant level. The descriptive results also demonstrated that students' essays were at medium level of coherence; that is, they were not well-organized and did not have a perfect logical structure; nevertheless, their writings were not completely incoherent but lacked some transitions showing logical division of ideas.

After being analyzed in terms of basic dimensions, students' essays were evaluated for overall writing quality by the same rubric used in the analysis of coherence. The results displayed that both rhetorical mode of writing and task complexity had significant but small effect on students' general writing achievement. In this regard, whereas descriptive writing of students was found to have the highest quality, their narrative writing had the lowest. Similarly, their complex tasks performed without having time pressure had higher writing quality than simple tasks. Difference between simple and complex tasks is slight in mean scores, but it was the greatest in descriptive writing.

In brief, the first factor, rhetorical mode of writing was found to have significant effect on almost all dimensions of this study except for coherence and three components of syntactic complexity. However, it significantly affected one component of syntactic complexity, the length of production, but had no effect on the other three components, sentence complexity, coordination, and subordination. However, as for the effect of task complexity, it influenced just one component of cohesion, deep cohesion, and writing quality but had no impact on other dimensions in the current study.

5.3. Implications

In line with the main findings of this study, many important implications for writing instruction, writing assessment, and task-based research were obtained.

5.3.1. Implications for writing instruction

Firstly, adopted in the current study, task-based instruction provides great opportunities for natural learning by putting emphasis on meaning over form (Arslanyilmaz, 2013; Ellis, 2009; Willis, 1996). Writing can be defined as a pedagogical task which is the primary unit of task-based language teaching since it meets the four criteria of Skehan (1998) and Ellis (2009). As clearly seen in the current study, task-based approach can be effectively adopted while teaching writing. However, writing instructors should pay utmost attention to selecting or designing the task. Besides fundamental elements proposed for the sequencing of task such as goals, input, procedures, learner and teacher roles, other factors such as students' level of knowledge

and proficiency in writing were also found to be of importance in designing a writing task.

The results of this study also provided clear implications for writing instruction. It was seen that while teaching writing, its dimensions cannot be ignored; that is, linguistic complexity, cohesion, and coherence should have place in syllabus and material design for writing courses (Lu, 2011). In other words, writing courses should be designed so as to explicitly teach (Lee, 2002) and seperately evaluate these dimensions besides general writing proficiency. However, some caution should be taken while relating the results of these dimensions to writing quality. As seen in this study, an essay that is found higher qualified does not mean that it is cohesive, coherent, or linguistically complex; or vice versa; that is, a text with higher coherence does not yield higher quality (Crossley & McNamara, 2012). Although an effectively written text is supposed to have both syntactic and lexical complexity as much as coherence and cohesion, it should be never disregarded that a text having all these dimensions is not meant to have naturally higher writing quality. That is, they may not have predictive power of writing quality.

On-line planning modified at two levels as pressured and unpressured was found to have almost no significant effect on the results of writing dimensions in the current study. However, as suggested by Skehan and Foster (1997), planning performed as a way of pre-task activity, is an important aspect of task performance since a well-prepared task is more likely to result in a successful performance.

5.3.2. Implications for writing assessment

One of the four basic skills of language, writing is a complex process in that it indicates a persons' ability to use language besides the ability to express ideas (Liu & Braine, 2005). Not only the process itself but its evaluation also requires much more time and skills. In this sense, automated programs are very helpful in evaluation of written production from different perspectives, particularly in evaluation of the components involved in the text itself such as lingusitic complexity and cohesion. On the contrary, manual analysis is more time-consuming and subjective. The programs used for analysis of lexical complexity (Lu's Lexical Compelxity Analzyer), syntactic

complexity (Lu's L2 Syntactic Complexity Analyzer), and cohesion (Coh-Metrix) were easy and free to use. Furthermore, they enabled to assess the written performance of students from various perspectives at a time.

Instead of holistic evaluation of writing which provides narrower ranges for writing achievement (Ortega, 2003), an analytic rubric consisting of five dimensions was used in this study. Written productions of students were simultaneously evaluated in terms of organization, style, ideas, mechanics, and support and thus from a broader perspective. Therefore, the use of analytic rubrics is corroborated in this study. It is also suggested that rubrics used to evaluate general writing achievement should have specific descriptors for syntactic and lexical complexity.

5.3.3. Implications for task-based research

As suggested by Ellis (2009), task-based language teaching builds a bridge between theory of second language acquisition and the actual process of language teaching and learning. In this regard, it has attracted many researchers admiring to investigate natural language learning process. Nevertheless, a major problem exists for task-based researchers while selecting and sequencing the task. Furthermore, operationalization of task complexity is another important issue for researcher since one task may be regarded more complex by some learners although others consider it as simple. For instance, in the present study the first tasks of each rhetorical mode that were carried out under strategic planning seemed to be more complex due to time pressure learners would have while writing; however, those tasks were conceptualized as simple tasks as they required less cognitive demands. In this sense, researchers should be careful about difference between two terms "task difficulty" and "task complexity" which can be interchangeably used.

Moreover, the automated analysis tools used in the current study were proved to facilitate the process of evaluation of written texts in many aspects at a time. In this sense, the researchers who want to study written task performance are advised to use such programs to get quick and versatile results. Another striking implication to be suggested with respect to this study is dependent variables which show difference from the task-based studies in the literature. Although previous studies focus on the measures

of CAF (Beers & Nagy, 2009; Frear & Bitchener, 2015; Kuiken et al., 2005; Kuiken & Vedder, 2007b; Kuiken & Vedder, 2008; Lu & Ai, 2015; Ong & Zhang, 2010; Yang et al., 2015), the current study also investigated task performance in terms of other dimensions such as coherence, cohesion, and writing quality as well as linguistic complexity. It was demonstrated that additional to CAF measures, new dimensions of written performance can be investigated to see the effect of task complexity on them.

5.4. Limitations and suggestions for further studies

The current study has some limitations advising to take some caution in interpreting its main findings. These limitations also suggest directions for future research. The first limitation is the participants of this study. They were Turkish EFL university students which may pose an obstacle in generalizing the results of the current study to other EFL/ESL contexts or populations. The number of participants (41) that may be a strength for a study following repeated-measures design is also seen as a limitation. Future studies may benefit from a larger number of participants to have more generalizable results.

In addition, the findings of this study may be adversely affected by the fact that the participants of this study were nonnative speakers of English. However, native speakers generally produce more complex and longer utterances compared to non-native speakers (Ai & Lu, 2013; Tavakoli, 2009b). At least, in order to minimize such an effect, their proficiency in English can be taken into consideration to have more similar results to those of native speakers. Furthermore, as reported by Wigglesworth (1997) and Kawauchi (2005), the effects of planning may also differ according to proficiency level. In this sense, L2 proficiency ignored in the current study may be also investigated by further studies to see its impact on the results of students' writing (Ishikawa, 2006; Iwashita, 2006; Lu & Ai, 2015).

Another limitation of this study is caused by the design itself we followed. Although repeated-measures design gives the researcher strength to avoid unsystematic variation, it may also result in some problems such as practice effect and boredom effects. However, we may overcome these effects, which seems impossible to be completely eliminated, by counterbalancing the order we used performing the tasks

(Field, 2012). That is, we may first perform the simple task and then complex task for the second or third rhetorical task to avoid those effects. Furthermore, although location may not be a threat for repeated-measures designs (Fraenkel et al., 2012), it may seem to pose a threat in our study since complex tasks were required to complete out of the classroom and the place where the participants performed the writing task would have bilateral effects on the outcomes.

This study also neglected the effect of time and instruction on results. However, as demonstrated by Crossley and McNamara (2014) that time spent in a writing class results in a significant increase in syntactic complexity, further research may consider time effect on writing outcomes to obtain significant results. Similarly, as also suggested by Ortega (2003), Mazgutova and Kormos (2015) a three or four-month advanced level instruction may yield higher syntactic complexity, lexical complexity, coherence, and cohesion. Additionally, in order to see the effect of instruction on the results, future research can follow a true experimental design instead of repeated-measures design and thus compare instructed versus uninstructed learners through a control and experimental group to which participants are assigned randomly.

The current study considered just the effect of planning and rhetorical mode; however, further research might be necessary to see how different types of task influence the results of EFL writing. In this regard, new types of task manipulated along cognitive complexity may be performed by future research. Similarly, besides or except for the rhetorical modes employed in the current studies, new modes of writing may be investigated. Furthermore, the dimensions of writing in the current study may be analyzed with extra tools to have more reliable results.

Probably the most salient limitation of the current study is the measures used. Although almost all task-based research depends on three measures of CAF by Skehan, the two measures, accuracy and fluency, were neglected to be assessed. However, particularly the first factor of this study, strategic or on-line planning, is found to have great effect on these measures (Skehan & Foster, 2005). Furthermore, compared to strategic planning, on-line planning is regarded to be more relevant to accuracy. In this regard, the further research is suggested to evaluate also accuracy and fluency of students' writing as well as linguistic complexity.

REFERENCE LIST

Abrams, Z. I., & Byrd, D. R. (2016). The effects of pre-task planning on L2 writing: Mind-mapping and chronological sequencing in a 1st-year German class. *System, 63*, 1-12.

Adams, R., Nik Mohd Alwi, N. A., & Newton, J. (2015). Task complexity effects on the complexity and accuracy of writing via text chat. *Journal of Second Language Writing, 29*, 64-81. doi:10.1016/j.jslw.2015.06.002

Adrian, M. M., & Mangado, M. J. G. (2015). L1 Use, Lexical Richness, Accuracy and Syntactic Complexity in the Oral Production of CLIL and NON-CLIL Learners of English. *Atlantis-Journal of the Spanish Association of Anglo-American Studies, 37*(2), 175-197.

Ahangari, S., & Abdi, M. (2011). The effect of pre-task planning on the accuracy and complexity of Iranian EFL learners' oral performance. *Procedia - Social and Behavioral Sciences, 29*, 1950-1959. doi:10.1016/j.sbspro.2011.11.445

Ahmadian, M. J., Abdolrezapour, P., & Ketabi, S. (2012). Task difficulty and self-repair behaviour in second language oral production. *International Journal of Applied Linguistics, 22*(3), 310-330.

Ai, H., & Lu, X. (2013). A corpus-based comparison of syntactic complexity in NNS and NS university students' writing. *Studies in Corpus Linguistics, 59*, 249-264.

Arslanyilmaz, A. (2012). An online task-based language learning environment: Is it better for advanced - or intermediate- level second language learners? *TOJET: The Turkish Online Journal of Educational Technology, 11*(1), 20-35.

Arslanyilmaz, A. (2013). Computer-assisted foreign language instruction: task based vs. form focused. *Journal of Computer Assisted Learning, 29*(4), 303-318. doi:10.1111/jcal.12003

Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing, 22*(2), 185-200. doi:10.1007/s11145-007-9107-5

Blair, T. K., & Crump, W. D. (1984). Effects of discourse mode on the syntactic complexity of learning disabled students' written expression. *Learning Disability Quarterly, 7*(1), 19-29.

Bonate, P. L. (2000). *Analysis of pretest-posttest designs*. Boca Raton: Chapmann & Hall / CRC.

Breen, M. (1987). Learner contribution to task design. In C. Cadlin and D. Murphy (eds) *Language Learning Tasks* (pp.23-46). Englewood Cliffs NJ: Prentice-Hall

Brindley, G., & Slatyer, H. (2002). Exploring task difficulty in ESL listening assessment. *Langauge Testing, 19*(4), 369-394. doi:10.1191/0265532202lt236oa

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21-45). Amsterdam: John Benjamins Publishing Company.

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing, 26*, 42-65. doi:10.1016/j.jslw.2014.09.005

Bussmann, H. (2006). *Routledge dictionary of language and linguistics*. New York: Routledge.

Candlin, C. (1987). Towards task-based language learning. In C. Candlin & D. Murphy (Eds.), *Language learning tasks*. London: Prentice Hall.

Chen, M., & Zechner, K. (2011, 19-24 June). *Computing and evaluating syntactic complexity features for automated scoring of spontaneous non-native speech.* Paper presented at the 49th Annual Meeting of the Association for Computational Linguistics, Portland, Oregon.

Cohen, L., Manion, L., & Morrison, K. (2005). *Research methods in education* (5th ed.). London and New York: Routledge.

Covington, M. A., He, C., Brown, C., Naci, L., & Brown, J. (2006). *How complex is that sentence? A proposed revision of the Rosenberg and Abbeduto D-Level*

*Scale. (*A CASPR research report by the University of Georgia)**.** Retrieved from http://athenaeum.libs.uga.edu/xmlui/bitstream/handle/10724/19857/2006-01-Covington.pdf?sequence=1

Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (2nd ed.). Upper Seaddle River: Merrill.

Crossley, S. A., Kyle, K., & McNamara, D. S. (2016). The development and use of cohesive devices in L2 writing and their relations to judgments of essay quality. *Journal of Second Language Writing, 32*, 1-16.

Crossley, S. A., & McNamara, D. S. (2009). Computational assessment of lexical differences in L1 and L2 writing. *Journal of Second Language Writing, 18*(2), 119-135.

Crossley, S. A., & McNamara, D. S. (2010). *Cohesion, coherence, and expert evaluations of writing proficiency.* Paper presented at the Proceedings of the 32nd annual conference of the Cognitive Science Society, Portland, Oregon, USA.

Crossley, S. A., & McNamara, D. S. (2011). *Text coherence and judgments of essay quality: Models of quality and coherence.* Paper presented at the Proceedings of the 29th Annual Conference of the Cognitive Science Society, Nashville, Tenessee.

Crossley, S. A., & McNamara, D. S. (2012). Predicting second language writing proficiency: the roles of cohesion and linguistic sophistication. *Journal of Research in Reading, 35*(2), 115-135.

Crossley, S. A., & McNamara, D. S. (2014). Does writing development equal writing quality? A computational investigation of syntactic complexity in L2 learners. *Journal of Second Language Writing, 26*, 66-79. doi:10.1016/j.jslw.2014.09.006

Crowhurst, M. (1990). Reading/writing relationships: An intervention study. *Canadian Journal of Education/Revue canadienne de l'éducation, 15*(2), 155-172.

Crowhurst, M., & Piche, G. L. (1979). Audience and mode of discourse effects on syntactic complexity in writing at two grade levels. *Research in the Teaching of English, 13*(2), 101-109.

Crystal, D. (1992). *An encylopedic dictionary of language & languages*. Oxford: Blackwell.

Crystal, D. (2003). *A dictionary of linguistics and phonetics* (5th ed.). Massachusetts: Blackwell.

Crystal, D. (2010). *The Cambridge encyclopedia of language* (3rd ed.). Cambridge: Cambridge University Press.

Ellis, R. (2000). Task-based research and language pedagogy. *Language Teaching Research, 4*(3), 193-220.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Ellis, R. (2005). Planning and task performance: Theory and research. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 3-34). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics, 30*(4), 474-509. doi:10.1093/applin/amp042

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics, 19*(3), 221-246.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition, 26*, 59-84. doi:10.1017/S0272263104261034

Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167-192). Amsterdam: John Benjamins.

Engelhard Jr, G., Gordon, B., & Gabrielson, S. (1992). The influences of mode of discourse, experiential demand, and gender on the quality of student writing. *Research in the Teaching of English, 26*(3), 315-336.

Example of a Grading Rubric For a Term Paper in Any Discipline. Retrieved on 07.06.2016, from https://course1.winona.edu/shatfield/air/termpaper.pdf

Field, A. (2012). *Discovering statistics using IBM SPSS statistics* (4th ed.). London: SAGE Publications.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education*. New York: Mc Graw Hill.

Frear, M. W., & Bitchener, J. (2015). The effects of cognitive task complexity on writing complexity. *Journal of Second Language Writing, 30*, 45-57. doi:10.1016/j.jslw.2015.08.009

Fulcher, G., & Reiter, R. M. (2003). Task difficulty in speaking tests. *Language Testing, 20*(3), 321-344.

Gan, Z. (2012). Complexity measures, task type, and analytic evaluations of speaking proficiency in a school-based assessment context. *Language Assessment Quarterly, 9*(2), 133-151. doi:10.1080/15434303.2010.516041

Genç, Z. S. (2012). Effects of strategic planning on the accuracy of oral and written tasks in the performance of Turkish EFL learners. In A. Shehadeh & C. A. Coombe (Eds.), *Task-Based Language Teaching in Foreign Language Contexts* (pp. 67-88). Amsterdam: John Benjamins Publishing.

Graesser, A. C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix: Providing Multilevel Analyses of Text Characteristics. *Educational Researcher, 40*(5), 223-234. doi:10.3102/0013189x11413260

Green, S. B., & Salkind, N. J. (2004). *Using SPSS for Windows and Macintosh - Analyzing and Understanding Data* (4th ed.). New Jersey: Pearson Education.

Halliday, M. A. K., & Hassan, R. (1976). *Cohesion in English*. Hong Kong: Longman.

Hamp-Lyons, L. (1993). Second language writing: Assessment issues. In B. Kroll (Ed.), *Second Langauge Writing* (pp. 69-87). New York: Cambridg University Press.

Hamp-Lyons, L. (2003). Writing teachers as assessors of writing. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 162-89). New York: Cambridge University Press.

Hedge, T. (1988). *Writing*. Oxford: Oxford University Press.

Hinkel, E. (2004). *Teaching academic ESL writing: Practical techniques in vocabulary and grammar*. London: Lawrence Erlbaum Associates.

Ho, R. (2006). *Handbook of univariate and multivariate data analysis and interpretation with SPSS*. Boca Raton: Chapmann & Hall / CRC.

Hunt, K. W. (1965). Grammatical Structures Written at Three Grade Levels. NCTE Research Report No. 3.

Hyland, K. (2002). *Teaching and researching writing*. Malaysia: Longman Pearson Education.

Ishikawa, T. (2006). The effect of task complexity and language proficiency on task-based language performance. *The Journal of Asia TEFL, 3*(4), 193-225.

Iwashita, N. (2006). Syntactic complexity measures and their relation to oral proficiency in Japanese as a foreign language. *Language Assessment Quarterly, 3*(2), 151-169. doi:10.1207/s15434311laq0302_4

Johansson, V. (2009). Lexical diversity and lexical density in speech and writing: a developmental perspective. *Working Papers in Linguistics, 53*, 61-79.

Johnson, B., & Christensen, L. (2012). *Education research: Quantitative, qualitative, and mixed approaches*. USA: SAGE Publications.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130-144.

Kawauchi, C. (2005). The effects of strategic planning on the oral narratives of learners with low and high intermediate L2 proficiency. In R. Ellis (Ed.), *Planning and task performance in as second langauge* (pp. 143-164). Amsterdam: John Benjamins.

Kırkgöz, Y. (2014). Task-based language teaching. In S. Çelik (Ed.), *Approaches and principles in English as a foreign language (EFL) education* (pp. 225-40). Ankara: Eğiten Kitap.

Knoch, U. (2007). 'Little coherence, considerable strain for reader': A comparison between two rating scales for the assessment of coherence. *Assessing Writing, 12*(2), 108-128. doi:10.1016/j.asw.2007.07.002

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing, 20*(2), 148-161.

Kuiken, F., Mos, M., & Vedder, I. (2005). Cognitive task complexity and second language writing performance. In S. Foster-Cohen, M. d. P. Garcia Mayo, & J. Cenoz (Eds.), *Eurosla Yearbook* (Vol. 5, pp. 195-222). Amsterdam: John Benjamins.

Kuiken, F., & Vedder, I. (2007a). Task complexity and measures of linguistic performance in L2 writing. *Iral-International Review of Applied Linguistics in Language Teaching, 45*(3), 261-284.

Kuiken, F., & Vedder, I. (2007b). Task complexity and measures of linguistic performance in L2 writing. *IRAL - International Review of Applied Linguistics in Language Teaching, 45*(3). doi:10.1515/iral.2007.012

Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing, 17*(1), 48-60.

Lantolf, J. P. (2011). The sociocultural approach to second langauge acquisition: Socicultural theory, second langauge acquision, and artificial L2 development. In D. Atkinson (Ed.), *Alternative approaches to second langauge acquisition* (pp. 24-47). New York: Routledge.

Larsen-Freeman, D., & Anderson, M. (2011). *Techniques & principles in language teaching* (Third ed.). New York: Oxford University Press.

Lawal, B. (2014). *Applied statistical methods in agriculture, health and life sciences.* Malete, Nigeria: Springer.

Lee, I. (2002). Teaching coherence to ESL students-a classroom inquiry. *Journal of Second Language Writing, 11*, 135-159. doi:160-3743/02/$

Leeser, M. J. (2004). The effects of topic familiarity, mode, and pausing on second language learners' comprehension and focus on form. *Studies in Second Language Acquisition, 26*(4), 587-615.

Levkina, M., & Gilabert, R. (2012). The effects of cognitive task complexity on L2 oral production. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency* (pp. 171-98). Amsterdam: John Benjamins Publishing Company.

Liu, M., & Braine, G. (2005). Cohesive features in argumentative writing produced by Chinese undergraduates. *System, 33*(4), 623-636.

Lix, L. M., & Keselman, H. J. (2010). Analysis of variance: Repeated-measures design. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 15-28). New York: Routledge.

Long, M. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hylstenstam & M. Pienemann (Eds.), *Modelling and Assessing Second Language Acquisition* (pp. 77-99). Clevedon: Multilingual Matters.

Long, M., & Norris, J. (2000). Task-based teaching and assessment. In M. Byram (Ed.), *Encyclopedia of Language Testing* (pp. 597-603). London: Routledge.

Lu, X. (2008). *Automatic measurement of syntactic complexity using the revised developmental level scale.* Paper presented at the Proceedings of the Twenty-First Internatioanl FLAIRS Conference, Key Largo, FL, USA.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics, 15*(4), 474-496.

Lu, X. (2011). A Corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *Tesol Quarterly, 45*(1), 36-62. doi:10.5054/tq.2011.240859

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal, 96*(2), 190-208.

Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing, 29*, 16-27. doi:10.1016/j.jslw.2015.06.003

Makinen, L., Loukusa, S., Nieminen, L., Leinonen, E., & Kunnari, S. (2013). The development of narrative productivity, syntactic complexity, referential cohesion and event content in four-to-eight-year-old Finnish children. *First Language, 34*(1), 24-42. doi:10.1177/0142723713511000

Malicka, A., & Levkina, M. (2012). Measuring task complexity: Does EFL proficiency matter? In A. Shehadeh & C. A. Coombe (Eds.), *Task-based language teaching in foreign langauge context* (pp. 43-66). Amsterdam: John Benjamins Publishing.

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing, 29*, 3-15. doi:10.1016/j.jslw.2015.06.004

McCulley, G. A. (1985). Writing quality, coherence, and cohesion. *Research in the Teaching of English, 19*(3), 269-282.

McMillan, J. H., & Schumacher, S. (2006). *Research in education: Evidence-based inquiry*. USA: Pearson

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2009). Linguistic Features of Writing Quality. *Written Communication, 27*(1), 57-86. doi:10.1177/0741088309351547

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. New York: Cambridge University Press.

McNamara, D. S., Louwerse, M. M., & Graesser, A. C. (2002). *Coh-Metrix: Automated cohesion and coherence scores to predict text readability and facilitate comprehension*. Retrieved from http://129.219.222.66/Publish/pdf/IESproposal.pdf

McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-Metrix: Capturing Linguistic Features of Cohesion. *Discourse Processes, 47*(4), 292-330. doi:10.1080/01638530902959943

Merriam-Webster - Online Dictionary (2016). Retrieved from http://www.merriam-webster.com/dictionary/sentence on 10/25/2016

Muijs, D. (2004). *Doing quantitative research in education with SPSS*. London: SAGE Publications.

Nemati, M. (1999). *The effect of mode of discourse on EFL writing performance.* Unpublished doctoral dissertation, University of Leicester-Tehran.

Norris, J., Brown, J., Hudson, T., & Yoshioka, J. (1998). *Designing second language performance assessments*. Honolulu, HI: University of Hawaii Press.

Nunan, D. (2004). *Task-based language teaching*. New York: Cambridge University Press.

Ong, J., & Zhang, L. J. (2010). Effects of task complexity on the fluency and lexical complexity in EFL students' argumentative writing. *Journal of Second Language Writing, 19*(4), 218-233. doi:10.1016/j.jslw.2010.10.003

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition, 21*(1), 109-148.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*(4), 492-518.

Oshima, A., & Hogue, A. (2006). *Writing academic English*. New York: Pearson Longman.

Parsons, G. (1991). Cohesion-coherence: Scientific texts. In E. Ventola (Ed.), *Functional and systemic linguistics: Approaches and uses* (pp. 415-431). Berlin, New York: Mouton de Gruyter.

Polio, C. (2000). Research methodology in second language writing research:The case of text-based studies. In T. Silva & P. K. Matsuda (Eds.), *On Second Language Writing* (pp. 91-115). New Jersey: Lawrence Erlbaum Associates.

Prater, D. (1985). The effects of modes of discourse, sex of writer, and attitude toward task on writing performance in grade 10. *Educational and Psychological Research, 5*, 241-259.

Prater, D., & Padia, W. (1983). Effects of modes of discourse on writing performance in grades four and six. *Research in the Teaching of English, 17*(2), 127-134.

Préfontaine, Y., & Kormos, J. (2015). The relationship between task difficulty and second language fluency in French: A mixed methods approach. *The Modern Language Journal, 99*(1), 96-112. doi:10.1111/modl.12186

Rahimpour, M., & Hazar, F. (2007). Topic familiarity effect on accuracy, complexity, and fluency of L2 oral output. *The Journal of Asia TEFL, 4*(4), 191-211.

Ravid, D. (2004). Emergence of linguistic complexity in later language development: Evidence from expository text construction. In D. D. Ravid & H. B.-Z. Shyldkrot (Eds.), *Perspectives on language and language development* (pp. 337-355). Dordrecht: Kluwer Academic Publishers.

Raykov, T., & Marcoudiles, G. A. (2008). *An introduction to applied multivariate analysis*. New York, London: Routledge.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Revesz, A. (2013). Towards a fuller assessment of cognitive models of Task-Based Learning: Investigating task-generated cognitive demands and processes. *Applied Linguistics, 35*(1), 87-92. doi:10.1093/applin/amt039

Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition, 31*(3), 437. doi:10.1017/s0272263109090366

Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences - A classroom-based study. *The Modern Language Journal, 95*(Supplementary Issue), 162-181. doi:10.1111/j.1540-4781.2011.01241

Révész, A., & Brunfaut, T. (2013). Text characteristics of task input and difficulty in second language listening comprehension. *Studies in Second Language Acquisition, 35*(1), 31-65. doi:10.1017/s0272263112000678

Richards, B. (1987). Type/token ratios: What do they really tell us? *Journal of Child Language, 14*(2), 201-209.

Richards, J. C., & Rodgers, T. (2001). *Approaches and methods in language teaching*. New York: Cambridge University Press.

Richards, J. C., & Schmidt, R. (2010). Dictionary of language teaching & applied linguistics. Harlow: Longman.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27-57.

Robinson, P. (2003). The Cognition Hypothesis, task design and task-based language learning. *Second Langauge Studies, 21*(2), 45-107.

Robinson, P. (2005). Cognitive complexity and task squencing: A review of studies in a Componential Framework for second language task design. *International Review of Applied Linguistics in Language Teaching, 43*(1), 1-33.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning effects of L2 speech production, interaction, uptake and perceptions of task difficulty. *Iral-International Review of Applied Linguistics in Language Teaching, 45*, 161-176.

Robinson, P., Cadierno, T., & Shirai, Y. (2009). Time and motion: Measuring the effects of the conceptual demands of tasks on second language speech production. *Applied Linguistics, 30*(4), 533-554. doi:10.1093/applin/amp046

Robinson, P., & Gilabert, R. (2007). Task complexity, the Cognition Hypothesis and second language learning and performance. *Iral-International Review of Applied Linguistics in Language Teaching, 45*(3), 161-176.

Rosenberg, S., & Abbeduto, L. (1987). Indicators of linguistic competence in the peer group conversational behavior of mildly retarded adults. *Applied Psycholinguistics, 8*(1), 19-32.

Ruiz-Funes, M. (2015). Exploring the potential of second/foreign language writing for language learning: The effects of task factors and learner variables. *Journal of Second Language Writing, 28,* 1-19. doi:10.1016/j.jslw.2015.02.001

Ryshina-Pankova, M. (2015). A meaning-based approach to the study of complexity in L2 writing: The case of grammatical metaphor. *Journal of Second Language Writing, 29,* 51-63. doi:10.1016/j.jslw.2015.06.005

Salahshuri, S. (2011). The role of background knowledge in foreign language listening comprehension. *Theory and Practice in Language Studies, 1*(10), 1446-1451.

Salimi, A., Alavinia, P., & Hosseini, P. (2012). The effect of strategic planning time and task complexity on L2 written accuracy. *Theory and Practice in Language Studies, 2*(11). doi:10.4304/tpls.2.11.2398-2406

Salimi, A., & Dadaspour, S. (2012). Task complexity and SL development: Does task complexity matter? *Procedia - Social and Behavioral Sciences, 46*, 726-735.

Salimi, A., Dadaspour, S., & Asadollahfam, H. (2011). The effect of task complexity on EFL learners' written performance. *Procedia - Social and Behavioral Sciences, 29*, 1390-1399. doi:10.1016/j.sbspro.2011.11.378

Schmidy-Rinehart, B. C. (1994). The effects of topic familiarity on second language listening comprehension. *The Modern Language Journal, 78*(2), 179-189.

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

Shehadeh, A. (2005). Task-based language learning and teaching: Theories and applications. In C. Edwards & J. Willis (Eds.), *Teachers exploring tasks in English language teaching* (pp. 13-29). New York: Palgrave Macmillan.

Shimoda, T. A. (1993). The effects of interesting examples and topic familiarity on text comprehension, attention, and reading speed. *The Journal of Experimental Education, 61*(2), 93-103.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Skehan, P. (2003). Task-based instruction. *Language Teaching, 36*(1), 1-14. doi:10.1017/s026144480200188x

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510-532. doi:10.1093/applin/amp047

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research, 1*(3), 185-211. doi:10.1177/136216889700100302

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49*, 93-120.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 183-205). New York: Cambridge University Press.

Skehan, P., & Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.), *Planning and task performance in a second language*. Amsterdam: John Benjamins.

Stapleton, P. (2001). Assessing critical thinking in the writing of Japanese university students insights about assumptions and content familiarity. *Written Communication, 18*(4), 506-548.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics*. Boston: Pearson Education.

Tavakoli, P. (2009a). Assessing L2 task performance: Understanding effects of task design. *System, 37*(3), 482-495. doi:10.1016/j.system.2009.02.013

Tavakoli, P. (2009b). Investigating task difficulty: learners' and teachers' perceptions. *International Journal of Applied Linguistics, 19*(1), 1-25.

Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 58*(2), 439-473.

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239-76). Amsterdam: John Benjamins.

Tedick, D. J. (1988). *The effects of topic familiarity on the writing performance of non-native writers of English at the graduate level.* Unpublished doctoral dissertation, Ohio State University.

Tercanlıoğlu, L. (2000). *Linguistics for TEFL students*. İstanbul: Multilingual.

Todd, R. W., Thienpermpool, P., & Keyuravong, S. (2004). Measuring the coherence of writing using topic-based analysis. *Assessing Writing, 9*(2), 85-104. doi:10.1016/j.asw.2004.06.002

Trask, L. (2014). *A student's dictionary of language and linguistics.* New York: Routledge.

Van den Branden, K. (2006). Introduction: Task-based language teaching in a nutsell. In K. Van den Branden (Ed.), *Task-based language education* (pp. 1-16). New York: Cambridge University Press.

Vyatkina, N. (2015). New developments in the study of L2 writing complexity: An editorial. *Journal of Second Language Writing, 29*, 1-2. doi:10.1016/j.jslw.2015.06.001

Watson Todd, R., Khongput, S., & Darasawang, P. (2007). Coherence, cohesion and comments on students' academic essays. *Assessing Writing, 12*(1), 10-25. doi:10.1016/j.asw.2007.02.002

Way, D. P., Joiner, E. G., & Seaman, M. A. (2000). Writing in the secondary foreign language classroom: The effects of prompts and tasks on novice learners of French. *The Modern Language Journal, 84*, 171-184. doi:0026-7902/00

Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.

Williamson, G. (2014). Definition of vocabulary. *Language.* Retrieved from http://www.sltinfo.com/lexical-density/

Willis, D., & Willis, J. (2007). *Doing task-based teaching*. New York: Oxford University Press.

Willis, J. (1996). *A framework for task-based learning.* Palatino, Italy: Longman.

Yang, W. (2014). *Mapping the relationships among the cognitive complexity of independent writing tasks, L2 writing quality, and complexity, accuracy and fluency of L2 writing.* (Unpublished Dissertation Thesis), Georgia State University. Retrieved from http://scholarworks.gsu.edu/alesl_diss/29/

Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing, 28*, 53-67. doi:10.1016/j.jslw.2015.02.002

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*(1), 1-27.

Yule, G. (2010). *The study of language*. Cambridge: Cambridge University Press.

APPENDICES

APPENDIX A. RESULTS OBTAINED FROM AUTOMATED TOOLS AND MANUAL ANALYSIS FOR EACH ESSAY

Description of Writing Tasks

A-Descriptive Writing - Simple Task

B- Descriptive Writing - Complex Task

C- Narrative Writing - Simple Task

D- Narrative Writing - Complex Task

E- Cause-and-Effect - Writing-Simple Task

F- Cause-and-Effect - Complex Task

A. 1. Results of L2 Syntactic Complexity Analyzer by Lu

STUDENT	SYNTACTIC STRUCTURE																	
	Word Count						Sentence						Verb Phrase					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
1	226	194	348	278	335	278,00	17	12	21	15	22	17,00	32	24	57	51	59	47,00
2	195	147	275	178	192	322,00	13	13	17	11	15	17,00	31	21	48	33	28	52,00
3	21	220	360	232	293	446,00	20	21	29	17	20	34,00	28	34	66	41	44	73,00
4	155	469	323	337	285	213,00	15	49	34	43	26	18,00	22	69	60	64	44	25,00
5	253	255	330	312	223	319,00	29	30	45	48	28	38,00	46	42	61	71	38	48,00
6	187	204	241	277	274	287,00	24	17	33	36	30	31,00	33	26	51	60	50	46,00
7	147	248	226	314	207	200,00	13	19	22	31	25	19,00	22	42	41	59	34	30,00
8	157	232	270	471	247	271,00	17	21	30	45	18	24,00	29	36	43	94	46	43,00
9	180	314	307	583	221	221,00	17	32	35	68	19	22,00	28	51	57	107	31	32,00
10	202	182	257	629	400	348,00	16	16	18	53	26	21,00	29	27	32	121	49	47,00
11	242	358	307	330	280	292,00	20	30	36	43	20	19,00	32	37	61	69	42	38,00
12	217	297	305	240	264	527,00	20	25	31	29	23	50,00	30	39	53	42	40	84,00
13	150	182	410	437	242	249,00	17	20	52	54	27	25,00	25	31	84	86	39	37,00
14	212	294	342	480	283	345,00	28	30	40	64	29	36,00	35	40	65	98	41	56,00
15	148	275	331	499	185	188,00	18	35	40	56	21	20,00	25	47	67	96	33	30,00
16	135	382	340	378	295	298,00	11	34	41	38	26	27,00	22	60	68	77	50	50,00
17	209	269	366	493	331	537,00	15	17	29	33	24	34,00	26	35	62	78	49	80,00
18	291	515	309	538	303	373,00	32	58	41	73	31	36,00	44	86	61	122	49	63,00
19	309	302	359	251	215	202,00	8	6	12	31	20	19,00	33	43	59	52	33	32,00
20	211	321	359	251	317	319,00	3	7	12	31	10	32,00	34	31	59	52	49	48,00
21	250	246	393	366	198	263,00	19	20	37	36	20	25,00	33	35	79	79	33	46,00
22	293	391	296	362	313	445,00	26	42	44	53	38	50,00	47	51	60	69	46	65,00

Continued

23	250	259	257	366	259	228,00	23	22	24	44	27	25,00	50	45	53	77	49	42,00
24	273	381	258	648	236	509,00	20	18	19	40	15	28,00	39	48	40	114	29	60,00
25	207	226	356	221	273	247,00	17	18	28	24	23	19,00	24	28	63	42	57	37,00
26	242	194	448	530	277	239,00	27	20	48	66	33	26,00	42	33	91	92	46	43,00
27	249	205	344	333	324	305,00	31	19	35	39	33	34,00	51	38	72	63	53	52,00
28	286	332	380	464	220	345,00	28	31	41	40	20	29,00	41	44	71	72	33	51,00
29	256	286	217	399	267	301,00	35	36	26	50	32	28,00	48	42	46	89	47	41,00
30	242	258	416	365	292	217,00	23	24	41	35	21	18,00	38	32	77	68	48	27,00
31	242	570	561	336	304	331,00	24	45	52	31	27	28,00	38	67	90	57	44	47,00
32	174	311	306	289	234	231,00	16	19	27	28	19		21	35	57	51	35	34,00
33	489	342	429	582	348	301,00	45	31	39	59	30	21,00	74	48	81	112	48	41,00
34	280	408	509	494	383	208,00	17	32	38	43	23	15,00	41	65	94	105	67	34,00
35	176	270	341	346	316	291,00	17	29	34	41	34	28,00	27	43	67	68	53	52,00
36	195	142	590	246	279	298,00	10	8	39	24	22	26,00	32	25	96	57	53	45,00
37	236	351	297	407	286	385,00	21	27	31	43	21	29,00	41	51	50	80	48	52,00
38	262	343	330	219	253	342,00	22	32	33	24	22	23,00	40	55	55	37	33	47,00
39	195	229	239	295	235	202,00	18	24	20	20	18	16,00	34	42	38	48	45	32,00
40	166	211	254	260	242	280,00	14	17	21	24	18	23,00	24	27	35	42	29	36,00
41	310	320	270	266	383	303,00	39	33	40	33	43	32,00	52	41	54	54	64	47,00

STUDENT	SYNTACTIC COMPLEXITY																	
	LENGTH OF PRODUCTION																	
	Mean Length of Sentence						Mean Length of T-Unit						Mean Length of Clause					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
1	13,29	16,17	16,57	18,53	15,23	16,35	11,89	11,41	10,24	12,09	14,57	14,63	8,37	9,24	7,91	8,18	8,82	9,27
2	15,00	11,31	16,18	16,18	12,80	18,94	13,00	11,31	10,19	12,71	10,67	16,10	7,50	7,35	7,05	8,48	7,68	8,47
3	10,80	10,48	12,41	13,65	14,65	13,12	9,82	10,48	9,73	12,21	13,95	12,05	8,31	8,15	7,06	10,55	8,37	7,43
4	10,33	9,57	9,50	7,84	10,96	11,83	9,69	8,85	7,88	7,66	9,83	11,83	9,12	9,02	7,18	6,48	8,38	10,14
5	8,72	8,50	7,33	6,50	7,96	8,39	8,72	8,50	7,50	6,24	7,96	9,11	5,88	7,97	7,67	6,24	7,19	7,60
6	7,79	12,00	7,30	7,69	9,13	9,26	7,48	11,33	7,09	7,10	8,84	9,26	5,84	7,85	5,74	6,44	7,41	6,67
7	11,31	13,05	10,27	10,13	8,28	10,53	11,31	11,81	9,83	8,72	7,96	10,00	8,17	6,70	6,11	6,16	7,14	8,33
8	9,24	11,05	9,00	10,47	13,72	11,29	8,26	8,92	8,18	7,36	10,74	10,04	6,54	6,82	6,43	5,61	7,26	7,53
9	10,59	9,81	8,77	8,57	11,63	10,05	9,47	9,81	8,30	7,99	11,63	10,05	7,50	7,48	5,79	6,63	8,50	8,19
10	12,63	11,38	14,28	11,87	15,38	16,57	11,88	10,11	10,71	8,99	13,33	12,00	8,42	8,27	8,86	6,17	9,76	9,67
11	12,10	11,93	8,53	7,67	14,00	15,37	11,52	11,55	7,49	7,33	12,73	13,90	8,34	10,53	5,48	5,79	10,00	10,81
12	10,85	11,88	9,84	8,28	11,48	10,54	9,86	11,42	8,71	8,00	9,78	9,76	8,35	8,03	6,35	6,49	8,00	7,42
13	8,82	9,10	7,88	8,09	8,96	9,96	8,33	9,10	7,88	7,53	8,96	10,38	7,14	7,28	5,39	5,99	6,91	7,78
14	7,57	9,80	8,55	7,50	9,76	9,58	7,31	9,48	7,28	6,58	9,76	8,85	6,24	8,65	5,90	5,52	7,86	6,90
15	8,22	7,86	8,28	8,91	8,81	9,40	8,22	7,64	8,28	8,46	8,41	9,89	6,43	5,85	5,91	6,24	6,61	8,95
16	12,27	11,24	8,29	9,95	11,35	11,04	10,38	10,61	7,91	8,40	10,93	9,31	6,14	7,49	5,76	5,91	7,38	7,10
17	13,93	15,82	12,62	14,94	13,79	15,79	13,93	14,94	11,44	12,97	13,24	15,79	9,50	9,61	7,32	8,36	8,49	9,26
18	9,09	8,88	7,54	7,37	9,77	10,36	8,82	8,58	7,19	6,40	8,91	9,56	7,46	6,87	5,62	5,12	8,19	7,74
19	38,63	50,33	29,92	8,10	10,75	10,63	25,75	20,13	17,95	8,10	10,75	10,63	10,30	7,55	6,65	6,44	8,60	7,77
20	70,33	45,86	29,92	8,10	31,70	9,97	70,33	26,75	17,95	8,10	31,70	9,11	6,59	12,84	6,65	6,44	9,61	8,86
21	13,16	12,30	10,62	10,17	9,90	10,52	13,89	11,71	9,36	8,93	9,00	10,52	9,26	8,20	5,54	6,20	6,83	7,11

Continued

22	11,27	9,31	6,73	6,83	8,24	8,90	10,10	9,09	6,43	6,83	8,46	8,56	6,23	7,98	5,29	6,03	8,03	7,54
23	10,87	11,77	10,71	8,32	9,59	9,12	11,90	10,79	10,71	7,79	8,63	9,12	6,25	6,48	5,14	5,15	7,40	6,51
24	13,65	21,17	13,58	16,20	15,73	18,18	14,37	12,70	10,32	12,00	15,73	16,42	8,53	9,29	7,59	7,62	11,24	13,76
25	12,18	12,56	12,71	9,21	11,87	13,00	10,35	11,30	8,28	8,19	11,87	12,35	9,86	10,27	6,59	6,31	8,27	9,50
26	8,96	9,70	9,33	8,03	8,39	9,19	8,96	9,24	8,00	7,91	8,66	9,56	6,72	7,19	5,46	6,79	7,69	8,24
27	8,03	10,79	9,83	8,54	9,82	8,97	7,78	10,79	9,30	8,33	9,00	8,97	5,41	6,41	4,91	5,84	6,89	6,78
28	10,21	10,71	9,27	11,60	12,00	11,90	9,23	10,38	7,92	9,87	9,57	11,50	7,94	9,22	6,33	8,29	8,15	8,85
29	7,31	7,94	8,35	7,98	8,34	10,75	6,92	7,94	7,48	6,54	7,42	10,03	5,69	7,33	5,29	5,70	6,51	8,14
30	10,52	10,75	10,15	10,43	13,90	12,06	9,31	10,32	8,32	8,49	13,90	11,42	7,12	8,60	6,40	6,29	8,59	9,04
31	10,08	12,67	10,79	10,84	11,26	11,82	8,96	11,88	10,20	9,60	12,16	10,68	7,12	8,91	7,10	6,86	8,22	8,07
32	10,88	16,37	11,33	10,32	12,32	12,16	10,24	15,55	9,00	7,81	10,17	10,04	9,16	10,03	6,51	6,42	8,07	7,22
33	10,87	11,03	11,00	9,86	11,60	14,33	9,23	9,50	8,41	8,20	9,94	11,15	7,19	8,14	6,04	6,47	8,29	8,60
34	16,47	12,75	13,39	11,49	16,65	13,87	13,33	10,74	9,09	8,52	13,68	10,40	11,20	7,42	6,06	6,02	7,51	8,00
35	10,35	9,31	10,03	8,44	9,29	10,39	8,80	8,71	9,22	6,78	8,78	9,39	7,33	7,30	7,58	5,86	7,71	6,77
36	19,50	17,75	15,13	10,25	12,68	11,46	16,25	11,83	9,37	9,46	9,30	9,03	7,50	6,76	6,86	5,47	7,15	6,93
37	11,24	13,00	9,58	9,47	13,62	13,28	10,26	12,54	8,49	8,31	11,44	13,28	7,15	8,36	6,46	5,99	7,94	9,17
38	11,91	10,72	10,00	9,13	11,50	14,87	9,36	9,03	8,46	8,42	10,54	12,21	7,49	7,15	7,17	8,11	8,16	10,36
39	10,83	9,54	11,95	14,75	13,06	12,63	8,13	7,63	8,85	14,05	10,68	1,88	6,72	6,74	7,03	8,68	8,70	6,92
40	11,86	12,41	12,10	10,83	13,44	12,17	11,86	11,72	12,10	9,63	11,52	9,66	7,90	8,12	8,19	7,43	8,64	1,26
41	7,95	9,70	6,75	8,06	8,91	9,47	7,56	9,14	6,28	8,06	8,70	9,18	6,89	8,21	5,51	5,91	7,09	7,21

A. 2. Results for Lexical Complexity Analyzer by Lu

STUDENT	LEXICAL COMPLEXITY																							
													Lexical Variety											
	Lexical Density						Lexical Sophistication						Type-token Ratio						Lexical Word Variation					
	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F	A	B	C	D	E	F
1	0,53	0,5	0,47	0,44	0,49	0,53	0,16	0,14	0,2	0,14	0,24	0,24	0,43	0,4	0,42	0,39	0,46	0,47	0,55	0,65	0,66	0,64	0,56	0,49
2	0,51	0,51	0,49	0,48	0,49	0,56	0,25	0,42	0,32	0,26	0,33	0,3	0,52	0,54	0,48	0,5	0,51	0,6	0,66	0,75	0,78	0,78	0,65	0,66
3	0,5	0,5	0,5	0,53	0,53	0,51	0,43	0,34	0,23	0,28	0,29	0,25	0,41	0,52	0,44	0,42	0,54	0,45	0,78	0,62	0,58	0,67	0,55	0,59
4	0,51	0,52	0,48	0,5	0,56	0,52	0,33	0,34	0,21	0,24	0,25	0,39	0,48	0,48	0,49	0,51	0,62	0,46	0,84	0,69	0,69	0,7	0,64	0,67
5	0,51	0,55	0,46	0,42	0,57	0,51	0,27	0,43	0,23	0,29	0,31	0,23	0,47	0,37	0,5	0,41	0,53	0,6	0,67	0,82	0,7	0,59	0,62	0,57
6	0,43	0,54	0,49	0,52	0,54	0,53	0,35	0,32	0,19	0,26	0,31	0,3	0,51	0,42	0,38	0,4	0,45	0,5	0,69	0,65	0,7	0,58	0,51	0,55
7	0,5	0,48	0,46	0,44	0,6	0,58	0,35	0,31	0,19	0,18	0,27	0,18	0,54	0,44	0,51	0,57	0,52	0,48	0,65	0,68	0,79	0,7	0,61	0,65
8	0,44	0,49	0,41	0,42	0,5	0,48	0,39	0,37	0,27	0,21	0,3	0,22	0,41	0,37	0,45	0,4	0,45	0,51	0,62	0,71	0,6	0,59	0,58	0,58
9	0,53	0,5	0,44	0,49	0,54	0,57	0,4	0,35	0,19	0,2	0,29	0,38	0,4	0,38	0,47	0,48	0,49	0,45	0,65	0,65	0,59	0,58	0,59	0,6
10	0,45	0,5	0,53	0,44	0,6	0,55	0,22	0,24	0,24	0,18	0,17	0,22	0,54	0,33	0,37	0,37	0,45	0,49	0,64	0,71	0,74	0,54	0,47	0,51
11	0,49	0,53	0,46	0,44	0,53	0,51	0,36	0,44	0,22	0,17	0,26	0,27	0,42	0,42	0,46	0,48	0,51	0,5	0,73	0,67	0,6	0,62	0,66	0,69
12	0,58	0,58	0,5	0,57	0,58	0,6	0,4	0,49	0,25	0,34	0,34	0,28	0,5	0,52	0,44	0,66	0,54	0,56	0,75	0,71	0,73	0,61	0,57	0,61
13	0,6	0,53	0,48	0,5	0,58	0,53	0,33	0,32	0,22	0,23	0,29	0,27	0,4	0,41	0,5	0,55	0,54	0,52	0,62	0,74	0,62	0,6	0,61	0,74
14	0,48	0,51	0,5	0,48	0,46	0,47	0,37	0,24	0,2	0,19	0,26	0,29	0,45	0,37	0,39	0,43	0,44	0,41	0,64	0,59	0,69	0,56	0,57	0,62
15	0,57	0,52	0,51	0,53	0,58	0,61	0,39	0,38	0,35	0,3	0,28	0,29	0,44	0,4	0,49	0,52	0,61	0,53	0,76	0,78	0,62	0,56	0,55	0,62
16	0,58	0,51	0,44	0,49	0,52	0,47	0,27	0,42	0,13	0,17	0,23	0,21	0,43	0,42	0,39	0,44	0,59	0,5	0,74	0,73	0,66	0,61	0,53	0,66
17	0,58	0,49	0,48	0,48	0,52	0,49	0,34	0,37	0,21	0,19	0,24	0,21	0,46	0,38	0,49	0,38	0,6	0,48	0,7	0,69	0,68	0,6	0,62	0,56
18	0,54	0,5	0,52	0,49	0,5	0,53	0,42	0,34	0,3	0,22	0,3	0,22	0,48	0,37	0,49	0,46	0,49	0,4	0,67	0,62	0,69	0,58	0,7	0,65
19	0,51	0,55	0,53	0,55	0,56	0,54	0,43	0,48	0,4	0,29	0,27	0,35	0,49	0,48	0,51	0,55	0,53	0,59	0,77	0,84	0,74	0,64	0,66	0,7
20	0,51	0,59	0,58	0,57	0,53	0,55	0,5	0,59	0,41	0,4	0,41	0,31	0,55	0,46	0,38	0,41	0,49	0,57	0,73	0,82	0,73	0,64	0,52	0,57
21	0,47	0,54	0,42	0,47	0,54	0,56	0,36	0,31	0,18	0,19	0,23	0,23	0,37	0,4	0,58	0,53	0,53	0,56	0,78	0,75	0,59	0,58	0,69	0,68
22	0,51	0,52	0,47	0,57	0,53	0,49	0,29	0,38	0,24	0,26	0,3	0,29	0,47	0,42	0,47	0,44	0,48	0,48	0,63	0,65	0,71	0,52	0,67	0,65
23	0,52	0,52	0,45	0,49	0,51	0,59	0,28	0,41	0,18	0,17	0,31	0,23	0,52	0,42	0,49	0,56	0,52	0,52	0,72	0,76	0,79	0,62	0,71	0,71

Continued

24	0,51	0,55	0,53	0,48	0,56	0,58	0,35	0,36	0,31	0,25	0,36	0,29	0,51	0,36	0,55	0,41	0,45	0,46	0,61	0,66	0,65	0,56	0,72	0,54
25	0,57	0,55	0,47	0,46	0,64	0,6	0,36	0,3	0,24	0,26	0,19	0,27	0,38	0,48	0,41	0,43	0,54	0,46	0,68	0,6	0,57	0,66	0,47	0,51
26	0,26	0,51	0,42	0,51	0,58	0,55	0,51	0,2	0,16	0,3	0,29	0,21	0,35	0,3	0,38	0,68	0,44	0,44	0,58	0,55	0,56	0,41	0,47	0,46
27	0,51	0,54	0,41	0,51	0,51	0,47	0,24	0,22	0,18	0,3	0,22	0,22	0,42	0,3	0,46	0,44	0,52	0,5	0,69	0,64	0,68	0,41	0,64	0,65
28	0,46	0,45	0,46	0,51	0,51	0,48	0,18	0,21	0,21	0,28	0,24	0,26	0,43	0,43	0,49	0,39	0,4	0,4	0,6	0,64	0,63	0,6	0,66	0,56
29	0,53	0,57	0,47	0,54	0,63	0,57	0,32	0,45	0,28	0,3	0,27	0,36	0,47	0,4	0,45	0,43	0,38	0,41	0,56	0,56	0,65	0,56	0,53	0,53
30	0,48	0,48	0,45	0,49	0,54	0,54	0,38	0,34	0,18	0,17	0,27	0,3	0,35	0,39	0,39	0,5	0,44	0,41	0,62	0,58	0,56	0,59	0,52	0,63
31	0,48	0,49	0,48	0,42	0,47	0,51	0,35	0,41	0,26	0,17	0,28	0,27	0,34	0,38	0,41	0,46	0,51	0,41	0,74	0,62	0,52	0,63	0,58	0,62
32	0,56	0,53	0,43	0,53	0,55	0,53	0,44	0,39	0,17	0,21	0,31	0,3	0,41	0,43	0,45	0,48	0,53	0,52	0,68	0,71	0,62	0,59	0,57	0,66
33	0,45	0,5	0,43	0,46	0,48	0,53	0,35	0,35	0,23	0,14	0,31	0,22	0,37	0,35	0,35	0,43	0,41	0,45	0,65	0,66	0,56	0,58	0,48	0,57
34	0,45	0,48	0,46	0,51	0,47	0,52	0,22	0,35	0,21	0,26	0,24	0,28	0,33	0,33	0,32	0,48	0,36	0,41	0,52	0,64	0,49	0,48	0,49	0,58
35	0,48	0,49	0,45	0,52	0,49	0,47	0,4	0,38	0,28	0,24	0,26	0,25	0,37	0,41	0,44	0,38	0,53	0,44	0,76	0,69	0,51	0,56	0,62	0,57
36	0,45	0,46	0,45	0,48	0,46	0,46	0,36	0,26	0,27	0,16	0,28	0,24	0,27	0,44	0,47	0,49	0,51	0,58	0,79	0,86	0,4	0,59	0,72	0,77
37	0,49	0,52	0,45	0,43	0,47	0,54	0,26	0,28	0,1	0,19	0,22	0,24	0,46	0,41	0,4	0,44	0,48	0,46	0,66	0,66	0,68	0,65	0,57	0,59
38	0,54	0,5	0,47	0,5	0,54	0,54	0,29	0,39	0,21	0,37	0,3	0,21	0,41	0,47	0,43	0,39	0,43	0,4	0,58	0,6	0,61	0,62	0,55	0,53
39	0,49	0,5	0,56	0,49	0,6	0,51	0,29	0,3	0,3	0,28	0,31	0,27	0,54	0,42	0,47	0,46	0,51	0,49	0,73	0,72	0,72	0,62	0,56	0,58
40	0,55	0,49	0,55	0,53	0,51	0,51	0,31	0,32	0,14	0,31	0,29	0,26	0,5	0,5	0,44	0,49	0,54	0,51	0,66	0,69	0,67	0,64	0,56	0,63
41	0,48	0,53	0,52	0,52	0,53	0,56	0,23	0,37	0,32	0,23	0,28	0,26	0,46	0,44	0,39	0,49	0,39	0,47	0,57	0,63	0,59	0,57	0,51	0,61

A. 3. Results of Cohesion by Coh-Metrix 3.0

STUDENT	COHESION											
	Referential Cohesion						Deep Cohesion					
	A	B	C	D	E	F	A	B	C	D	E	F
1	43	62	50	96	73	77	92	92	78	56	99	99
2	81	16	10	44	45	44	52	6	22	94	99	57
3	19	52	56	39	75	37	70	48	100	75	99	100
4	31	15	45	43	39	32	52	32	98	59	100	99
5	47	2	8	44	11	50	85	21	78	61	84	62
6	70	12	21	60	70	58	98	97	94	65	98	83
7	34	38	23	67	33	14	85	70	63	95	100	47
8	39	20	69	62	97	82	97	99	93	42	47	64
9	73	38	89	45	86	54	31	92	87	84	94	90
10	31	66	42	58	67	81	83	99	22	50	37	95
11	34	12	57	77	47	38	96	77	78	65	95	99
12	21	3	35	50	86	27	25	26	53	91	100	96
13	26	21	80	29	45	11	97	78	93	89	100	89
14	42	48	40	52	64	44	86	70	89	53	100	94
15	3	12	31	64	73	23	48	40	53	94	83	97
16	18	17	35	78	67	44	92	94	98	90	100	96
17	24	50	44	71	27	37	60	21	87	50	90	95
18	8	11	35	64	11	20	22	72	71	30	90	81
19	14	20	36	38	29	24	100	89	95	15	99	85
20	77	22	47	35	99	44	100	31	85	83	100	98
21	41	28	87	62	32	21	17	33	69	47	100	99
22	24	13	25	31	8	22	51	38	99	83	81	77
23	13	24	68	65	49	30	87	91	100	88	99	93
24	50	71	67	44	35	36	99	93	62	95	98	99
25	13	38	69	81	64	74	85	92	89	40	100	80
26	53	90	93	96	50	87	42	95	42	59	100	100
27	23	30	74	44	27	27	95	43	96	93	96	96
28	47	61	17	34	50	54	16	18	69	53	97	88
29	81	77	62	65	64	74	85	68	69	92	99	89
30	64	55	73	62	85	74	77	71	94	34	57	99
31	25	33	38	69	60	16	97	66	82	62	89	66
32	46	18	46	90	67	33	71	90	59	76	99	95
33	59	18	86	68	81	55	69	89	71	70	100	100
34	79	42	89	86	99	79	67	71	57	85	99	92
35	59	59	58	83	40	85	92	72	64	66	91	91
36	55	26	57	49	71	41	98	82	90	76	93	48
37	41	27	28	44	76	32	75	73	57	93	99	94
38	82	61	62	57	84	63	78	58	25	40	99	60
39	46	31	13	81	40	71	83	62	97	52	100	32
40	55	61	42	47	50	83	98	54	62	96	98	46
41	40	16	22	50	28	20	92	71	89	92	100	99

A. 4. Results of Coherence by the Analytic Rubric

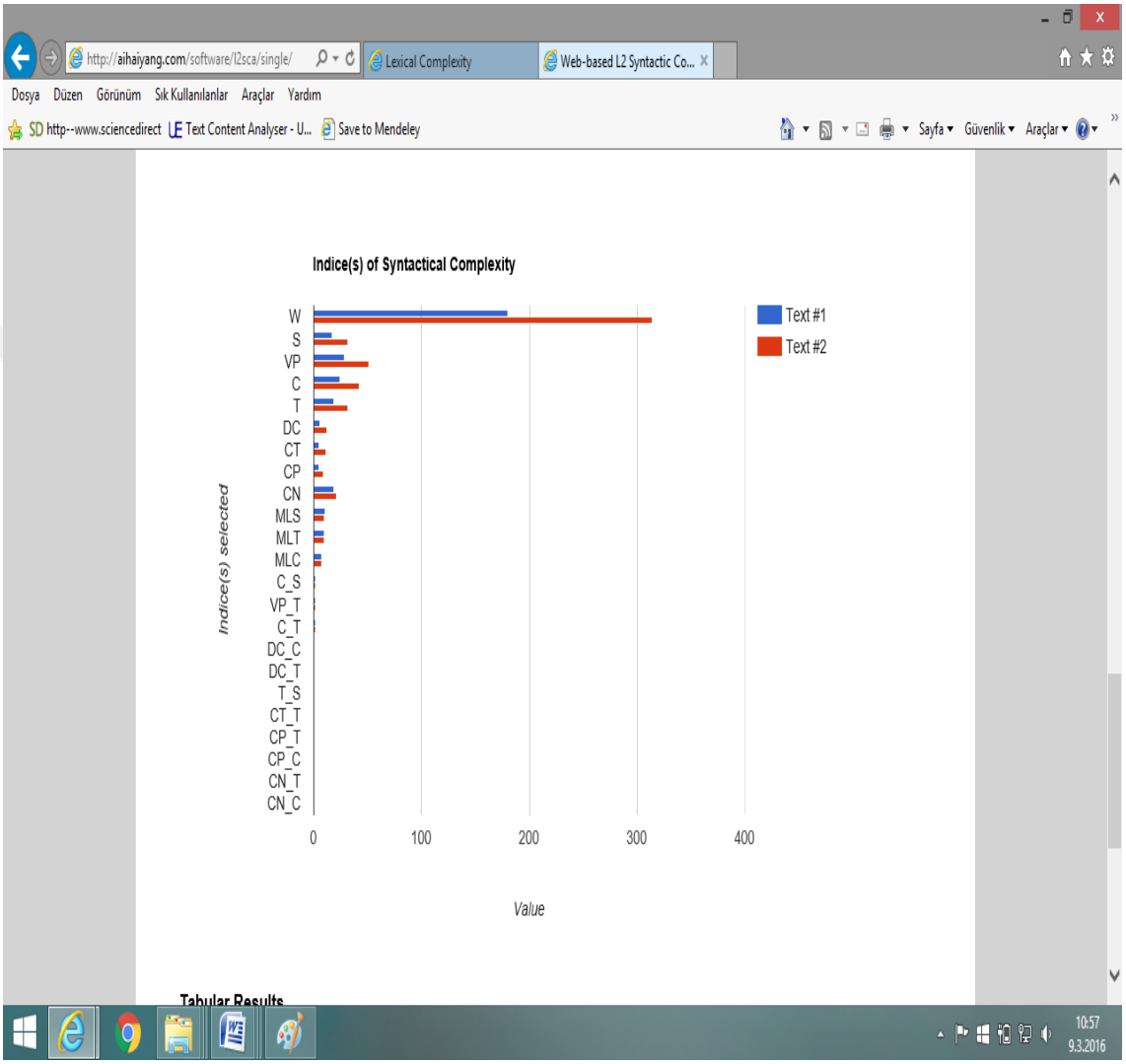
STUDENT	COHERENCE					
	A	**B**	**C**	**D**	**E**	**F**
1	4	3,5	3,5	3,5	3	3
2	4	2,5	2,5	3	2,5	3
3	4	3	2,5	4	2,5	3,5
4	3	4,5	3	4	3,5	3,5
5	3	3,5	3,5	4	3	3,5
6	2,5	3,5	2,5	3,5	3,5	3,5
7	3	4	4	4	3	3,5
8	3	3,5	4	3	3	3
9	3	3	4	3	3	3
10	3	3	3,5	2,5	3,5	3,5
11	3,5	3	3,5	3,5	3,5	3,5
12	3,5	3,5	3,5	4	3,5	3,5
13	3	3	4	3,5	3,5	3,5
14	3	3	4	4	3	3,5
15	3,5	4,5	3,5	3	3,5	3
16	3,5	4,5	3,5	3,5	4	4
17	4,5	5	5	4,5	4,5	5
18	4	4,5	5	4	3,5	4
19	3,5	3,5	3,5	4	4	2,5
20	3	3,5	4	2,5	3	3
21	4	4	3,5	4,5	4	3,5
22	3,5	3,5	3	3	3	3
23	3	2,5	3	3	2,5	3
24	3	4	4,5	2,5	3	3,5
25	3	3,5	4	4	3,5	3,5
26	3,5	4	3	3,5	3,5	3
27	2,5	2,5	3	3,5	3	2,5
28	3,5	4,5	4,5	3,5	2,5	3,5
29	3,5	3	3	3,5	2,5	3
30	4	3,5	3,5	3	4	4
31	2,5	3,5	2,5	3	3	3
32	3	3	2,5	2,5	3	3
33	4	4,5	3,5	3,5	4	4
34	3	4	4	3	3	3
35	3,5	3	3,5	2,5	4	3
36	3	4	2,5	3,5	3,5	3,5
37	3,5	3,5	3	2,5	3	3
38	4	3,5	3	3	4	4
39	3,5	3,5	2,5	2	3	2,5
40	3,5	3	2,5	3	2,5	2,5
41	3,5	3,5	3	2,5	4	4

A. 5. Results of Overall Writing Quality by the Analytic Rubric

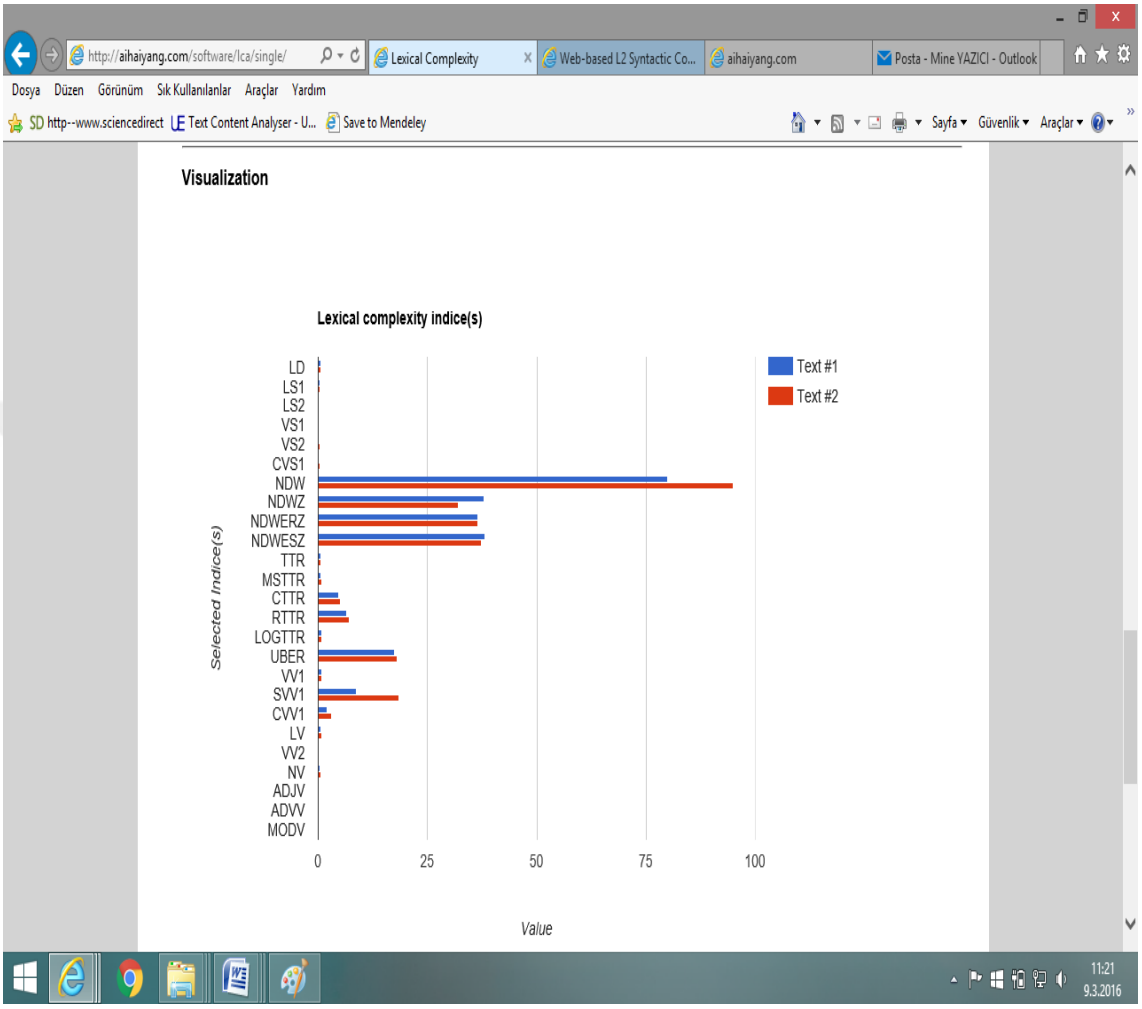
STUDENT	OVERALL WRITING QUALITY					
	A	**B**	**C**	**D**	**E**	**F**
1	17,5	16,5	14,5	18,5	16,5	15
2	16,5	13,5	12,5	15,5	13,5	15,5
3	18,5	16	13	18,5	14,5	17
4	15	21	15,5	16,5	18,5	17,5
5	16	18	18,5	19	16,5	18
6	15	18	14,5	16	17	17,5
7	15,5	18,5	18	18,5	16	17,5
8	14	15,5	17,5	13	16,5	15,5
9	16	17	19	13,5	14	15
10	14	15,5	15	14	18	17,5
11	17,5	18	15,5	15	18	19
12	17,5	17,5	15,5	17,5	18,5	16,5
13	14	15	16,5	16	17,5	19
14	14,5	16,5	19	18,5	16	16
15	16,5	20,5	16,5	14,5	16,5	15,5
16	16	21,5	16,5	16	18,5	20
17	21,5	24	22,5	20,5	22	23
18	18,5	22	23,5	19	18,5	19,5
19	17,5	19	16,5	17	20	13,5
20	17	18	20	15	16	16
21	18,5	19	18	20,5	20	18,5
22	17	20	15	15,5	15,5	15,5
23	15,5	15	16	15,5	16	16,5
24	14,5	16	19	13,5	16	20
25	16,5	18	19	18	17,5	18,5
26	16	18	17,5	17	17	15
27	15	15,5	15	16	14,5	15
28	15,5	22	20	17	15	17
29	18	15	16	16	15	16
30	17,5	15,5	18	16,5	18	20
31	14	18	13,5	15	15,5	16
32	15,5	15,5	14,5	14,5	17	16
33	19,5	22	19,5	17,5	20,5	21,5
34	16,5	23	20	16,5	17,5	16
35	16	15	16	14,5	19,5	15,5
36	17	20	14,5	17,5	19	19,5
37	15,5	16	15,5	14,5	15,5	15,5
38	17	20,5	14,5	16,5	20	19
39	17	18	14,5	12,5	16,5	14
40	15,5	14,5	13	15	13,5	15
41	16,5	18,5	15	14	19	18,5

APPENDIX B. SAMPLES FOR DATA ANALYSIS BY AUTOMATED ANALYZER TOOLS

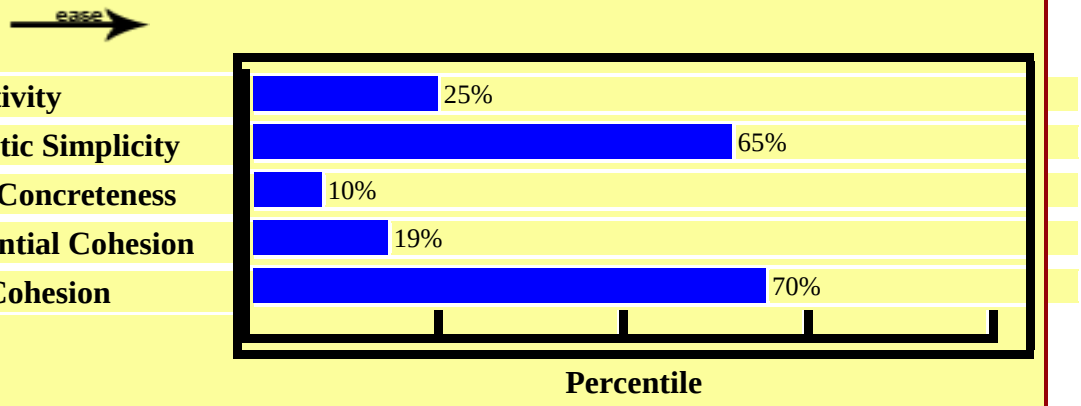
B. 1. Sample Analysis of Syntactic Complexity by Lu's L2 Syntactic Complexity Analyzer



B. 2. Sample analysis of lexical complexity by Lu's Lexical Complexity Analyzer



B.3. Sample analysis of cohesion by Coh-Metrix 3.0



Flesch Kincaid Grade Level 7

This text is low in narrativity which indicates that it is less story-like and may have less familiar words. Less story-like texts are usually harder to comprehend. It has low word concreteness, which means there are many abstract words that are hard to visualize. Abstract texts may be more difficult to understand. This text has low referential cohesion, indicating little overlap in words and ideas between sentences. Cohesion gaps require the reader to make inferences, which can be challenging and even unsuccessful without sufficient prior knowledge. It is high in deep cohesion. There are relatively more connecting words to help clarify the relationships between events, ideas, and information. Because of this added support, comprehension may be facilitated, especially when the topic is unfamiliar.

	The A Paper	The B Paper	The C Paper	The D Paper	The F Paper
Ideas	Excels in responding to assignment. Interesting, demonstrates sophistication of thought. Central idea/thesis is clearly communicated, and is worth developing but limited enough to be manageable. Paper recognizes some complexity of its thesis: may acknowledge contradictions, qualifications, or limits of claims and follow out to their logical conclusions. Understands and critically evaluates its sources, while appropriately limiting and defining terms.	A solid paper, responding appropriately to assignment. Clearly states a thesis/central idea, but may have minor lapses in development. Begins to acknowledge the complexity of central idea and the possibility of other points of view, but may fail to develop these insights. Shows careful reading of sources, but may not evaluate them critically. Attempts to define terms, not always successfully.	Adequate but weaker and less effective, possibly responding less well to assignment. Presents central idea in general terms, often depending on platitudes or clichés. Usually does not acknowledge other views or counterarguments. Demonstrates basic comprehension of sources, perhaps with lapses in understanding. If it defines terms, it may depend on dictionary definitions.	Does not have a clear central idea or does not respond appropriately to the assignment. Thesis may be too vague or obvious to be developed effectively. Paper may misunderstand sources.	Does not respond to the assignment, lacks a thesis or central idea, and may neglect to use sources where necessary.
Organization & Coherence	Uses a logical structure that is appropriate to paper's subject, purpose, audience, thesis, and disciplinary field. Employes sophisticated transitional sentences which often develop one idea from the previous one or identify their logical relations. It clearly guides the reader through the chain of reasoning or progression of ideas.	Shows a logical progression of ideas and uses fairly sophisticated transitional devices; e.g., may move from least to more important idea. Some logical links may be faulty, but each paragraph clearly relates to paper's central idea.	May list ideas or arrange them randomly rather than using any evident logical structure. May use transitions, but they are likely to be sequential (first, second, third) rather than logic-based. While each paragraph may relate to central idea, logic is not always clear. Paragraphs have topic sentences but may be overly general, and arrangement of sentences within paragraphs may lack coherence.	May have random organization, lacking internal paragraph coherence and using few or inappropriate transitions. Paragraphs may lack topic sentences or main ideas, or may be too general or too specific to be effective. Paragraphs may not all relate to paper's thesis.	No appreciable organization; lacks transitions and coherence.
Support	Uses evidence appropriately and effectively, providing sufficient evidence and explanation to convince. Evaluates the strengths and weaknesses of evidence, and offers clear reasons for which evidence is strongest, and why this evidence is compelling. Provides clear citations and uses quotes well.	Begins to offer reasons to support its points, perhaps using varied kinds of evidence. Begins to interpret the evidence and explain connections between evidence and main ideas. Its examples bear some relevance. More implied as compared to an A Paper's explicitness. Usually provides adequate citations and good use of quotations.	Often uses generalizations to support its points. May use examples, but they may be obvious or not relevant. Often depends on unsupported opinion or personal experience, or assumes that evidence speaks for itself and needs no application to the point being discussed. Often has lapses in logic, and incomplete citations with poorly used or explained quotations.	Depends on clichés or overgeneralizations for support, or offers little evidence of any kind. May rely on personal narrative rather than evidence-based argument, or summary rather than analysis. Often fails to include appropriate citations: quotes are either absent or overly-long.	Uses irrelevant details or lacks supporting evidence entirely. May be unduly brief. Usually fails to include appropriate citations.
Style	Chooses words for their precise meaning and uses an appropriate level of specificity. Sentence style fits paper's audience and purpose. Sentences are varied, yet clearly structured and carefully focused.	Generally uses words accurately and effectively, but may sometimes be too general. Sentences generally clear, well-structured, and focused, though some may be awkward or ineffective.	Uses relatively vague and general words, may use some inappropriate language. Sentence structure is generally correct, but sentences may be wordy, unfocused, repetitive, or confusing.	May be too vague and abstract, or very personal and specific. Usually contains several awkward or ungrammatical sentences; sentence structure is simple or monotonous.	Usually contains many awkward sentences, misuses words, employs inappropriate language.
Mechanics	Almost entirely free of spelling, punctuation, and grammatical errors.	May contain a few errors, which are obvious to the reader but do not impede understanding.	Usually contains several mechanical errors, which may temporarily confuse the reader but not impede the overall understanding.	Usually contains either many mechanical errors or a few important errors that prevent the reader from understanding and impede her ability to see connections between thoughts.	Usually contains so many mechanical errors that it is impossible for the reader to follow the thinking from sentence to sentence.

(Based on a rubric used by the UC Davis English Department Composition Program, located online at www.winona.edu/AIR/documents/termpaper.pdf)

APPENDIX D. SAMPLE ESSAYS WRITTEN BY STUDENTS

D.1. A Sample essay for the complex task of descriptive writing

Pearl of The Black Sea

I think all people have a special place which makes them happy in their lives. This special place is Samsun in my life. There are many factors that make it a wonderful place to me.

Firstly, Samsun is meaningful to me because it is my hometown. It is part of my childhood. This factor creates an emotional bond between me and there. When I walk on one of its streets, my memories come to my mind, and I definitely see many familiar faces. I do not myself like a stranger in Samsun.

Secondly, Samsun is located in Black Sea where blue and green bring together. It has such a vast sea that you dream away while watching it. It also has such a huge urban forest that you feel as if you breathed for the first time in your life while walking among the trees. All these mean peace.

Another factor that makes Samsun meaningful to me is its natural beauty. Bird Paradise which hosts numerous bird species is just one natural beauty of it. You can go there and have a picnic. You can also go fishing in the lake which has a wonderful view. Samsun offers many places where you can spend your time nicely.

In sum, factors which make Samsun a special place to me are its location, its natural beauty, its facilities and being my hometown. I believe that it should be in the list of places to visit.

D. 2. A Sample Essay for the Simple Task of Cause-and-Effect Essay

Effects Of Overusing The Internet

Today, Internet is a communication tool that everybody uses it unexaptionally. Internet has such a big place in our lifes that it is impossible to live without it. It has many functions which are useful to us. But like everything when we overuse the internet, it can effect our lifes in different ways. Three of them, which effects our lifes at most are addiction, insomia and being asocial.

Firstly, people who use internet for everything become addicted to it after a while, even if accidentally. At the first times of using internet, everything seems to be normal. But when we get used to using it, we spend more than the normal time that is supposed to be spent. There is no reason for anyone to spend his/her all day just with internet. Due to the internet, people often forget that there is also a life out of the internet. They become addicted and they don't spend time even with their family. In some situations, internet can be so effective that people can forget even eating something.

Our second important effect is insomia, which means "lack of sleep". Every child has to go to school until a certain age. That means all of them have to get up early. But most of the children sit in front of the computer for surfing on the internet. They log in the "Facebook", watch a video or maybe a movie on "youtube" for

hours without understanding how they waste their time. Suddenly, they look at the clock and realize that it is very late. On the next day, because of overusing internet, they go to school with homeworks they have forgotten to do. And at school, most of them fell asleep due to wasting time with internet for hours. The experts have explained for a thousand time how important sleeping is for human, especially for children.

And the third, according to me the most important effect is being asocial. Overusing the internet is dangerous for children, young people, old people, . . . actually for everyone. People use internet because they think it necessary for every section. Even housewifes use it to quiet their children by opening a cartoon movie or music which soothes them thanks to the internet. They sit their children in front of the computer and take a rest. But they don't have the finest idea how that will effect their children in the future. According to some investigations, 75 percent of the children are becoming asocial day by day owing to the lack of communication. After a while, parents observe their children and realize that they don't talk as much as before. A social life requires communicating. And because of the lack of the communication, many young people don't have a social life that they can share something with other people.

. As explained above, we can see how overusing the internet can effect our lifes. Addiction, insomnia and associality are just three items that effect our living styles in negative ways. So, I'd recommend everyone to be aware of the other beautiful things in the world, not to waste their time in vain and to enjoy their lifes.

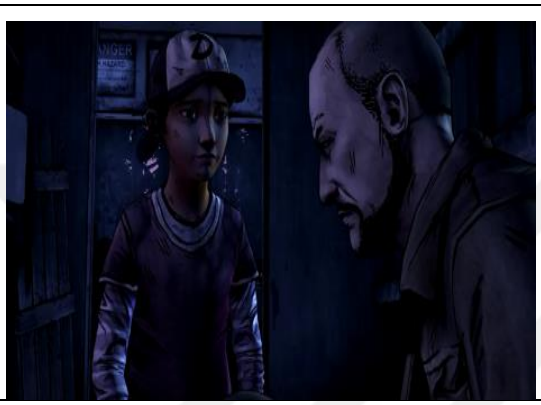
APPENDIX E. PICTURES FOR NARRATIVE WRITING TASKS

E. 1. Picture for Simple Narrative Writing Task

It was taken from https://www.youtube.com/watch?v=qz3zmegJzO8 through snapshot

E. 2. Pictures for Complex Narrative Writing Task

They were taken from https://www.youtube.com/watch?v=qz3zmegJzO8 through snapshot




lead cinematic artist
REBEKAH CAMIN AFGOVITCH


Someone came here looking for you. A man.

APPENDIX F. CONSENT FORM FOR PARTICIPANTS

KATILIMCI ONAY FORMU

Atatürk Üniversitesi, Eğitim Bilimleri Enstitüsü, İngilizce Öğretmenliği Bilim Dalı'nda doktora yapmaktayım. Doktora tezi çalışmam, işlem karmaşıklığının İngilizce Öğretmenliği Bölümü öğrencilerinin yazma çalışmaları üzerine sözcüksel ve tümcesel karmaşıklık, genel yazım kalitesi açısından bir etkisi olup olmadığını araştırmayı amaçlamaktadır. Bu amaç doğrultusunda, 2015-2016 Güz yarıyılında yürütmüş olduğum İleri Okuma ve Yazma I dersi esnasında sizler tarafından hem ders içerisinde hem ders sonrasında ödev olarak üretilen yazma çalışmalarının bir kısmını, tümcesel karmaşıklık ve tutarlılık, bütünlük ve içerik olmak üzere genel yazım kalitesi açısından incelemek amacıyla kullanmak istiyorum. Hiçbir kişisel ve akademik bilginiz, kesinlikle, araştırmanın herhangi bir kısmında, başka bir araştırmacı ya da farklı bir araştırmada kullanılmayacaktır.

Arş. Gör. Mine YILDIZ

Katılımcı Onayı

Yukarıda yer alan bilgileri okudum ve katılmam istenen çalışmanın kapsamını ve amacını, gönüllü olarak üzerime düşen sorumlulukları anladım. Çalışma hakkında yazılı ve sözlü açıklama aşağıda adı belirtilen araştırmacı tarafından yapıldı. Kişisel bilgilerimin özenle korunacağı konusunda yeterli güven verildi.

Bu koşullarda söz konusu araştırmaya kendi isteğimle, hiçbir baskı olmaksızın katılmayı kabul ediyorum. Yukarıda belirtildiği gibi tarafımdan üretilen çalışmaların, bilgilerim paylaşılmadan kullanılmasında hiçbir sakınca yoktur.

Katılımcının
Adı-Soyadı: Tarih:
İmzası:

Araştırmacının

Adı-Soyadı: Tarih:
İmzası:

APPENDIX G. DOCUMENTS FOR ETHICS COMITTEE APPROVAL

G. 1. Application of the researcher

KAZIM KARABEKİR EĞİTİM FAKÜLTESİ
YABANCI DİLLER EĞİTİMİ BÖLÜMÜ
İNGİLİZ DİLİ EĞİTİMİ ANABİLİM DALI BAŞKANLIĞINA

Anabilim dalımızda, Yrd. Doç. Dr. Savaş YEŞİLYURT'un danışmanlığında "Görev zorluğunun öğrencilerin ikinci dilde yazma çalışmalarına karmaşıklık düzeyi ve yazım kalitesi açısından etkisi" konulu doktora tez çalışmamı yapmaktayım. Bu çalışmam için 2015-2016 Güz Yarıyılında ders hocası Okt. Yılmaz YAZICI gözetiminde yürüttüğüm (103-AL) İleri Okuma ve Yazma I dersinde, ders kapsamı içerisinde öğrenciler tarafından üretilen hem dersiçi hem de ödev olarak istenilen yazma çalışmalarının bir kısmını, ekteki listede belirtilen öğrencilerin iznini de alarak kullanmak istiyorum. Bu konuda, anabilim dalımızın yazılı iznini talep etmekteyim.

Gereğini ve bilgilerinizi arz ederim. 17.05.2016

Doktora Öğrencisi

Mine YILDIZ

Adres: Atatürk Üniversitesi Kazım Karabekir Eğitim Fakültesi

İngiliz Dili Eğitimi Anabilim Dalı No:416 Yakutiye/Erzurum

Tel: 0537 954 57 92

442 231 42 02

G.2. Application of the advisor

KAZIM KARABEKİR EĞİTİM FAKÜLTESİ
YABANCI DİLLER EĞİTİMİ BÖLÜMÜ
İNGİLİZ DİLİ EĞİTİMİ ANABİLİM DALI BAŞKANLIĞINA

Doktara tez danışmanlığını yapmakta olduğum Mine Yıldız'ın "Görev zorluğunun öğrencilerin ikinci dilde yazma çalışmalarına karmaşıklık düzeyi ve yazım kalitesi açısından etkisi" konulu doktora tez çalışması için, 2015-2016 Güz Yarıyılında kendisi tarafından ders hocası Okt. Yılmaz YAZICI gözetiminde yürütmüş olduğu (103-AL) İleri Okuma ve Yazma I dersinde öğrenciler tarafından üretilen hem dersiçi hem de ödev olarak istenilen yazma çalışmalarının bir kısmını, ekteki listede belirtilen öğrencilerin iznini de alarak kullanmak istiyoruz. Bu konuda, anabilim dalımızın yazılı iznini talep etmekteyiz.

Gereğini ve bilgilerinizi arz ederim. 17.05.2016

Danışman

Yrd. Doç. Dr. Savaş YEŞİLYURT

Adres: Atatürk Üniversitesi Kazım Karabekir Eğitim Fakültesi

İngiliz Dili Eğitimi Anabilim Dalı No:418 Yakutiye/Erzurum

Tel: 0505 224 23 46

G.3. Response of Ethics Committee

T.C.
ATATÜRK ÜNİVERSİTESİ REKTÖRLÜĞÜ
Sosyal ve Beşeri Bilimler Etik Kurul Başkanlığı

Sayı : 75513667-020-E.1600129407 31.05.2016
Konu : Y.Doç.Dr. Savaş YEŞİLYURT-
Arş.Gör. Mine YILDIZ Etik
Kurul Onay Belgesi

DAĞITIM YERLERİNE

İlgi : 26.05.2016 tarihli ve 77040475-000-E.1600125633 sayılı belge.

Üniversitemiz Kazım Karabekir Eğitim Fakültesi Yabancı Diller Eğitimi Anabilim Dalı İngilizce Eğitimi Bilim Dalı öğretim üyelerinden Yrd.Doç.Dr. Savaş YEŞİLYURT'un danışmanlığını yürütmüş olduğu doktora öğrencisi **Arş.Gör. Mine YILDIZ**'ın yapmakta olduğu ***"Görev Zorluğunun Öğrencilerin İkinci Dilde Yazma Çalışmalarına Karmaşıklık Düzeyi ve Yazım Kalitesi Açısından Etkisi"*** konulu doktora tez çalışmasında öğrenciler tarafından *103 AL İleri Okuma Yazma I* dersinde üretilen yazma çalışmalarının bir kısmı söz konusu çalışmada kullanılması için Katılımcı Onay Formu, Katılımcı Listesi ve dilekçeler ekte sunulmuş ve katılımcı listesindeki tüm öğrencilerin yazılı izinleri alınmış olup, söz konusu çalışmanın Etik Kurul tarafından incelenerek Uygunluk-Onay belgesi talebi ile ilgili, ***Sosyal ve Beşeri Bilimler Etik Kurulu'nun 30.05.2016*** gün ve ***10.13*** sayılı kararı aşağıya çıkarılmıştır.

Gereğini bilgilerinize arz ederim.

Karar - 13:

Üniversitemiz Kazım Karabekir Eğitim Fakültesi Yabancı Diller Eğitimi Anabilim Dalı İngilizce Eğitimi Bilim Dalı öğretim üyelerinden Yrd.Doç.Dr. Savaş YEŞİLYURT'un danışmanlığını yürütmüş olduğu doktora öğrencisi **Arş.Gör. Mine YILDIZ**'ın yapmakta olduğu ***"Görev Zorluğunun Öğrencilerin İkinci Dilde Yazma Çalışmalarına Karmaşıklık Düzeyi ve Yazım Kalitesi Açısından Etkisi"*** konulu doktora tez çalışmasında öğrenciler tarafından *103 AL İleri Okuma Yazma I* dersinde üretilen yazma çalışmalarının bir kısmı söz konusu çalışmada kullanılması için Katılımcı Onay Formu, Katılımcı Listesi ve dilekçeler ekte sunulmuş ve katılımcı listesindeki tüm öğrencilerin yazılı izinleri alınmış olup, söz konusu çalışmanın Etik Kurul tarafından incelenerek Uygunluk-Onay belgesi talebi ile ilgili husus görüşüldü.

Yapılan görüşmelerden sonra; adıgeçenin ***"Görev Zorluğunun Öğrencilerin İkinci Dilde Yazma Çalışmalarına Karmaşıklık Düzeyi ve Yazım Kalitesi Açısından Etkisi"***

Atatürk Üniversitesi Merkez Kampüsü 25240 Yakutiye/ERZURUM Bilgi: Süleyman ÇİĞDEM
Elektronik Ağ: http://www.atauni.edu.tr

Kep Adresi: atauni@hs01.kep.tr

başlıklı tezle ilgili yapılacak araştırma için, araştırmaların gerekçe, amaç, yaklaşım ve yöntemleri dikkate alınarak konuyla ilgili çalışmaların gerçekleştirilmesinde etik ve bilimsel yönden sakınca bulunmadığına,

oy birliği ile karar vermiştir.

Prof.Dr. Süleyman ÇİĞDEM
Kurul Başkanı

Ek :
1 - Etik Kurul Kararı
2 - Yazışma Evrakları

Dağıtım:
Gereği:
Sayın Yrd.Doç.Dr. Savaş YEŞİLYURT
Hukuk Müşavirliğine

Atatürk Üniversitesi Merkez Kampüsü 25240 Yakutiye/ERZURUM
Elektronik Ağ: http://www.atauni.edu.tr
Bilgi: Süleyman ÇİĞDEM
Kep Adresi: atauni@hs01.kep.tr

CURRICULUM VITAE

Personal Information

Name Surname: Mine YILDIZ

Place of Birth / Date of Birth: Erzurum / 08.07.1988

Education

Primary Education: Mehmetçik Primary School / Erzurum – 2002

Secondary Education: Erzurum Central Anatolian High School / Erzurum -2006

University: Atatürk University, KKEF, ELT Department /Erzurum – 2011

M. A.: Atatürk University, Institute of Educational Sciences, FLE Department – 2013

PhD: Atatürk University, Institute of Educational Sciences, FLE Department – 2017

Professional Qualifications

Language: English, French

Employment History

2012- …. Atatürk University – Research Assistant

Contact Information

Adress: Department of English Language Teaching

Kazım Karabekir Faculty of Education

Atatürk University / Erzurum

e-mail: yldizmine@gmail.com