THE EFFECTS OF PLANNING TIME ON ENGLISH LANGUAGE LEARNERS' COMPLEXITY, ACCURACY, AND FLUENCY IN WRITING

İREM TÜMER GÖRGÜLÜ

BOĞAZİÇİ UNIVERSITY

2019

THE EFFECTS OF PLANNING TIME ON ENGLISH LANGUAGE LEARNERS' COMPLEXITY, ACCURACY, AND FLUENCY IN WRITING

Thesis submitted to the

Institute for Graduate Studies in Social Sciences

in partial fulfillment of the requirements for the degree of

Master of Arts

in

English Language Education

by

İrem Tümer Görgülü

Boğaziçi University

2019

The Effects of Planning Time on English Language Learners'

Complexity, Accuracy, and Fluency in Writing

The thesis of İrem Tümer Görgülü

has been approved by:

Assist. Prof. Şebnem Yalçın
(Thesis Advisor)

Assist. Prof. Sibel Tatar

Assist. Prof. Derya Altınmakas
(External Member)

November 2019

DECLARATION OF ORIGINALITY

I, İrem Tümer Görgülü, certify that

- I am the sole author of this thesis and that I have fully acknowledged and documented in my thesis all sources of ideas and words, including digital resources, which have been produced or published by another person or institution;
- this thesis contains no material that has been submitted or accepted for a degree or diploma in any other educational institution;
- this is a true copy of the thesis approved by my advisor and thesis committee at Boğaziçi University, including final revisions required by them.

Signature 

Date 06.11.2019

ABSTRACT

The Effects of Planning Time on English Language Learners'

Complexity, Accuracy, and Fluency in Writing

The aim of study is to find out if elementary level English as a foreign language students` written accuracy, fluency, and syntactic complexity levels are affected by planning condition of a task. Secondarily, the relationship between writing anxiety and the three language dimensions is examined. The results of study are reported by following two influential task complexity hypotheses: the Cognition and the Limited Attentional Capacity. The former claims that increasing complexity level of a task leads to an increase in accuracy and complexity thanks to multiple attentional capacity of human brain. The latter supports limited attentional capacity models and states that paying simultaneous attention to both accuracy and syntactic complexity is not probable. In the study, twenty university students studying English produced two different narratives based on two sets of pictures under two different task conditions: pre-task planning (PTP) and no pre-task planning (NP). Accuracy of written language production was measured by ratio of errors per T-unit and errors per 100 words. On the other hand, written fluency was measured by counting words and syllables per minute. Syntactic complexity was measured by mean length of T-units and clauses per T-units. Moreover, writing anxiety was measured by Second Language Writing Apprehension Inventory. As a result of the analysis, it was found out that students produced more error free narratives under simpler task condition (PTP) but the planning type did not impact either fluency or syntactic complexity levels. Regarding writing anxiety, correlation was only found with accuracy.

ÖZET

Planlama Zamanının İngilizce Öğrenenlerin Yazmadaki Sözdizimsel Karmaşıklık, Hatasızlık ve Akıcılık Üzerindeki Etkileri

Bu çalışmanın amacı, başlangıç seviyesinde İngilizce öğrenen öğrencilerin sözdizimsel karmaşıklık, hatasızlık ve akıcılık düzeyleri bakımından yazılı üretiminin dil öğrenme ödevinin planlama koşulundan ne şekilde etkilendiğini bulmaktır. İkincil amaç ise yazılı dil ölçütleri ile yazma kaygısı arasındaki korelasyonu bulmaktır. Çalışmanın sonuçları, iki etkili bilişsel ödev zorluğu hipotezini izleyerek rapor edilmiştir: Biliş Modeli ve Sınırlı Dikkat Kapasitesi Modeli. Modellerden ilki, insanlardaki çoklu dikkat kaynakları sayesinde bir ödevin bilişsel karmaşıklık seviyesinin artmasının, doğruluk ve karmaşıklığının artmasına yol açtığını iddia ediyor. İkincisi, sınırlı dikkat kapasite modellerini desteklemektedir ve hem doğruluk hem de sözdizimsel karmaşıklığa eşzamanlı dikkat göstermenin muhtemel olmadığını belirtmektedir. Çalışmada bilişsel zorluk ödev öncesi planlama koşulu yönünden değiştirilmiştir. Çalışmaya üniversitede İngilizce öğrenen yirmi öğrenci katılmış ve bu katılımcılardan kendilerine verilen resimlerden yola çıkarak iki adet farklı hikaye yazmaları istenmiştir. Fakat bunları iki farklı görev koşulu -görev öncesi planlama (PTP) ve görev öncesi planlama (NP) - altında yapmışlardır. Yapılan analiz sonucunda öğrencilerin daha basit ödev koşulunde (PTP) daha hatasız anlatılar ürettikleri, ancak planlama türünün akıcılık veya sözdizimsel karmaşıklık seviyelerini etkilemediği tespit edilmiştir. Yazma kaygısı ile yazınsal ölçütlerden ancak biri ile korelasyon bulunmuştur. O da hatasızlıktır.

ACKNOWLEDGEMENTS

First, I would like to express my heartfelt gratitude to my thesis advisor Assist. Prof Şebnem Yalçın for her never-ending encouragement and guidance. Her inspiring comments have enlightened me during the thesis process. More than a thesis advisor, she was also my psychological supporter. Whenever I lost my hope in me, she was always there with her encouraging words.

I would like to offer my further thanks to Aysen Tuzcu for her never-ending sincere support and sharing her insights about my thesis subject.

I would also like to thank Fazilet Sönmez for her invaluable help and patience during the data coding process.

My deepest thanks and gratitude go to my dear friend and colleague Duygu Çınga for her precious support in data collection. I also thank my friends and my parents who have always believed in me and supported me. I am really grateful to them for bearing with me through my thesis process.

Above all, I would like to thank my husband, Çetin Görgülü, for his unconditional love, endless support and never-ending encouragement.

My special thanks go to my mother-in-law and father-in-law, Nazife Görgülü and Alpaslan Görgülü for looking after my little son with patience while I was writing my thesis. If it had not been their support, I would not have been able to finish this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Tasks are motivating communication activities which offer learners a reason to use language for a real purpose. They provide learners with meaningful and authentic opportunities to improve their interlanguages (Norris, 2009, pg. 580). Research on task-based language teaching (TBLT) pedagogy has focused on investigating the role of tasks in language learning since the late 1980s. In TBLT, syllabus design should be based on pedagogic tasks in which the communication of meaning is fundamental with a learning goal and evaluation of achievement is based on the outcome of an authentic activity (Skehan, 1998a). Tasks should be sequenced according to some well-chosen criteria so that they can promote an effective balance between fluency and accuracy while restructuring the previously learned language (Skehan, 1996a). To achieve this goal, TBLT syllabus designers benefit from two influential frameworks for task selection and sequencing. These frameworks are Triadic Componential Framework (Robinson, 2001a & b) and Skehan`s (1996a) three-way distinction for task sequencing. Both of them comprise a task-complexity rationale to sequence tasks from the least to the most complex. They base their propositions about task complexity on cognitive models of attention and attempt to shed light on how task complexity influence the quality of the learners` second language (L2) performance.

Rather than being a unitary construct, L2 performance is thought to comprise multiple components which are named as complexity, accuracy, and fluency (CAF)

(Housen & Kuiken, 2009; Housen, Kuiken & Vedder, 2012). Improving these three language areas is the goal of the TBLT as stated by Skehan (1996a). Therefore, they are the most widely used measures to assess the linguistic outcomes of a task in the literature. The first one, complexity, is defined as the size, richness, and diversity of the learners` L2 linguistic systems (Housen & Kuiken, 2009). However, one must be cautious while interpreting the results of complexity measures as higher complexity does not necessarily mean higher performance but it can be a stylistic choice (Palotti, 2009). The second one, accuracy, is `how well the target language is produced in relation to the rule system of target language` (Skehan, 1996b, pg. 23). However, a well-developed interlanguage does not necessarily mean more accurate production. Learners who do not like risk-taking will be more accurate ones (Skehan, 1996a). Finally, the last goal, fluency, is explained as the words that can be accessed by the learner in a limited time (Wolfe-Quintero, Inagaki, & Kim, 1998). Although these language aspects are distinct, they interact with each other in unpredictable ways (Norris & Ortega, 2009, pg. 2).

Two influential hypotheses in task complexity research try to predict how three dimensions of CAF interact with each other in relation to the complexity of tasks. In this account, Skehan proposes Limited Attentional Capacity Model (Skehan, 1998; Skehan & Foster, 1999) whereas Robinson posits Cognition Hypothesis (Robinson, 2001a, 2001b, 2003, 2005, 2007). They have distinct propositions about the effects of attentional capacity on language production. Limited Attentional Capacity Model proposes that learners have limited attention capacity for different language areas (CAF) that might compete when performing a cognitively complex task (Skehan, 1998; Skehan & Foster, 1999). On the other hand, Cognition Hypothesis proposes that a learner has multiple

attentional resources so his/her CAF levels do not necessarily have to compete during the language production (Robinson, 2001a, 2001b, 2003, 2005, 2007).

Although there is extensive research on task complexity and speaking, the studies investigating task complexity, especially in terms of planning, in written production is rather limited. Therefore, the main purpose of the present study is to explore the effect of task complexity on elementary level English as a foreign language (EFL) learners` written narrative performance when one dimension of task complexity (i.e. planning time) is manipulated. 45 undergraduate students with beginner and elementary level proficiency participated in the study and wrote two different narratives under two different planning conditions. Their written L2 production was evaluated by using CAF framework.

Furthermore, writing anxiety is another independent variable which was investigated in relation to CAF in the current study. In Triadic Componential Framework of Task Complexity, it is classified as one of the learner factors which helps investigating performance differences between language learners. However, there are very few studies scrutinizing the relationship between writing anxiety and CAF in TBLT research. Thus, the secondary purpose of the study is to scrutinize how writing anxiety affects learners` written output quality.

The study investigated the following research questions:

1. How does manipulating task complexity along +/- pre-task planning time affect the syntactic complexity, accuracy, and fluency levels of EFL learners` written performance?
2. Is there any relationship between WA levels and CAF levels of EFL learners under no-planning condition?

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

First of all, this chapter aims to describe the cognitive processes of writing by reviewing an influential cognitive writing hypothesis. Secondly, it examines the constructs that constitute second language performance, namely, complexity, accuracy and fluency. The following part will focus on two prominent hypotheses explaining the effects of task complexity on L2 performance; Skehan`s (1998b) Limited Attentional Capacity Hypothesis and Robinson`s (2001a) Cognition Hypothesis. Next, the role of individual differences in L2 research will be overviewed and the role of writing anxiety in L2 acquisition will be focused on. Finally, it will present an overview of task complexity research manipulating planning time. Also, previous studies on writing apprehension in relation to CAF will be summarized in the last section.

2.2 The role of planning in written language production

Planning is one of the several processes involved in written language production. Investigating the role of planning in written language production cannot be complete without considering other cognitive processes in relation to it. Cognitive writing models which do not see writing process as a step by step action but an interactive action done by the writer aim to explain the relationship between cognitive processes of writing. Although there are several writing models of L1 (e.g. Bereiter & Scardamalia, 1987; Flower & Hayes, 1981a; Kellogg, 1996; Grabe & Kaplan, 1996; Zimmerman, 2000),

there is not any globally accepted L2 writing model (Cumming,2001). Among these writing models, Flower and Hayes (1980a&b, 1981a), Hayes (1996) and Kellogg (1996) have explained the process of writing cognitively and emphasized the role of attentional capacity in the writing process. Similarly, attentional capacity also plays a central role in task-complexity research. Thus, in this literature review only Kellogg`s writing model will be mentioned in relation to the processes of writing as it draws heavily on previous models (Flower and Hayes, 1980a&b,1981a). Furthermore, types of planning will be explained in relation to writing models and the task-based language teaching.

Kellogg`s model (1996), as shown in Figure 1, distinguishes three recursive and interactive systems in text production, namely formulation, execution, and monitoring. Each of these systems consists of two basic level processes. Formulation involves planning and translation of ideas into the text whereas execution comprises creation of the text with handwriting or typing. Monitoring stage involves reading and editing of the text produced. In this model, planning stage is conceptualized as a process when the writer establishes the writing purposes, comes up with ideas related to them, and organizes these related ideas. In other words, the writer thinks about the content of the writing in the planning stage.

This model draws on Baddeley`s WM model (1986) and puts great emphasis on working memory. It relates all the processes of the writing with the components of WM. All major processes except execution place major demands on central executive system of working memory. The main proposition of Kellogg`s model is that central executive has limited capacity so writers may have to decide on which writing process to prioritize when they are under time pressure. However, according to Brown, et al. (1988) study,

formulation demands always take the priority over execution and monitoring as it is the most crucial process in writing.

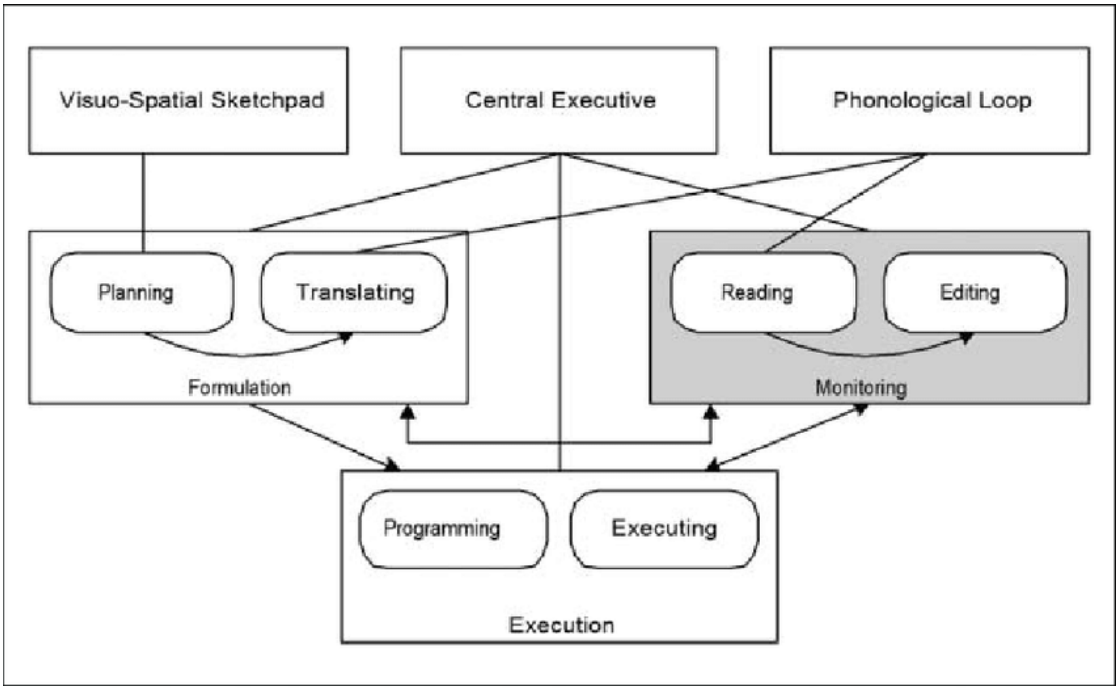


Figure 1. Writing Model of Kellog (1996)

Although designed to account for L1 writing, L2 writing research can also draw on Kellog's (1996) because it explains cognitive processes which are not specific to any language. However, it is important to scrutinize in what ways L2 writing differs from L1 writing. One difference can be the limited proficiency. De Larios, Martin, and Murphy. (2001) found that low level L2 writers need to concentrate on translation of sentences at the expense of planning and revising. That is, they tend to pay more attention to formulation stage. To explain this, they proposed that central executive hinders some cognitive processes (i.e. planning) to guarantee the completion of a task.

Planning seems like one intact cognitive process but there are different types of it depending on the time when it occurs and how it occurs. In their review about planning, Hayes & Nash (1996) present taxonomy of planning in their L1 writing model. They distinguish between process planning and text planning. The focus of the process planning stage is the writer and how the task to be performed. Text planning is divided into two: abstract planning and language planning. Abstract text planning entails planning of ideas and concepts to be written whereas process planning entails planning a string of words, or one or two clauses, and then writing them down verbatim. Also, Hayes & Nash (1996) posits that planning and text production can happen simultaneously during the process of writing. They also distinguished planning from other reflective processes in that it may happen in an environment other than the task environment. That is, it can occur even before a writer starts the production process. Accordingly, Ellis and Yuan (2004) propose that whether planning occurs inside or outside the task environment can be a determining factor for the planning type in TBLT. As a result, Ellis (2005) suggests the main classification of planning by using the premises of the planning taxonomy mentioned above. He suggests two major types of task planning conditions: pre-task planning and within-task planning (pg. 3-4). If the planning takes place before the task, it is called pre-task planning (PTP). On the other hand, if the planning process takes place during the performance of the task, it is named with-in task planning or online planning (OP). There are two types of pre-task planning: rehearsal, in which the planning time is used as an opportunity to perform the complete task once before performing the real task, and strategic planning, in which the planning time is used to think about the content and what language to use but without rehearsing the complete task (Ellis, 2005, pg.3).

Strategic planning can also be implemented in a number of ways: unguided, guided with content focus, guided with language focus (Ellis, 2018, pg. 241). OP can be implemented in two different ways by giving limited time, that is pressured, or by not giving any specific time, that is unpressured, to finish the task (see Figure 2).

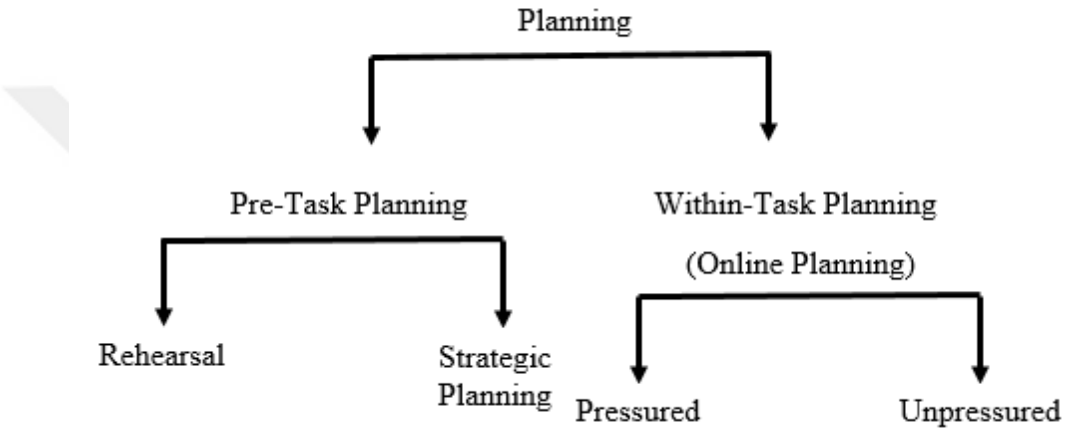


Figure 2. Types of planning
Source: [Ellis, 2005]

For the effect of giving pre-task planning time, Ellis (2005) states that PTP lessens the burden on learners` working memory which allows learners to attend to form while they are conveying the message. To put it simply, pre-task planning creates a context in which learners are provided with the opportunity to map form onto meaning through available linguistic knowledge that is not yet automatized (Ellis, 2005, pg. 4).

In a nutshell, since writing entails a complex interaction between a wide range of different cognitive processes, these processes place extremely high demands on the limited capacity of working memory especially when language

production is not automated. Therefore, it is highly possible that providing PTP time to L2 learners will ease that burden on WM, and will give them the opportunity to focus on form. As a result, more accurate and complex language production can be expected from L2 writers.

2.3 Assessment of language performance: Complexity, accuracy, and fluency

Many researchers in Second Language Acquisition (SLA) emphasized the multi-componential nature of L2 performance and proficiency. As such, they purported that the notions of complexity, accuracy, and fluency can comprehensively capture this nature to investigate learners' language performance deeply (Skehan, 1996b; Robinson, 2001a; Ellis & Barhuizen, 2005; Wolfe-Quintero et al. 1998). Therefore, complexity, accuracy, and fluency (CAF) have been appeared as major research variables in SLA research, and they have been used both as performance descriptors for the oral and written assessment of language learners as well as indicators of learners' proficiency underlying their performance. For task-based SLA studies, the first CAF framework which has distinguished three aspects of L2 linguistic performance was introduced by Skehan (1996a) as language learning goals of task-based language teaching. Furthermore, Skehan (2009) added lexical variety to this triad, and it has been used in a number of studies since then.

Complexity is the most complicated and less straightforward term to define and measure among those linguistic aspects. Housen et al. (2012) outline complexity as "the ability to use a wide and varied range of sophisticated structures and vocabulary in the L2" (p.2). According to this definition, complexity of production has two components:

syntactic and lexical. Syntactic complexity in second language acquisition was defined by (Ortega, 2003) as the following: "range of forms that surface in language production and sophistication of such forms". Palotti (2015) defines a text`s structural complexity as the number of different linguistic elements and the interconnections between them, which in return result in a longer text production. Therefore, length of production unit, amount of subordination, range of structural types, and sophistication of the particular structures should play crucial roles as a unit of measurement while measuring language complexity (Ortega, 2003). Furthermore, Norris and Ortega (2009) identify three sub-constructs of syntactic complexity: (1) complexity by subordination, which is measured by diving the number of multi-clausal units with subordinated clauses; (2) general complexity, which is measured by any length based unit such as words, clauses divided by multiple-clausal unit; (3) sub-clausal complexity by phrases which is measured by mean length of a phrase. They add two more sub-components which have been underutilized in SLA studies so far: coordination and sophistication of forms. However, Palotti (2015) states that defining some syntactic structures as more complex than others is problematic. Hence, researchers should focus on stylistic syntactic complexity which can be measured by length of phrase, number of phases per clause, and number of clauses per unit. Lexical diversity is also utilized as a measure of complexity because some learners may not use a large variety of grammatical structures but a large variety of words. Therefore, it can give a more comprehensive picture of learner language when it is analyzed in terms of complexity (Ellis & Barkhuizen, 2005). It is generally measured by type-token ratio which is calculated by dividing the number of word types to total number of words in a text.

Accuracy is probably the most transparent and most consistent construct of the triad, referring to the degree of deviancy from a particular norm (Wolfe-Quintero et al. 1998). Deviations from the norm are usually characterized as errors. Although this definition seems straightforward, it raises the thorny issue of criteria for evaluating accuracy and identifying errors. Whether the accuracy criteria should follow prescriptive or descriptive grammar rules (Housen & Kuiken, 2009) is another controversy. Accuracy is measured via various types of measures, some of which focus on the learners` accurate performances (e.g. the ratio of error-free units or clauses) whereas some of which focus on errors in the performance (e.g. the mean number of errors per units or 100 words). Also, a number of language specific measures can be used such as, correct use of articles or verb forms.

Fluency is related to the speed and ease of accessing relevant information from L2 interlanguage in order to meaningfully communicate in real time, so it is basically about learners` control over their interlanguage (Housen & Kuiken, 2009). It has been defined as a multi-componential construct by Tavakoli and Skehan (2005). Its sub-components include (1) speed or rate e.g. number of words per minute; (2) silence or breakdown e.g. amount, location, and duration of pauses; (3) repair e.g. false starts, repetitions. Although it was first used as a speaking performance measure in SLA, more and more written production studies have been utilizing it. For SLA writing research, fluency is a very controversial construct as it is hard to measure it truly due to the recursive nature of writing process (Wolfe-Quintero, et al., 1998). Many written production studies utilize measures of rate such as number of words or syllables produced per minute (Ong &

Zhang, 2010), and measures of length such as average length of sentence (Johnson, et al.2012).

Still, answers of many questions remain unclear about the operationalization of those constructs while assessing the relation of language output with other variables such as learners` proficiency, task complexity, and interlanguage development (Michel, 2017). In this study, accuracy, fluency and complexity dimensions of written language production will be analyzed as dependent variables. The independent variable to investigate in relation to CAF will be the pre-task planning time and writing anxiety.

2.4 Task complexity in TBLT

"Task is the unit of analysis throughout the design, implementation, and evaluation of a TBLT program" (Long, 2015; pg. 6). Identifying criteria for grading and sequencing tasks in terms of their difficulty is an essential part of syllabus design in TBLT (Ellis, 2018; pg.12). According to Robinson (2001b), task complexity occurs as a result of attentional, memory, reasoning, and other information processing demands caused by the structure of the task on the language learner. In TBLT research context, the word `complexity` does not refer to the difficulty of linguistic form used in a pedagogic task but rather refers to how difficult it is to execute a pedagogic task. Criteria for task complexity have been proposed by different scholars about TBLT with the aim of providing a principled basis for grading tasks. However, there is still no globally accepted task grading and sequencing criteria (Baralt, Gilabert, Robinson, 2016, pg. 1).

2.4.1 Early approaches to task complexity

Scholars have tried to classify tasks as `simple` and `difficult` and have proposed generalizable frameworks for task grading since the early 1980s. One of the first attempts to design task grading framework was made by Brown, Anderson and Yule (1984). Brown et al. (1984) conducted a series of speaking studies which manipulated the task difficulty in terms of type of information contained and +/- few elements. According to results of these studies, they proposed a matrix of difficulty with two dimensions. First dimension is information type which produces three distinct task types, namely, static, dynamic, and abstract. Static tasks are regarded as the simplest one because the information presented is the same through the task implementation. Describing a diagram could be an example of this type. Whereas in dynamic tasks, the information presented by task can be changed during the task performance. For instance, narrating a car crash involves both a sequence of events and causality, so they are more challenging than static ones. The most difficult type is abstract tasks in which an idea is presented and requires learners to present their opinions about it. Second dimension is about the number of elements involved and the relationship between them. That is, the more elements involve the more difficult a task is. Similarly, the more intricate the relationship between elements, the more difficult a task is.

Another pedagogic attempt for task grading comes from Prabhu (1987) from the Bangalore Project. The project aimed to teach English in India by creating meaningful conditions for communication. In that sense, it was an early attempt to design a TBLT syllabus based on classroom experience. In the project, teachers present learners a pre-task to demonstrate the main task and ask them to assesses the main task`s difficulty, and

then teacher adapt the task according to students` needs. However, the tasks are not structured according to a grammar point, instead they focus on the function of language and "let the language relevant to task to come into play" (Prabhu, 1987; pg. 276). Explicit error correction and feedback are deliberately avoided and language learning is expected to be incidental. Prabhu suggested the use of challenging classroom activities which engaged learners in reasoning, inferring or problem solving. He stated that their "demand on thinking should be just above the level which learners could meet without help" (p. 24). Based on that criterion, tasks were divided into three types: information gap, reasoning gap, and opinion gap activities (pg.47). Based on thinking demand of a task, he proposed that opinion-gap tasks were the most difficult one due to their unpredictability. Whereas reasoning gap activities were more challenging than info-gap activities but less complex than opinion-gap activities. As an outcome of the project, Prabhu (1987) concluded that reasoning-gap tasks were the most successful tasks as for producing useful language as well as being reasonably challenging for students (pg. 47).

Different from previous researchers, Candlin (1987, as cited in Skehan,1998) established a more extensive criteria for sequencing and grading tasks. Candlin (1987) proposed six criteria to select and grade tasks. These criteria are named as cognitive load (general cognitive complexity of a task), communicative stress (i.e. high proficiency of interlocutor), particularity and generalizability (having a clear goal for the task), code complexity and interpretative density (the linguistic and argumentative complexity of texts used in tasks), process continuity (learners` capacity to relate tasks). Later, these criteria have influenced Skehan`s (1998) three-way distinction framework for task sequencing.

Similar to Prabhu, Long & Crookes (1992) also advocated designing a task based syllabus according to needs analysis. Once the needs of learners are identified, target tasks such as serving lunch for a course designed for flight attendants are defined. After that, task types are identified in order to design the task-based syllabus. Grading of tasks from simple to complex should not be based on linguistic elements but rather be based on certain characteristics of tasks. These characteristics are the number of steps involved, the number of solutions to a problem, the number of parties involved and the saliency of their distinguishing features, the location (or not) of the task in displaced time and space, the amount and kind of language required, the number of sources competing for attention, and other aspects of the intellectual challenge. Therefore, they advocate that tasks should be sequenced according to their cognitive complexity levels since this sequence could help learners build on their existing knowledge and achieve the target tasks in L2 (Prabhu, et al., 1992, pg. 45).

These previous attempts to categorize task as complex or simple have provided a basis for the development of two most influential frameworks and hypotheses for task sequencing and grading in TBLT. These are Skehan`s (1998) three-way distinction framework and the Limited Attentional Capacity Hypothesis and Robinson`s (2001a) Triadic Componential Framework and the Cognition Hypothesis. Both models attempt to categorize tasks and explain the effects of task complexity on linguistic outcomes of tasks. As they have different theoretical foundations, they predict the effect of task complexity on learners` performances and L2 learning differently. In the next two sections, Skehan`s and Robinson`s models will be reviewed.

2.4.2 Skehan's task sequencing framework and Limited Attentional Capacity Hypothesis

According to Skehan (1992; 1996; 1998a) and Skehan & Foster (2001) the cognitive complexity of tasks can be manipulated in various dimensions. Following the task categorization premises of Candlin (1987), Skehan (1998a) categorizes the task dimensions affecting task complexity into three areas: Code complexity, cognitive complexity, and communicative stress (see Table 1).

Table 1. Skehan`s Three-Way Distinction for Task Sequencing

Code complexity	Cognitive complexity	Communicative stress
– linguistic complexity and variety – vocabulary load and variety	– cognitive familiarity – familiarity of topic and its predictability – familiarity of discourse genre – familiarity of task – cognitive processing – information organization – amount of 'computation' – clarity of information given	– time pressure – scale – number of participants – length of texts used – modality – stakes – opportunity for control

Source: [Skehan, 1998b]

Skehan (1996, 1998a&b, 2009, 2014) proposes the Trade-off Hypothesis (Limited Attentional Capacity Hypothesis) to predict the effects of manipulating tasks along these three dimensions on language production. He bases his hypothesis on two assumptions. The first one is that humans have limited attentional resources and working memory capacities so they cannot pay concurrent attention to content and form during execution of language learning tasks. He supports his hypothesis with a study conducted by VanPatten (1990). In the study, a group of Spanish L2 learners were randomly assigned to four listening task conditions. All task groups including the control group were instructed to listen to a text for its content, whereas three experimental groups were also asked to pay attention to either a particular lexical item or a grammar point or a verb form

along with paying attention to content. The results suggested that learners experience great difficulty in directing their attention to both form and content simultaneously because attending to grammar resulted in less comprehension. Drawing on VanPatten‟s (1990) study, Skehan (1998) proposes that while dealing with a cognitively complex task, L2 learners tend to allocate most of their attentional resources to the meaning since these tasks require extensive communication for meaning. As a result, they will be paying less attention to form of the language they use during the task. In other words, while performing a task there is a constant competition between meaning for "getting a task done" and form leading to "language focus and development" (Skehan & Foster, 2001, p. 190). To explain this competition between meaning and form of language in terms of complexity, accuracy, and fluency dimensions of language production, Skehan & Foster (2001) present dimensions of language performance like in Figure 3 (pg. 190). The Figure 3 shows that a commitment to carrying out the meaning requires more fluent language use whereas a commitment to attend the form of language requires more complex and accurate use of language forms.

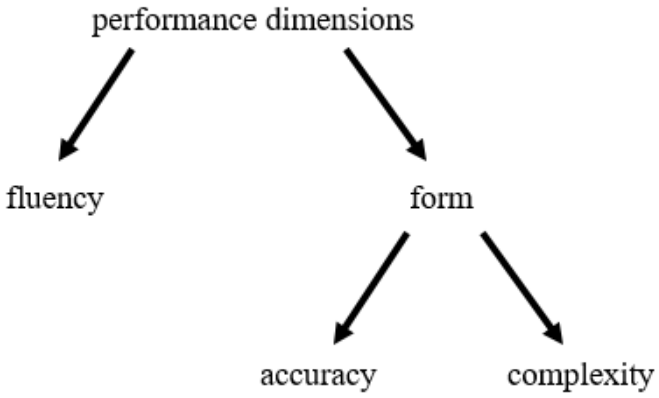


Figure 3. Theorizing dimensions of performance
[Souce: Skehan & Foster, 2001]

The second assumption of the Limited Attentional Capacity Hypothesis comes from Skehan and Foster`s (2001) factor analysis of their 1996 and 1997 studies. They found out that even if accuracy and complexity can load on the same factor, the complexity has a negative loading while accuracy has a positive loading, and this results in higher accuracy when there is lower complexity in language output, and vice versa. The researchers argued that there were trade-off effects between accuracy and complexity and that achieving both greater complexity and greater accuracy at the same time was not possible while performing a cognitively complex task.

As a result, the Trade-off Hypothesis (Skehan,1996; 1998a & b; 2009; 2014) claims that when the complexity of a task increases, fluency of learners performances will increase due to the need of high attentional demands to the meaning. However, there will be trade-off effects between accuracy and complexity. When one of these competing aspects gets attention, the other will not be attended due to the limited attentional capacity of learners. Which of these two dimensions gets more attention depends on the individual differences. A learner may not pay attention to the use of complex structures for the sake of being accurate (safety first approach) while another learner may choose to use more complex sentences in spite of inaccuracy of these sentences (accuracy last approach) (Skehan & Foster, 2001, pg. 189).

2.4.3 Robinson's Triadic Componential Framework and the Cognition Hypothesis

In his task sequencing framework, Robinson (2001 a & b) distinguishes between task difficulty and task complexity. He proposes that task complexity is a result of the cognitive demands imposed by task structure on the language learner whereas task difficulty is a consequence of learner factors such as motivation, proficiency aptitude and so on. which may make the task more or less difficult. Also, he states that task conditions which refer to interactional demands of a task are the third factor that affects the implementation of task-based instruction.

Depending on this framework, the primary claim of Cognition Hypothesis is that pedagogical tasks in L2 learning should be sequenced on the basis of increases in cognitive complexity (Robinson, 2003; 2011). That is, language learners should start with cognitively less complex tasks, and then the level of complexity should be elevated step by step in order to promote L2 development. (Robinson, 2003). Moreover, Robinson (2003) claims that increasing the cognitive demands of tasks along certain dimensions will push L2 learners to produce more accurate and complex language to meet the functional / communicative demands of a task, will increase the attention to output, and will lead to greater input processing depth. As a result, performing tasks in increasing complexity will result in automaticity of the linguistic structures.

After a decade, Robinson (2011) improves his framework and names it as `Triadic Componential Framework` (see Table 2.), which is a part of the `Cognition Hypothesis`. In this recent framework of Robinson, he identifies cognitive dimensions of task complexity in a more detailed way. He bisects cognitive factors as (1) resource-directing and (2) resource-dispersing, and states that these dimensions affect attention allocation

during task performance. Resource-directing dimension of task demands is explained by Robinson (2003) as following:

> …..tasks which require reasoning and reference to many elements and which are displaced in time and space are more complex and attention demanding than their simpler counterparts, but these dimensions have the potential to direct learner`s attentional resources to needed aspects of language code, such as conjunctive coordinators to establish causality, past tense morphology and temporal expressions, and complex nominalizations…. (p. 18)

On the other hand, increasing the task complexity along presence or absence of prior knowledge, presence or absence of planning time, or single vs dual activity dimensions depletes the attention available for performing the task and disperses it over those non-linguistic dimensions (Robinson, 2003). Robinson calls this dimension as resource-dispersing (2011). That is, tasks which are consisted of single activity and where not only prior knowledge but also planning time is available are simpler than the tasks which require the opposite. Robinson (2001a, 2005, 2007) purports that increasing task complexity with respect to the resource-directing dimensions will result in more accurate and complex oral production as learners have to pay more attention to linguistic demands of the task. However, it will result in dysfluent production, as learners have to process language deliberately. In contrast, increasing task complexity along the resource-depleting dimensions demands greater attentional resources and working memory. Moreover, it promotes reinforcement of learnt structures and fast accessibility to the structures that are already in the interlanguage of the learner (Robinson, 2011, pg. 17). As a result, "this can result in less accurate and less complex language production since it complicates learners` attempt to access their current repertoire of L2 knowledge" (Robinson, 2005, pg. 7). However, fluency of production will not be affected since

learners are expected to automatically access their already established interlanguage system to meet the task demands (Robinson, 2007).

Table 2. The Triadic Componential Framework for Pedagogic Task Classification

Task Complexity (Cognitive Factors)	Task Condition (Interactive Factors)	Task Difficulty (Learner Factors)
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	+/- open solution	h/l working memory
+/- Few elements	+/- one way flow	h/l reasoning
+/- Spatial reasoning	+/- convergent solution	h/l task switching
+/- causal reasoning	+/- few participants	h/l aptitude
+/- intentional reasoning	+/- few contributions	h/l field independence
+/- perspective taking	needed	h/l mind-reading
	+/- negotiation not needed	
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	+/- same proficiency	h/l openness
+/- prior knowledge	+/- same gender	h/l control of emotion
+/- single task	+/- familiar	h/l task motivation
+/- task structure	+/- shared content	h/l anxiety
+/- few steps	knowledge	h/l willingness to
+/- independency of steps	+/- equal status and role	communicate
	+/- shared cultural	h/l self-efficacy
	knowledge	

Source: [Robinson, 2011]

All in all, Robinson (2011) claims that increasing complexity through resource-directing dimension will lead to more linguistic complexity, lexical variety, and accurate speech production but to decreased fluency while increasing the task complexity along resource-dispersing dimensions will cause decrease in accuracy and complexity levels but not in fluency. Moreover, increasing complexity through both dimensions will cause learners to suffer in all these three areas of production.

In a nutshell, the claims of Trade-off Hypothesis and Cognition Hypothesis account are quite distinct from each other because they are based on different attentional resources hypotheses. While Trade-off hypothesis bases its claims on limited resource of attention, Cognitive hypothesis assumes that multiple resources of attention are available. Trade-off hypothesis asserts that if we use more challenging tasks, the meaning will be the primary focus of the learner, so the fluency will increase while accuracy and complexity decrease due to limited attentional resources. On the other hand, Cognition hypothesis claims that by increasing task complexity, we increase accuracy and complexity of the linguistic output, but not fluency. It is obvious that these two claims are contrasting, and we need more empirical research to settle parameters for sequencing tasks.

2. 5 Writing anxiety as an individual difference variable in second language acquisition

Writing anxiety (WA) was first defined by Daly (1978) as a "situation and subject-specific individual difference concerned with people's general tendencies to approach or avoid writing." Studies in L1 writing showed that WA as an individual difference has detrimental effects on the quality of written production (Daly & Miller, 1975; Daly, 1977; Faigley et al., 1981).

SLA studies have concluded that individual differences (IDs) are one of the most important predictors of achievement in SLA (Dörnyei, 2005). Specifically, Dörnyei (2009) conceptualized five best acknowledged ID factors in SLA research: language aptitude, motivation, learning styles and strategies, and anxiety. Language aptitude, along with learning styles and strategies are accepted

as cognitive factors while motivation and anxiety are regarded as affective factors in language learning studies (pg. 182-183). Robinson (2011) also conceptualizes anxiety as one of the affective variables under task difficulty dimension of his Triadic Componential Framework.

Actually, anxiety variable was first presented to the attention of L2 researchers in 1970s with Kleinmann`s (1977) study. Kleinmann (1977) measured syntactic avoidance behavior of L2 learners of English with different L1s (i.e. Arabic, Spanish, Portuguese) and found that the grammatical structures used by ESL learners varied with their level of facilitating anxiety avoidance behavior measured by several affective measures. Another important study was from Steinberg and Horwitz (1986) which was carried out with L2 learners of English. The participants were asked to describe a picture, and the anxiety condition was manipulated through the experimental conditions such as comfortable chair, and a smiling tester etc. The students in the anxiety condition produced more denotative (concrete) information whereas non-anxiety group produced more interpretive information while describing the given picture. However, the anxiety construct measured in those studies were not specific to foreign language learning context, instead they were general anxiety measures.

Foreign language classroom anxiety was first conceptualized as a situation-specific anxiety by the studies of Gardner (1985) and Horwitz et al. (1986). Horwitz et al.`s (1986) stated that " foreign language anxiety is a distinct set of beliefs, perceptions, and feelings in response to foreign language learning in the classroom and not merely a composite of other anxieties (pg. 25)." Furthermore, a number of studies have affirmed the existence of foreign language anxiety (Horwitz, Horwitz, &Cope, 1986; MacIntyre &

Gardner, 1989; Proulx, 1991; Truitt, 1995). However, the scales developed to measure foreign or second language anxiety in those studies were problematic because they were dominated by the items measuring anxiety about speaking a foreign language.

Cheng et al. (1999) was one of the first studies to differentiate between foreign language classroom anxiety and writing anxiety. The study investigated the constructs of second language classroom anxiety and second language writing anxiety empirically in EFL context with 433 Taiwanese adult learners of English. Foreign language writing anxiety was measured by the modified Chinese version of the Daly-Miller (1975) Writing Apprehension Test (WAT) which have been designed to the measure writing anxiety of native speakers of English. On the other hand, foreign language anxiety was measured by Horwitz et al (1986) Foreign Language Classroom Anxiety Scale (FLCAS). The results of the study revealed that EFL students can experience writing anxiety, and L2 writing anxiety was a language specific anxiety which is distinct from general classroom anxiety and L2 speaking anxiety. However, Cheng (2004) later stated several problems with Daly- Miller (1975). First of all, this anxiety test was designed to measure the writing anxiety of native speakers of English, and it was not applicable to foreign language environment setting as it includes culture specific items. The second problem was about its construct validity (Cheng, 2004). To clarify the problems concerning the applicability of the Daly-Miller WAT to an L2 learning context, Cheng (2004) developed an L2 writing apprehension scale by considering the multicomponential nature of the writing anxiety. Initially, to construct a pool of scale items, the researcher referred to 65 EFL learners' reports on their L2 anxiety experiences while writing in English. Later, a pilot test was conducted on the initial pool of items to help establish a preliminary

version of L2 writing anxiety scale for further refinement and evaluation in the formal study. Finally, the chosen items were validated with 421 Taiwanese university students. The validated Second Language Apprehension Inventory (SLWAT) was utilized to investigate the relationship between timed essay writings and L2 writing anxiety. Participants' performance on a timed English essay writing task was found to have a statistically significant negative association with writing anxiety. Furthermore, the results provided support for the need of SLWAT given in participants` native language. SLWAT was utilized in a number of EFL writing anxiety studies (Zhang, 2011; Saeedi & Farnia, 2017; Amiri & Saaedi, 2017), and negative associations between WA and students` written production quality was found.

2.6 Empirical research on task complexity in TBLT

Task complexity has been in the focus of many researchers and material designers in TBLT for its impact on L2 learning and, therefore, L2 performance. The two pioneering hypotheses which discuss the impact of task complexity on CAF were posited by Skehan (1998) and Robinson (2001a). Skehan (1998) asserts that it is not possible to allocate attention to both accuracy and linguistic complexity when the task is cognitively challenging because of our limited attentional capacity. On the other hand, Robinson supports the existence of multiple attentional resources. Therefore, he posits that increasing cognitive demands of a task will result in more accurate and more complex language production while fluency will suffer. However, this decrease in fluency can be compensated by giving learners pre-task planning time (Robinson, 2005b).

Following the premises of these two influential hypotheses, the effects of manipulating task complexity on CAF of learners' linguistic outcomes have been investigated extensively since the late 1990s. In these studies, task complexity has been manipulated through resource-directing dimensions (Gilabert, et al., 2011; Michel, 2011; Kuiken &Vedder, 2007; Ruiz-Funes, 2014), through resource-dispersing dimensions (Yuan & Ellis, 2003; Ong & Zhang, 2010; Tabari, 2016, Ellis & Yuan, 2004; Ghavamnia, et al., 2013), through structure of the tasks (Skehan & Foster, 1999; Tavakoli & Skehan, 2005), or through task conditions (Michel, et al. 2007; Michel, 2011). Furthermore, there are a number of studies investigating the influence of task complexity on different language modalities (i.e. speaking and writing) (Kormos, 2014; Tavakoli & Skehan, 2005). However, these studies have revealed mixed results for explaining the effects of task complexity on learners` linguistic production. These inconclusive findings are considered to be due to large variety of research designs and the abundancy of CAF measures (Jackson and Suethanapornkul, 2013).

The current study investigates the effects of task planning on written language production. In this section, therefore, the previous studies investigated the effect of task planning condition on oral and written production will be summarized and a general conclusion about the effect of task planning on CAF will be stated at the end.

2.6.1 Studies on task planning and oral production

There are a great number of studies that have investigated the effects of planning on L2 learners` interlanguage in terms of CAF in oral task performance (e.g. Ellis 1987; Foster&Skehan, 1996; Ortega, 1999; Robinson, 1995; Skehan & Foster 1997; Wendel, 1997; Yuan & Ellis, 2003). However, the ones which will be under the scrutiny of this literature review will be the ones that used a narrative task, operationalized task complexity by providing or not providing pre-task planning time, and investigated its effects on CAF indices. Some of studies (Foster & Skehan, 1996; Mochizuki & Ortega, 2008) provided students some guidance about how to use pre-task planning time while others (Ortega, 1999; Yuan & Ellis, 2003; Tavakoli & Skehan, 2005; Tavares, 2009) did not give any guidance. Here guided planning refers to a type of planning where learners receive detailed guidance about what to focus on during the pre-task planning time.

Foster and Skehan`s (1996) study compared planners and non-planners in three different oral production tasks varying in complexity. They explored participants` performance by personal information exchange, narrative, and decision-making tasks. Planning was operationalized at two levels. In the first level (unguided), participants were given 10-minute solitary planning time, in which they were required to make notes, but the notes were taken away at the end of the planning time. In the second (guided), subjects were given guidance as to how they might use the 10-minute planning time. The guidance took the form of suggestions that attention should be directed to anticipate the language which might be needed during the task. They reported that the guided planning condition produced greater complexity than the unguided planners, and slightly greater fluency. These effects tended to be greater for the more cognitively demanding narrative

and decision-making tasks compared to the personal task. Moreover, planners` language production was more accurate than the non-planners, and the most accurate performers of all were the unguided planners. Foster and Skehan (1996) interpret these results to suggest that the guided planners did indeed complexify the task they were doing, in contrast to the unguided planners who were hypothesized to have used preparation time to rehearse language.

Ortega (1999) investigated whether giving learners pre-task planning time will result in more accurate production or not, and how learners use this pre-task planning time. The participants (*N*=64) retold a story that they heard from the recording under two different planning conditions (-/+ pre-task planning), and after each narration they were interviewed about how they used the opportunity of planning. Results from the linguistic analysis showed that adult nonnative speakers of Spanish were able to produce significantly more fluent and complex language with pre-task planning, but that there were no effects either for lexical range or for accuracy. Results from analysis of the retrospective interview protocols suggested that planning before performing an L2 task can naturally promote a conscious focus on form as participants reported that they planned morpho-syntactic and utterance level aspects along with lexical and semantic aspects of their production.

Yuan and Ellis (2003) investigated the effects of both pre-task and on-line planning on L2 oral production. 42 full-time undergraduate English learners were divided into three groups according to the planning condition and narrated a story based on 6 related pictures provided. In no planning condition participants were given 0.5 sec for

checking pictures and 5 min to narrate whereas in pre-task planning condition, they were allowed to plan for 10 minutes and 5 minutes to narrate. Under online planning condition, participants were allowed to check picture strip for 0.5 minutes and given unlimited time to narrate. The results show that pre-task planning enhances grammatical complexity while online planning positively influences accuracy and grammatical complexity. The pre-task planners also produced more fluent and lexically varied language than the online planners.

Mochizuki and Ortega (2008) investigated the effect of guided pre-task planning related to a specific L2 form on oral performance of beginner English learners. High school students of English in Japan were asked to do an oral story-retelling task with a class partner under one of three conditions: without any prior planning (n = 17), after 5 minutes of unguided planning (n = 20), or after 5 minutes of guided planning (n = 19) with guidance in the form of a grammar handout about English relative clauses. The narratives were analyzed for quantity and correct use of relative clauses along with global complexity and fluency. It was found that the guided planning group produced more relative clauses with more accuracy in their narratives than the other two groups, while showing similar levels of complexity and fluency to those of the other two groups. The results show that guided planning can promote correct usage of grammatical forms which are practiced in the pre-planning stage.

Tavares (2009) investigated the relationship among pre-task planning, working memory capacity, and L2 speech production. 25 intermediate learners of English as a foreign language performed two narrative tasks under different planning conditions. The control group completed both first and second narrative tasks under a no-planning

condition, and the experimental group completed the first task under a no-planning and the second task under a planning condition. The first and second tasks were picture-cued narratives in which participants had to look at 9 pictures and organize them in a sequence in order to tell a story in 40 seconds. Both groups also performed a Speaking Span Test (SST) to assess their working memory capacity. For the task with pre-task planning condition, the results showed a significant increase in fluency and accuracy. However, differences in complexity measures were not significant.

The above mentioned pre-task planning studies (see Table 3 for a summary), especially the ones that gave 10 minutes of planning time, presented fairly similar results. Only one of these studies found a positive effect of pre-task planning on lexical variety, whereas almost all of the studies that operationalized pre-task planning as unguided found advantageous effects on syntactic complexity and fluency. As a result, it can be concluded that pre-task planning has a positive effect on fluency and complexity of oral production. When it comes to accuracy, however, the effects of pre-task planning for oral production are less certain. Another conclusion drawn from this table can be that pre-task planning, guided or not, never yielded a negative effect on any of the syntactic complexity, accuracy, lexical variety, or fluency indices.

These speaking studies in task complexity research can be used as a basis to make inferences on the written production as both language skills are productive and share some similarities in one way or another.

Table 3. Summary of Oral Production Studies

	Proficiency	Length of Planning	Type of Planning	C	A	L	F
Foster & Skehan	Pre-	10 min	Unguided	=	+	X	=
(1996)	intermediate		Guided	+	=	X	+
Ortega (1999)	Advanced	10 min	Unguided	+	=	=	+
Yuan & Ellis (2003)	Not stated	10 min	Unguided	+	=	+	+
Tavakoli & Skehan	Intermediate	5 min	Unguided	+	=	X	+
(2005)	Beginner						
Mochizuki & Ortega	Beginner	5 min	Unguided	+/-	=	X	=
(2008)			Guided	+/-	+	X	=
Tavares (2009)	Intermediate	10 min	Unguided	=	+	X	+

Note: A= Accuracy, F= Fluency, C= Syntactic Complexity, L= Lexical Variety + shows that planning had a positive effect, - = shows planning had a negative effect, = shows a neutral effect, X = not investigated; +/- = shows mixed effects between measurements.

2.6.2 Studies on pre-task planning time

Some studies specifically researched the impact of giving different planning time on L2 production (i.e. Wigglesworth, 1997; Mehnert, 1998). Wigglesworth (1997) researched the effects of planning time, but in the context of language testing. Using just one-minute pre-task planning, she was able to show that performance is improved but that this effect is mediated by the difficulty of the task undertaken and the proficiency level of the candidates.

Mehnert (1998) investigated the effects of systematically manipulating planning time on fluency, accuracy, and complexity of spoken L2 production in German. Different groups of participants were given no planning time, and 1, 5 and 10 minutes' planning time. The study showed that fluency, accuracy and complexity of language were all affected by the opportunity to plan, but in different ways. Increasing planning time led to

increase in fluency; however, the results were not statistically significant. An effect for accuracy was only shown between the non-planners and all the planning conditions. In other words, only one minute planning time is enough to see the effects of planning time on accuracy of the production. Finally, there was no difference in complexity between the no-planners, the one-minute planners and the five-minute planners, but 10-minute planners produced more complex language than all of the groups. The results propose that second language users prioritize different aspects of language during planning time. In other words, first priority is fluency as this was achieved with only 1 minute planning and did not improve with more time (i.e. 5 minutes and 10 minutes). Second priority is complexity as only 10-minute planners achieved it (Mehnert, 1998).

2.6.3 Studies on task planning and written production

A great number of studies have been carried out to investigate oral production differences with regard to task planning time so far, and they tend to indicate that planned speech differs from unplanned in the a number of aspects for L2 learners`. However, for the written production, the number of studies on planning is rather limited. Therefore, reagardless of the task type (i.e. narrative writing, argumentative essay, descriptive writing), all the studies that used task complexity frameworks and measured language production by CAF will be in the scope of this section.

Ellis and Yuan (2004) investigated the effects of pre-task planning (PTP), online planning (OP) and no planning (NP) conditions on fluency, accuracy, and complexity of 42 EFL university students' narrative writings. It was found out

that PTP resulted in increased fluency, and syntactic variety but less accuracy while OP group showed increased accuracy. However, the planning groups outperformed NP group in all of the measures.

Ong &Zhang (2010) study with 108 Chinese EFL univesity students invesigated the effects of increased task complexity through planning time with four conditions (extended pre-task,pre-task, free writing, and control group) on fluency, lexical complexity in the first drafts of students' argumentative essays. Another factor was the writing assitance provided in three distinct levels: topic given; topic and ideas given; topic, ideas, and macro-structure given. All of the four experimental groups were divided into subgroups according to the assisstance type given. It was revealed that fluency measured with mean number of words produced during the writing time was the highest for extended pre-task condition but the results were not significant. Also, the free-writing group had more lexical complexity than others. As for writing assitance provided, it resulted that it has no significant effect on fluency but has significant effect on lexical complexity for topic, ideas, and macro structure given group.

Rahmpour and Safarie (2011) investigate the role of PTP and OP on the performance of descriptive writing. The participants of the study were 37 Iranian sophomore students of Teaching English as a Foreign Language Department. The PTP group had 10 minutes to plan their performance and 17 minutes to finish the task. OP group was not given any time limitation to finish their writings. The findings of the study showed a significant difference between PTP and OP group only in term of fluency but not accuracy and complexity.

Ghavamnia, Tavakoli, and Esteki (2013) studied narrative writings of 40 Iranian EFL adult learners. There were two planning conditions: Pre-task and online. In PTP condition, students were allowed to plan their narratives for 10 minutes and to write their narratives in 20 minutes. On the other hand, in OP condition no time limitation was given for either planning or writing. General fluency, complexity and accuracy measures were used, and it was concluded that PTP group produced more fluent and complex language whereas OP group produced more accurate language. The results of this study affirms the Ellis & Yuan (2004) study. In both of the studies, PTP group outperformed the OP in terms of fluency and complexity.

Another recent study on planning time and writting performance was conducted by Tabari (2016). 78 Iranian undergraduate students' descriptive writings were evaluated in terms of complexity, accuracy, fluency, and lexical variety. Each participant was assigned to one of the planning conditions: NP, PTP, OP. The PTP group produced more fluent writing than the other groups. On the other hand, the PTP and NP group produced more complex language but the difference between them did not reach significant levels. Although it was not statistically significant, the OP group scored higher only in accuracy measures.

Considering the findings of the previous studies on written production, it can be concluded PTP groups tend to produce more fluent and complex writing than online planning groups while OP groups tend to outperform them in terms of accuracy only. However, these results are tentative as only the two studies used control groups (Ellis & Yuan, 2004; Tabari, 2016) and they have mixed results

regarding the NP condition. In Tabari (2016) NP group produced more complex language than OP group whereas in Ellis and Yuan (2004) that group scored the lowest in all of the measures. Nevertheless, it can be proposed that giving learners opportunity for planning before a task can yield more fluent written production (see Table 4).

Table 4. Summary of Written Production Studies

	Proficiency	Task type	Task time	Type of planning	C	A	L	F
Ellis & Yuan (2004)	Not stated * proficient	Narrative	17 min	NP		=	X	=
			10+17 min	PTP	=	--	X	+
			No limit	OP	+	+	X	--
					--			
Ong & Zhang (2010)	Not stated * proficient	Argumen-tative essay	30 min	NP	X	X	+	+/--
			10+20 min	PTP	X	X	--	--
			No limit	OP	X	X	--	--
Rahimpour & Safarie (2011)	Not stated * proficient	Descriptive (ceremony)	10+17 min	PTP	+	--	X	+
			No limit	OP	--	+	X	--
Ghavamnia, Tavakoli & Esteki (2013)	Intermediate	Narrative	10 + 20min	PTP	+	--	X	+
			No limit	OP	--	+	X	--
Tabari (2016)	Not stated *proficient	Descriptive (a picture)	8 min	NP	+	--	X	--
			10 + 8 min	PTP	+	--	X	+
			No limit	OP	--	+	X	--

Note: A= Accuracy, F= Fluency, C= Syntactic Complexity, L= Lexical Variety + shows that planning had a positive effect, - = shows planning had a negative effect, = shows a neutral effect, X = not investigated; +/- = shows mixed effects between measurements.

2.7 Studies on L2 writing anxiety and CAF

The literature has suggested that there is a negative relationship between WA and writing performance in English as a native language (Daly & Miller, 1975; Daly, 1977; Faigley et al., 1981). More specifically, L1 writing anxiety research has found out that native-speaking students with lower levels of WA write longer compositions, use more complex structure, evaluate their writings more highly on quality, and produce more intense messages (Daly, 1977; Daly & Miller, 1975). Similarly, Faigley, et al. (1981) also suggested that having high apprehensiveness results in shorter and less fluent writing than having low apprehensiveness. Moreover, highly apprehensive writers have a tendency to avoid the activities that is needed to be successful writers such as writing, practicing writing, and getting feedback on writing. As a result, many apprehensive learners end up showing quite poor performances in writing (Daly, 1978). Here, the studies investigating the correlation between L2 writing anxiety and quality of written production of EFL learners will be presented.

Zabihi (2017) is the only study that scrutinized the effects of writing anxiety along with self-efficacy and working memory on CAF. In the study, 232 upper-intermediate Iraninan EFL university students were given picture strips and were required to complete a narrative writing in a very limited time, i.e. 11 minutes without giving any planning time. The results presented that writing apprehension was negatively correlated with all CAF measures at significant levels.

Saeedi and Farnia (2017) investigated the interrelationship between EFL learners` writing apprehension levels and the linguistic quality of their narratives.

45 Iranian high intermediate university students completed the test of SLWAI (Cheng, 2004) and produced a narrative based on picture prompts in 3 minutes time. For investigating linguistic quality of their written language, measures of complexity, accuracy, and fluency were used. For complexity, the ratio of clauses to AS-units was utilized, and for accuracy the percentage of error free clauses to the total number of clauses was used. Regarding fluency, the total number of syllables per minute was calculated. The results presented a negative association between L2 writing anxiety and measures of complexity, accuracy, and fluency.

Haddad (2018) compared the writing anxiety levels and the number of subordinate clauses used in written compositions of Arabic speaking freshmen and senior students. Senior students writing anxiety levels were statistically lower than the freshman students and seniors tend to use more subordination than freshmen. Furthermore, a significant negative correlation was found between the level of writing anxiety and the grammatical complexity of the written compositions of the participants.

Lee (2005) investigated the effects of writing anxiety, writing blockage, free reading and writing on writing performance. 270 Taiwanese university students with ranging proficiency levels form low to high intermediate wrote a short essay on a given topic in 40 minutes. The relationship between WA and writing performance was found to be non-significant. However, it was revealed that the amount of free reading done by the participants was the only predictor of writing performance.

Choi (2014) examined how foreign language anxiety is related to foreign language writing anxiety. The study also investigated the correlation between writing anxiety and writing performance. 26 Korean college students submited a writing portfolio

at the end of semester. A statistically significant correlation between foreign language classroom anxiety and second language writing anxiety was one of the conclusions reached by the study. However, no statistically significant correlation between second language writing anxiety and actual writing performance was found as measured by the cumulative scores on the portfolio assignment.

The studies meantioned above report mixed results for correlations between WA and writing performance. However, the studies that used CAF framework (Zabihi, 2017; Saeedi & Farnia, 2017; Haddad, 2018) instead of holistic gradings or a rubrics to evaluate students` written production have presented negative correlations between writing anxiety and L2 measures. Therefore, it can be concluded that EFL students with lower WA levels tend to produce more complex, more accurate, and more fluent writing production than students with higher WA levels.

CHAPTER 3

METHODOLOGY

3.1 Introduction

A relatively under-researched area in the field of L2 writing is the role of task complexity in writing performance despite the fact that writing is an intricate skill both for language teaching pedagogy and for the ESL assessment. Gaining insights into how different task characteristics and conditions are associated with the linguistic and discourse features of written text production could assist both language teachers and test designers in selecting tasks that have the potential to elicit writing performance thoroughly and without biases. (Skehan, 2001)

Considering this research gap in the field, the purpose of this study is to explore the effects of task complexity on the accuracy, fluency and syntactic complexity of beginner and elementary level adult English language learners` narrative writings. In the present study, task complexity was manipulated by using one resource-dispersing factor: the availability of pre-task planning time. All of the participants carried out the tasks under two different planning conditions: pre-task planning available and pre-task planning unavailable. These conditions are named consecutively as no-planning (NP) and pre-task planning (PTP). Also, the mediating effect of writing anxiety on learners' written performances in terms of CAF framework was investigated in the current study. The reason why WA was investigated in relayion to CAF measures is that in the L1 and L2 writing literature, almost all studies on WA utilizes subjective criteria to assess writing performance; thus the results of these studies are not comparable and do not produce

consistent results. This study uses use objective indices in order to assess writing performance and, thus, scrutinize WA and writing performance by reducing the subjective factors to minimum.

3.2 Research questions and hypotheses

The main focus of the study is to investigate whether accuracy, fluency and syntactic complexity levels of beginner and elementary level EFL Turkish learners` written production improve in response to narrative writing tasks given under two planning conditions: NP (no-planning) and PTP (pre-task planning). Secondarily, it is aimed to find out the relationship between writing anxiety and CAF dimensions. Therefore, based on the previous research and hypotheses, this study was designed to answer the following research questions:

1. How does manipulating task complexity along +/- pre-task planning time affect the syntactic complexity, accuracy, and fluency levels of EFL learners` written performance?

2. Is there any relationship between WA levels and CAF levels of EFL learners under no-planning condition?

Related to these research questions, the following hypotheses are advanced:

1. Regarding the effect of pre-task planning on language performances of learners, Skehan (1998) and Robinson (2011a) have made different predictions. Following the Limited Attentional Capacity Hypothesis, Skehan (1996) asserts that language learners can focus more on conveying the meaning during the execution of a difficult task. As a result, they can prioritize fluency over complexity and accuracy when the task is cognitively demanding. Furthermore, Skehan (1998a) proposes that when learners priorotize meaning their fluency levels increase; however, with planning time either

accuracy or complexity along with fluency could increase due to the limited attentional resources of learners. Learners can allocate their attention to just one of them depending on the task condition and their individual preferences. For instance, guided planning can promote complexity while unguided one can promote accuracy. On the other hand, Robinson (2001a & 2011a) describes pre-task planning condition as a resource-dispersing dimensions of a task and proposes that giving planning time, in other words decreasing the complexity level of a task through resource-dispersing dimension, makes greater demands on attention and working memory, but do not direct resources to features of language code that can be used in completing the task.Therefore, accuracy and complexity levels can be affected adversely under the resource-dispersing complex task condition. Moreover, he claims that " Performing increasingly complex versions of tasks on resource-dispersing dimensions promotes …… consolidation and fast real-time access to existing interlanguage resources. (Robinson, 2011, pg 17)"

As for the availability of planning time, it can be inferred that Skehan expects increased fluency under NP condition whereas under PTP condition an increase in either accuracy or complexity as well as fluency was expected. On the other hand, Robinson expects an increase in learners` performances in terms of fluency under NP condition as well because this condition will push learners to access their interlanguages faster; however, accuracy and complexity will suffer under NP condition. Based on these claims, I expect participants` written productions tobe more fluent under NP condition, but to be more accurate and/or syntactically complex under PTP condition.

2. The effect of individual differences (ID) in terms of writing anxiety is expected to have an influence on text production in terms of CAF dimensions. According to studies

of Daly (1975) and Daly & Miller (1975), participants with less writing anxiety tend to produce longer narratives in their L1s than participants with high anxiety, that is, they were more fluent. In L2 studies, highly apprehensive participants produced less complex (Zabihi, 2017; Saaedi & Farnia, 2017; Haddad, 2018), less accurate (Zabihi, 2017; Saaedi & Farnia, 2017) and less fluent (Zabihi, 2017; Saaedi & Farnia, 2017) writing. In these L2 studies, all the participants were given limited time (i.e. 3 minutes) for completing their narratives without any extra planning time. As a result, I expect a negative correlation between writing anxiety and complexity, as well as fluency under NP condition. However, I expect positive correlations between writing anxiety and accuracy under NP condition because the accuracy measure used in this research is errors per T-unit which is the opposite of error free T-units measure used in the previous studies. In other words, the accuracy measure used in the current study focuses on the number of errors whereas the measures in previous studies focused on accurate language production. Therefore, the direction of the correlation is expected to be the opposite of the preceding studies.

3.3 Research design

This study is a quantitative repeated-measure design in which 45 Turkish speaking English learners performed two narrative tasks under two different planning conditions (i.e. no planning and pre-task planning). There are two independent (writing anxiety and pre-task planning time) and three dependent variables (syntactic complexity, accuracy, fluency) in the study. Each participant produced two written narratives based on two different sets of pictures (see Appendix A) under both planning conditions. Then,

their written narratives were analyzed in terms of syntactic complexity, fluency and accuracy.

Although the same participant produced the narratives under both conditions, there were two groups in the study due to participants being in two intact classes. Half of participants were in the first group and the other half was in the second group. The task orders and the narratives used were counterbalanced in each condition to avoid any confounding variables (see Table 4). Therefore, the researcher was able to eliminate the risk of any confounding variables that can be related to the story and the task planning type. Also, writing anxiety was investigated as a mediating independent variable to investigate the effect of individual differences on the written task production.

3.4 Participants

45 full-time undergraduate Turkish students studying English at a university preparatory school in Turkey participated in the study. The participant group consisted of 30 males and 15 females whose ages ranged between 18 and 22. All of the students were studying at English-medium technical university, and their departments are related to engineering. 44 participants are Turkish monolinguals and only one participant is a Turkish and Kurdish bilingual. None of the participants have lived in a foreign country before.

All the participants had started learning English at Grade 4. However, it was their first time to write a story based on pictures as none of them stated that they had done such an activity during their previous English lessons. According to Quick Oxford Placement test given by the researcher, their English levels are mixed. 30 of them are A2 level and 15 of them are A1 level. A1 is defined as the beginner level and A2 is defined as the

elementary level by the Common European Framework of Reference for Languages (CEFR).

Each student was asked to participate in the study voluntarily and signed a consent form. This study was carried out as a research project, and students were told that their written narrative would not be graded and affect their course grades. On the day of the data collection, while the participants attended their regular class, the researcher and the teacher collected the data.

3.5 Data collection instruments

Two equivalent tasks for narrative writing were designed to eliminate the practice effects under two planning conditions. A background questionnaire about their English learning history was given to all participants (see Appendix B), and in order to investigate individual differences, second language writing apprehension inventory in Turkish (see Appendix C and D) was given after the completion of the tasks.

3.5.1 Tasks

Two narrative writing tasks were assigned to the participants. In order to increase authenticity of the task and to provide a reason for communication, students were told that they should write the stories as if they were narrating it to a friend. Narrative tasks were chosen so that the results of this study would be comparable with the results of the previous studies which have investigated the effect of planning on similar oral tasks (e.g. Ellis, 1987; Foster & Skehan,1996; Ellis & Yuan, 2003), and on similar written tasks

(e.g. Ellis & Yuan, 2004; Ghavamnia, Tavakoli, and Esteki, 2013). Another reason to choose that kind of task is to ensure that the task was cognitively demanding for the participants with lower level English proficiency so that participants` linguistic competence could be stretched by the high demanding task. Previous research on oral tasks (e.g., Skehan & Foster, 1999) has indicated that this can be achieved by selecting a picture story that requires interpretation on the part of the learners. Moreover, both of the stories require some interpretation because we do not know where the dog is until the last scene or why the man is chasing the boy (Ellis &Yuan, 2004). Also, not only the foreground events but also the background events are presented in both picture strips. This is eminent because it was argued by Tavakoli & Foster (2008) that stories with not only foreground but also background events can lead to more subordination as students have to use conjunctions such as "when, while, before, after" to narrate what is happening in the background. Moreover, the same picture strips were used as a part of Tuzcu (2018) study and they were regarded as comparable in terms of structural complexity. Therefore, in the current study, I decided to utilize The Picnic and The Chase stories.

The two narrative writing tasks required participants to write a story based on a set of pictures given from the book "Compositions through pictures" (Heaton, 1975). One of the stories (the Chase task) is about a boy getting off the bus and dropping one of the packages he is carrying. The boy sets off for home without noticing that package. Then he realizes that a man with hat is following him. After realizing that, he starts to run because he is scared. Eventually, the man catches up with the boy and returns the package. However, his reason for chasing the boy is initially unclear, and it only reveals

in the last scene of the story. The second story (The Picnic task) is about siblings and their dog. The boy and the girl prepare sandwiches for a picnic. The puppy is checking the basket while their mother is showing them where to go for the picnic. The children say goodbye to their mother and climb up a hill to eat their sandwiches, but when they open the basket, the dog jumps out of it. After that, they realize that the dog has eaten all of the sandwiches (see Appendix A for the picture stories).

The task instructions were given in Turkish to avoid any misunderstanding. All of the participants were given the same prompts to start the stories. In order to eliminate the effect of the stories` content and language, half of the participants were given `The Chase` story first while the other got `The picnic` story first under the same planning condition. Taking into account the limited language proficiency of learners, the researcher provided one word (either in noun or verb form) for each picture in the stories in order to make sure that the participants have the required vocabulary to write the story in the no-planning conditions, as well.

3.5.2 Writing Apprehension Test

Second Language Writing Apprehension Inventory (SLWAI) developed by Cheng (2004) was translated into Turkish by the researcher. The translated version (see Appendix C) was checked by two other English teachers for any translation mistakes. SLWAI with .91 Cronbach Alpha has been reported to have a high level of internal consistency (Cheng, 2004). Cheng (2004) developed the items in this inventory by using learners` self reports of their writing anxiety experiences and utilized the previous L1 and L2 writing anxiety

scales. This test is utilized in the current study due to the fact that it is specifically designed for L2 learners and has internal consistency.

A Likert-type scale with five possible responses to each of the questions was used in scoring the test. The scale ranged from 1 (strongly disagree) to 5 (strongly agree). All of the positively worded statements, such as "I would not be nervous speaking English with native speakers" and "I enjoy writing in English," were reverse scored.

3.5.3 Background questionnaire

In the background questionnaire (see Appendix D), the students were asked to give information about their demographic information such as age, gender, and mother tongue. Additionally, the questions about how long they have been learning English, and where they have learnt it were asked. A question about if they have ever carried out a task like the one in the study was also included.

3.6 Pilot study

A comparable pilot study was used to set the time limit for writing the narratives. Participants from the first group was given a set of pictures which was also from Heaton`s book (1975) at least one month before the intervention to prevent the practice effect. The set of pictures given in the pilot study was totally different from the pictures used in the current study. In the pilot study, no time limit was set for the writing of the narrative, and the each participant`s writing time was noted. The pilot study had established that the fastest writer completed the story in 13 minutes whereas it took the

slowest 32 minutes. The word limit was set as 100 words in the pilot study because it was the criteria in Ellis & Yuan (2004).

3.7 Procedures

This study was conducted in the preparatory school of a state university in Turkey. Two intact classes were chosen. The participants performed the task in their regular classroom setting. Both their regular teacher and the researcher were in the class during the execution of the tasks.

The data collection took place in one session, in 90 minute block hour. As to eliminate the effect of stories on the conditions, the two intact groups wrote the same stories under different planning conditions. Table 5 explains the order of the stories and conditions for the two groups.

Table 5. Research Design

	Story	Task Condition
Group 1	Picnic	Pre-task Planning
	Chase	No Planning
Group 2	Chase	No Planning
	Picnic	Pre-task Planning

First, the students signed the consent form to confirm they volunteered to participate. Then, they were given the pictures of story for their first planning condition. After they had finished writing about the first story, the background questionnaire was

handed, and this allowed the student to rest for a while. After that, they were given the other story to complete. To measure students` writing anxiety, they also completed a questionnaire with Likert scale. Finally, the teacher distributed the 40-question Quick Oxford Placement Test, and gave students 30 minutes to complete it.

3.8 Task conditions

In this study, planning is operationalized at two levels: no-planning (NP), and pre-task planning (PTP). Table 6 shows the specifications of the task conditions.

In the NP condition, participants had only 13 minutes to finish their at least 100-word story. In this way, the participants were pressured to perform the task with limited opportunities for online planning. It is believed that 13 minute time limit is adequate to create a speeded writing condition, which is needed to limit participants` online planning opportunity, because the fastest writer finished his story in that amount of time in the pilot study and he was able to write more than 100 words. Therefore, in this condition students did not have the opportunity to plan their narratives in detail as they were under pressured online condition (see Appendix E and F for task instructions).

In the PTP condition, the picture cues were distributed to the participants, and they were allowed to look at the pictures and plan their narratives for 10 minutes which was based on Mehnert (1998). As in the NP condition, participants were requested to finish writing their stories within 13 minutes and to produce at least 100 words. To increase the chances that they would indeed engage in planning, the participants were provided a spare sheet and asked to make notes about what they were going to write in the narratives. However, they were told that these notes would be taken away before they

began to write. Moreover, the participants were not given any guidance about how to spend this pre-task planning time. Therefore, this condition was operationalized as unguided pre-task planning. When the planning time was over, their notes were collected by the teacher in order to ensure that the narrative was produced within the specified time limit (see Appendix E and F for task instructions). Also, the notes were considered as data to gain more insight into learners` planning preferences in the further analysis.

Table 6. Task Conditions

Task condition	Time for Pre-task planning	Timer for Online planning
No planning	None	13 minutes
Pre-task planning	10 minutes	13 minutes

As they were A2 level students, they had quite limited vocabulary, students were allowed to ask about any unknown vocabulary to their teacher or the researcher, and also for each picture, words were given as cues on the pictures under both task conditions. The researcher and the teacher helped the learners with vocabulary, but not for grammar as this would affect accuracy analysis.

3.9 Measures

Linguistic production was measured following a multidimensional approach with indices for syntactic complexity, fluency and accuracy that are distinct and complementary as

proposed by Norris and Ortega (2009). Measures of accuracy, complexity, and fluency are developed to evaluate the quality of the participants' written production. Insofar as possible, similar measures as those used in task complexity studies of written and oral production (Yuan & Ellis, 2004; Ghavamnia, et al., 2013; Ruiz-Funes, 2014; Mehnert, 1998; Gilabert, et al., 2011; Tabari, 2006; Kuiken & Vedder, 2007, 2008; Michel, et al., 2007; Michel, 2011; Tavakoli & Skehan, 2005) were employed to validate the results of this study.

3.9.1 Accuracy measures

Errors per T-unit and errors per 100 words measurements were utilized for the accuracy analysis (see Appendix G for the sample error analysis).

3.9.1.1 Errors per T-unit

This measure was calculated by dividing the total number of errors by the total number of T-units produced. In this study, errors per T-unit measurement was used instead of commonly used error-free measurements because these error-free indices have been criticized by Bardovi-Harlig and Bofman (1989) as a consequence of not being able to show how errors are distributed within the T-unit or any other unit of analysis. That is, a T-unit containing a single error is treated identically to a T-unit containing multiple errors. Also, another reason to utilize this measurement instead of error-free indices is that participants` writings contained a high number of mistakes, and even some of them hardly had any error-free T-units (Kuiken & Vedder, 2007). This probably was the result

due to the low proficiency level of students and that they have not reached much automaticity in production like the intermediate level learners in the other studies.

3.9.1.2 Errors per 100 words

This measurement is a typically used one in psycholinguistics, and is accepted more precise than general measures such as error-free T-units as error-free T-units measurements do not specify the number of errors within each T-Unit. (Gilabert et al. , 2011). To calculate this measure the total number of errors were divided by the total number of words. Then the result was multiplied by 100 to get the percentage of the errors to the total number of words.

3.9.2 Syntactic complexity measures

For assessing syntactic complexity, two different measures were used: mean length of T-units and subordination ratio. A measure for length was used as Palotti (2015) defines complexity as the number of different linguistic elements and the interconnections between them, which in return results in a longer text production. Therefore, it was inferred that the longer the T-unit, the more syntactically complex the T-unit is. On the other hand, following Norris and Ortega (2009), subordination ratio is also utilized to capture the syntactic complexity at sentential level.

3.9.2.1 Mean length of T-units

The mean length of T-unit is regarded as an overall syntactic complexity measurement (Norris & Ortega, 2009), and it was calculated by dividing the total number of words to

total number of T-units. This measure is used to reflect the global syntactic complexity of learners` performances.

3.9.2.2 Subordination ratio

The ratio of clauses to T-units, that is subordination ratio defined by Hunt (1965), was utilized in this study. It is calculated by dividing the number of clauses to the number of T-units in a student`s narrative writing.

3.9.3 Fluency measures

Two measures were employed to investigate the writing fluency of the participants. Tavakoli and Skehan (2005) defines three sub-dimensions for fluency: speed, breakdown, and repair. Due to the nature of the written task and the environment of study, the researcher was only able to use two measures which only could tap into the rate or speed sub-dimension of fluency (i.e. mean number of words per minute and mean number of syllables per minute). As the participants produced their narratives in a classroom setting with pen and pencil, the researcher did not have the chance to observe breakdown or repair fluency.

3.9.3.1 Mean number of words per minute

After the total number of words for each narrative was counted, the sum was divided by 13 which is the time spent during writing. This measure provides us with a general estimation of speed of the participants` writing.

3.9.3.2 Mean number of syllables per minute

The total number of syllables for each task was counted and divided by the time spent during writing. This measure enables us to measure the rate of writing in a more sensitive way as syllable count can provide more specific results than word count.

3.9.4 Writing Anxiety measure

The writing apprehension questionnaire used in the current study was Second Language Writing Anxiety Inventory (SLWAI) designed by Cheng (2004) to assess ESL writing anxiety. It consists of 22 items, and all the items were validated by Cheng (2004) in EFL context of Taiwan. The SLWAI scored on a five-point Likert response scale ranging from 1(strongly disagree) to 5(strongly agree). Six of the items (9, 10, 11, 15, 16 and 20) are negatively worded and require reverse scoring. The maximum score which indicates a hig level apprehension is 74 while for minimum score it is -14. The SLWAI has been proved to be a valid and reliable measure of L2 writing anxiety by Cheng (2004). In the current study, therefore, it was employed as the tool for measuring participants' writing

anxiety. The questions were translated into Turkish by the researcher and checked by another English teacher.

3.10 Units of data analysis

In the task complexity studies of written production, T-units and clauses are the most commonly used indices while analyzing the data. T-unit or minimal terminable unit was coined and defined by Hunt (1970) as main clause plus any subordinate clause(s) or non-clausal expression(s) that are attached to or embedded to the main clause. It is a linguistic measurement which aims to measure the smallest word group that can be defined as a grammatical sentence regardless of the punctuation. Any complex or simple sentence would be counted as one T-unit. Any compound or compound complex sentence would consist of two or more T-units (Hunt, 1965). Clause is defined as a subject or a coordinate subject plus a finite verb or coordinate verbs (Hunt, 1965). In his 1965 report, Hunt created this linguistic measurement to gauge how school children`s L1 syntactic maturity develops from 4^{th} to 8^{th} and then to 12^{th} grade. It is concluded in Hunt (1965) that in the earlier stages of writing development, students tend to use coordinated and run-on sentences which gradually decrease with the grades. In the later stages, they tend to produce increasing number of subordination, and embedded clauses. He also asserts that the length of T-unit is in line with a child`s cognitive development, so the T-unit length increases during developmental stages. Since then, the T-unit length has been used widely in both L1 and L2 studies to understand students` overall syntactic maturity both in their writing and speech samples (Hunt, 1965). Appendix G presents samples for the T-unit analysis used in the current study.

3.11 Coding

Students' narrative writings were coded to calculate students` accuracy, fluency, and syntactic complexity levels. For accuracy, the researcher and a second coder highlighted and counted the errors with respect to syntax, morphology, and lexical choice. Lexical errors were defined as errors in lexical form or collocation. The following are some examples of errors coded for each type of mistakes from the writings of the participants:

Syntax

(1) When children looked into the picnic basket, they can`t see anything.

(couldn`t)

(2) Put in place after basket, jumped out of a dog.

(sentence structure)

Morphology

(3) Because the dog eated all food.

(ate)

(4) ……. because they haven`t any food.

(didn`t have)

Lexical Form/ Collocation

(5) Children were very scary.

(scared)

(6) they decided to make a picnic on the hill.

(have)

No semantic mistakes were counted as errors due to the low proficiency level of students and limited lexical knowledge of students. Also, no article mistakes were analysed due to the same reason.

For syntactic complexity, the sentences in students` narratives were divided into T-units following the definition of Hunt (1965; 1970) and the number of clauses counted. T-unit is the smallest word group that can be defined as a grammatical sentence regardless of punctuation (Hunt,1965). A clause is defined as a subject or a coordinate subject plus a finite verb or coordinate verbs (Hunt, 1965). Following is a sample for T-unit and clause analysis:

(1) /They felt surprised and sad/, /because they haven`t any food/ (1 T-unit, 2 clauses)

(2) /When children looked to in the picnic basket/, /they can`t see anything/ /Because the dog ate everything in the picnic basket/ (1 T-unit, 3 clauses)

For fluency measures, the narratives were transcribed and the number of words counted with the help of Microsoft Office Word. Then the number of syllables for each narrative were counted by the researcher. The researcher used Cambridge Online Learners` Dictionary when she was in doubt about the number of syllables in a word.

3.12 Data analysis

Participants` hand written narratives were typed into Microsoft Office Word verbatim for the analysis. First of all, the number of words that participants produced under both task conditions was counted and the participants who wrote under 80 words at least in one task condition were eliminated. The researcher could not stick to the 100-word criterion that was set at the beginning of the research because most of the students were able to write less than 100 words at least under one task condition. This could be due to the time limitation given for tasks, or the opportunity to ask teacher vocabulary questions during the task might have lessened the time spend on the task. After eliminating the participants according to the 80-word criterion, only 20 students out of 45 were eligible for the further analysis. 17 of these eligible participants are A2 level and 3 of them are A1 level. The number of boys and girls is equal. Then the researcher and a second coder coded the data of these 20 participants independently. Before coding, the researcher sent guidelines (see Appendix H and I) to the second coder and talked about the coding process on the phone. The narratives were coded for clauses and T-units by the researcher and the assistant with an interrater reliability of 92.5% and 87.5% respectively. Inter-rater reliability was also calculated for the error analysis and reached 82.5%. All inter-coder reliability was calculated by Pearson Correlation. The researcher and the second rater could not persued any resolutions for disagreements, and the researcher used her own coding for the analysis.

The acculumated data was entered to SPSS version 21.0 for the analysis and checked with tests of normality and graphics such as histograms, box plots and P-P plots for the normal distribution. As a result, all of the data was found to be non-normally

distributed. In that case, non-parametric tests are to be used (Field, 2009). The non-parametric data for 20 students was examined with descriptive statistics, confidence intervals, effect sizes and test statistics. The Wilcoxon signed-rank, which is equivalent of dependent t-test, was applied for accuracy, complexity and fluency data, and Spearman`s Correlation Coefficient, which is a non-parametric version of Pearson`s Correlation, was used to investigate the relationship between writing anxiety and CAF.

Confidence interval graphs were reported for each data set because according to Field (2009) they are very useful for comparing the mean scores. If the mean score of one condition exceeds the upper 95% confidence interval of the other condition, it is highly likely to find out a significant difference between them (pg. 47).

According to Field (2009), calculating and reporting effect sizes is crucial because they are a standardized measure of the size of effect observed so people can compare them with other studies` (pg. 550). For effect size Pearson`s *r* was used in this study. As SPSS does not calculate it, I used the formula from Rosenthal (1991) as mentioned in Field (2009). The effect size for each measurement was calculated as:

$$r = \frac{z}{\sqrt{n}}$$

Here z is the z-score that SPSS produces and N is the size of the study (i.e. the number of total observations). For reporting how big an effect size is, the criterion that Cohen (1988) suggested were used while reporting the data. Following Cohen (1988), $r = .10$ is accepted as a small effect size whereas r = .30 is medium and r = .50 is a large one (pg. 77-81). Hence, the interpretations for effect sizes were made accordingly.

CHAPTER 4

RESULTS

4.1 Introduction

The data of 20 students who have met the criterion of word limit were analyzed and reported in this section. Firstly, I will present the results for accuracy, syntactic complexity and fluency measures consecutively. Then, the results of Spearman`s Correlation Coefficient for the writing anxiety survey will be reported. For each measurement, the descriptive statistics along with error graphs will be shown, and then non-parametric test results and effect sizes will be reported.

4.2 Accuracy measures

Accuracy was measured by the total number of errors per T-unit and errors per 100 words which are common accuracy measures in task-based research. Table 7 shows the means, standard deviations, minimum and maximum values of accuracy measures of 20 participants in two tasks.

Table 7. Descriptive Statistics for Accuracy Measures

	No-planning				Pre-task planning			
Accuracy Measures	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Errors per T-unit	.53	.22	.21	.94	.38	.19	.12	.91
Errors per 100 words	7.26	2.50	3.44	13.10	5.37	2.30	1.78	10.70

4.2.1 Errors per T-unit

According to the descriptive statistics results, as shown in Table 7, participants made more errors in NP condition (*M* = .53, *SD* = .22) than PTP condition (*M* = .38, *SD* = .19). When it comes to minimum and maximum values, only for the minimum value, there is a large difference between the conditions. On the other hand, maximum values were quite similar. When 95% confidence intervals were examined (see Figure 4 below), it is clear that error bars overlap by less than half. According to Hollands and Jarmasz (2010), this is an indication of a statically significant difference between the means. Therefore, it can be stated that there is a significant difference between the two conditions in terms of total number of errors per T-unit.

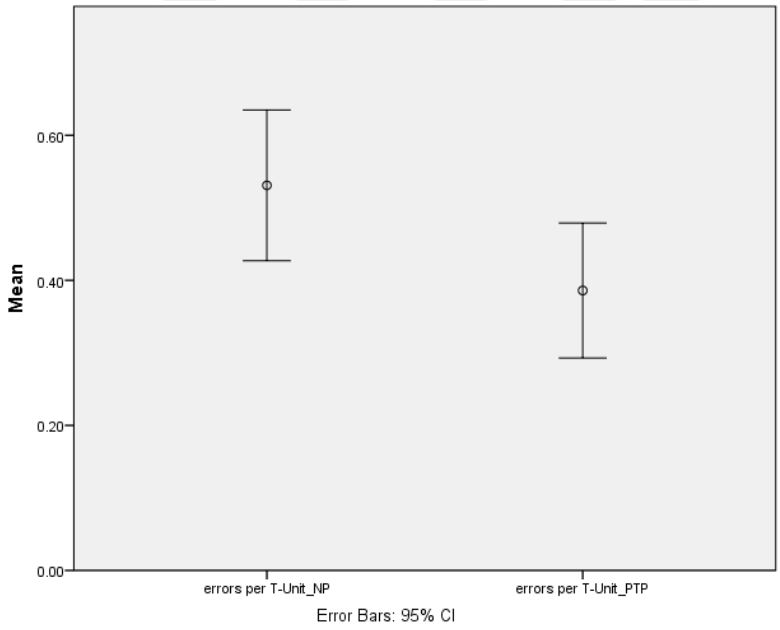


Figure 4. Errors per T-unit with 95% confidence intervals

When the Wilcoxon signed-rank test was conducted, the total number of errors per T-unit was significantly lower in pre-task planning condition (*Mdn* = .38) than in no-planning condition (*Mdn* = .52), $T = 6$, $p < .05$, $r = -.51$. This means that availability of pre-task

planning time significantly resulted in a decrease in the number of errors per T-unit. Moreover, the effect size is (r = -.51) a large one (Cohens, 1992).

4.2.2 Errors per 100 words

For the second accuracy measure, the Table 7 shows that under NP condition there were more mistakes per 100 words in students` narratives than under PTP condition. Standard deviations of both conditions were quite similar as it is a repeated measure design with the same participants. In this accuracy measurement, contrary to the previous measurement type, the minimum and maximum values of the task conditions were very different. However, according to the 95% confidence intervals, there is again an overlap between the two error bars by less than half of their length (see Figure 5). Following Holland and Jarmasz (2010), we can again state that this difference is also significant.

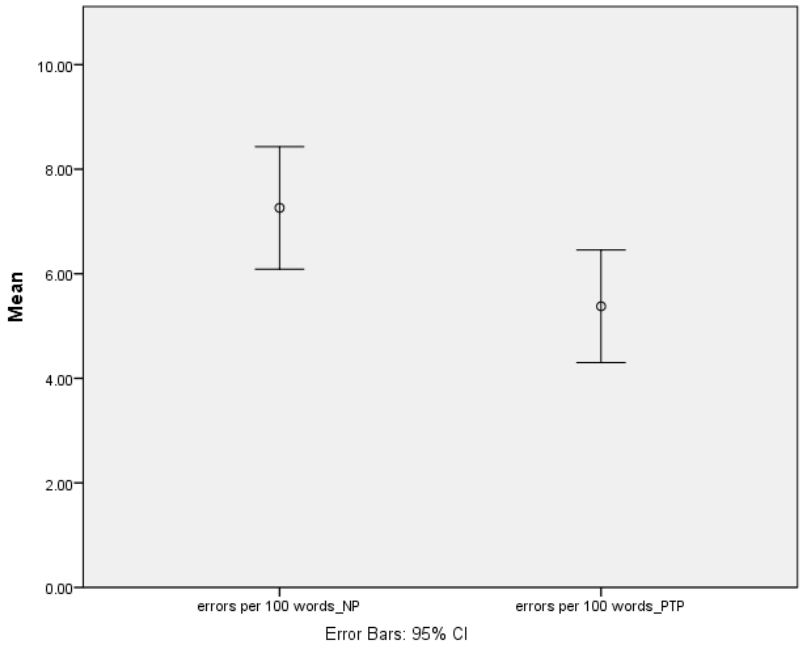


Figure 5. Errors per 100 words with 95% confidence intervals

Also, Wilcoxon signed-rank test was run for investigating the statistical significance of differences between the mean scores. Accordingly, the total number of errors per 100 words was significantly lower in PTP condition (*Mdn* = 5.37) than no planning NP condition (*Mdn* = 7.26), $T = 6$, $p < .05$, $r = -.50$. For this accuracy measurement also the effect size ($r = -.50$) represents a large one.

All in all, the accuracy scores for both measurements revealed a significant difference between the conditions and a large effect size for pre-planning conditions. Therefore, it can be stated that carrying out a narrative writing task with pre-planning time leads to greater accuracy levels and the relationship between planning condition and the accuracy levels is a strong one. Therefore, the statistical analysis results confirmed the hypothesis one.

4.3 Syntactic complexity measures

The study investigated syntactic complexity with two measures; the mean length of T-units and the ratio of subordinated clauses to total number of clauses. First one is an overall complexity measure while the second one is a more specific one. The Table 8 shows the descriptive statistics for both syntactic complexity measures.

Table 8. Descriptive Statistics for Syntactic Complexity Measures

	No-planning				Pre-task planning			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Length of T-units	7.03	1.25	4.89	9.14	7.18	1.46	5.19	11.10
Subordination Ratio	1.29	.21	1.00	1.66	1.26	.14	1.00	1.60

4.3.1 The mean length of T-units

When the results of descriptive data were examined for this overall syntactic complexity measure, it is clear that it is almost similar for both conditions: NP condition (*M* = 7.03, *SD* = 1.25) and PTP condition (*M* = 7.18, *SD* = 1.46). Also, the confidence interval Figure 6 with the overlapping error bars ratifies the descriptive statistics.

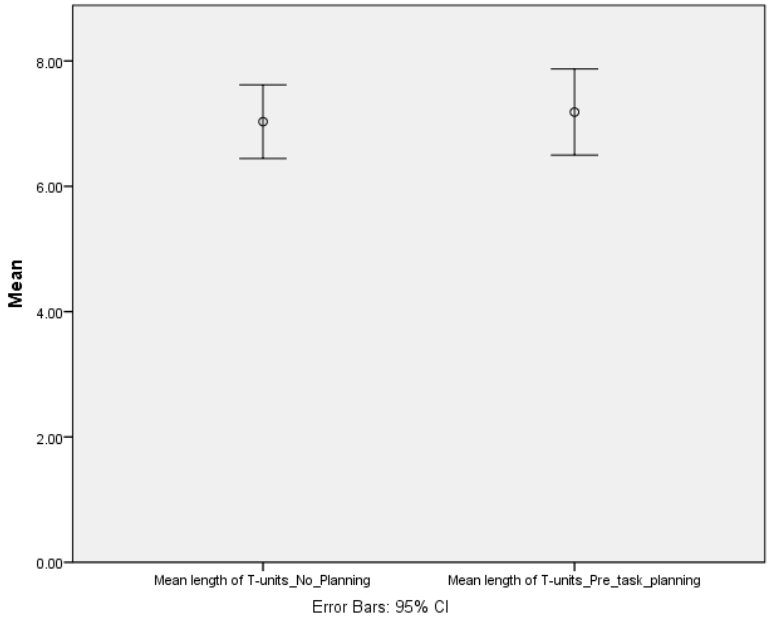


Figure 6. 95% Confidence intervals for T-unit length

When the Wilcoxon signed-rank test was conducted, it did not indicate any significant difference in the mean length of T-units between the PTP (*Mdn* = 6.88) and NP conditions (*Mdn* = 6.97), *T* = 10, *p* > .05, *r* = -.06. Also, the effect size is not even indicate a small effect of planning on the complexity levels. This means that the students produced T-units in similar lengths under these two different task conditions. In other words, manipulation of planning time did not have any impact on overall complexity of students` narratives.

4.3.2 Subordination ratio

The descriptive statistics for subordination indicates that the participants produced similar number of subordinated clauses under both task conditions. The mean for both simple and complex tasks was nearly equivalent; NP (*M* = 1.29) and PTP (*M* = 1.26). Parallel to that 95% confidence intervals show the similarity between the scores of participants under both conditions (see Figure 7).

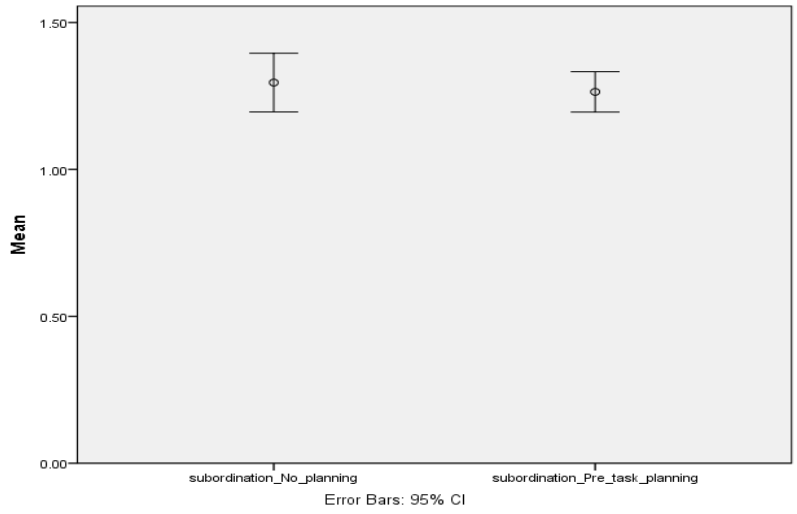


Figure 7. Subordination scores in two tasks with 95% confidence intervals

For complexity by subordination ratio measure, again there was no significant difference between the pre-task planning (*Mdn* = 1.25) and no-planning conditions (*Mdn* = 1.27), *T* = 9, *p* >.05, *r* = -.12 . Morover, the effect size is just over .10 which means the relationship between the variables is very weak. This result proves us that neither of the task planning conditions causes any significant change in the number of subordinated sentences in the narrative writing.

To sum up, manipulating tasks along planning conditions did not result in any statistically significant difference for syntactic complexity levels of participants` narratives. Therefore, the null hypothesis could not be rejected for hypothesis one in terms of syntactic complexity.

4. 4 Fluency measures

Fluency for the written tasks was measured by the mean number of words and syllables per minute. The number of words was counted with the help of Microsoft Office Word program and the syllables were calculated by the researcher with the help of online Cambridge Learners` Dictionary. After having total number of words and syllables, they were divided into 13 which is the minute spent for writing the tasks. The following table (Table 9) presents the means, standard deviations, minimum and maximum values of these measures.

Table 9. Descriptive Statistics for Fluency Measures

	No-planning				Pre-task planning			
	Mean	SD	Min.	Max.	Mean	SD	Min.	Max.
Words per minute	8.18	1.43	6.23	10.92	8.92	1.36	6.76	12.30
Syllables per minute	10.31	2.09	7.53	14.69	11.12	1.86	8.07	15.61

4.4.1 Mean number of words per minute

According to the descriptive statistics in Table 9, participants produced more words under PTP condition. In NP condition, the mean number of words per minute was 8.18 (SD = 1.43) while in PTP condition it was 8.92 (SD = 1.36). When the minimum values were checked, there is not any big difference. However, regarding the maximum values, the gap was wider. For affirming the result of descriptive statistics visually, the error bars checked (see Figure 8) and it was found out that they overlap slightly less than half. This can be a sign for significance for Hollands & Jarmasz (2010). However, the test statistics should be also checked to confirm this result.

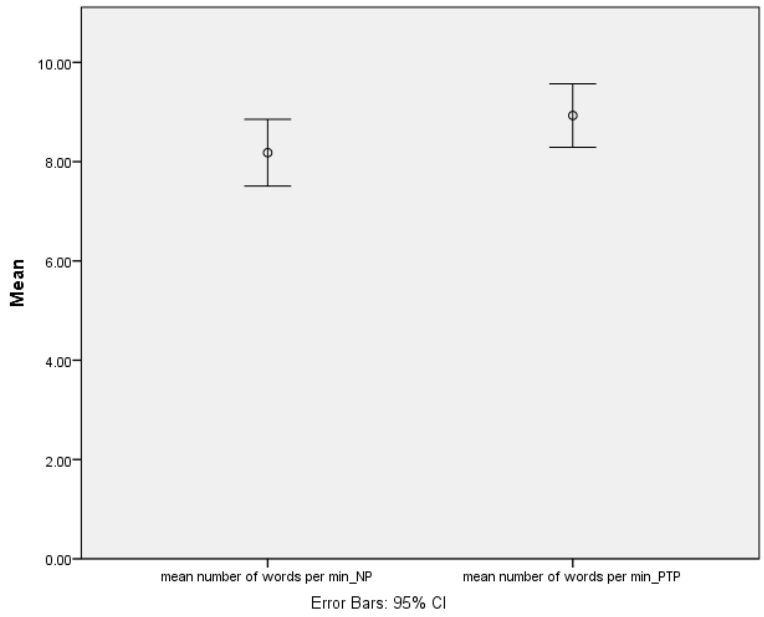


Figure 8. 95% Confidence intervals for words per minute measure

When the test statistics was checked by Wilcoxon signed-rank test for words per minute measurement, non-significant difference was found between NP (*Mdn* = 8.15) and PTP (*Mdn* = 8.57), T = 6, $p > .05$, $r = -.31$. This result shows us that the number of words written per minute under both task conditions are not statistically different. When we check the effect size ($r = -.31$), it points out a medium effect of task condition on words per minute measurement with gains under PTP condition.

4.4.2 Mean number of syllables per minute

Regarding the mean number of syllables per minute measurement for fluency, as shown in Table 9, participants produced similar number of syllables during writing time in NP condition (*M* = 10.31, *SD* = 2.09) and in PTP condition (*M* = 11.12, *SD* = 1.86).

Similarly, the minimum and maximum values were quite similar with less than 1 point. The confidence intervals (see Figure 9) affirm this similarity as the confidence intervals overlap by just half of the error bars which means the difference is non-significant (Hollands & Jarmansz, 2010).

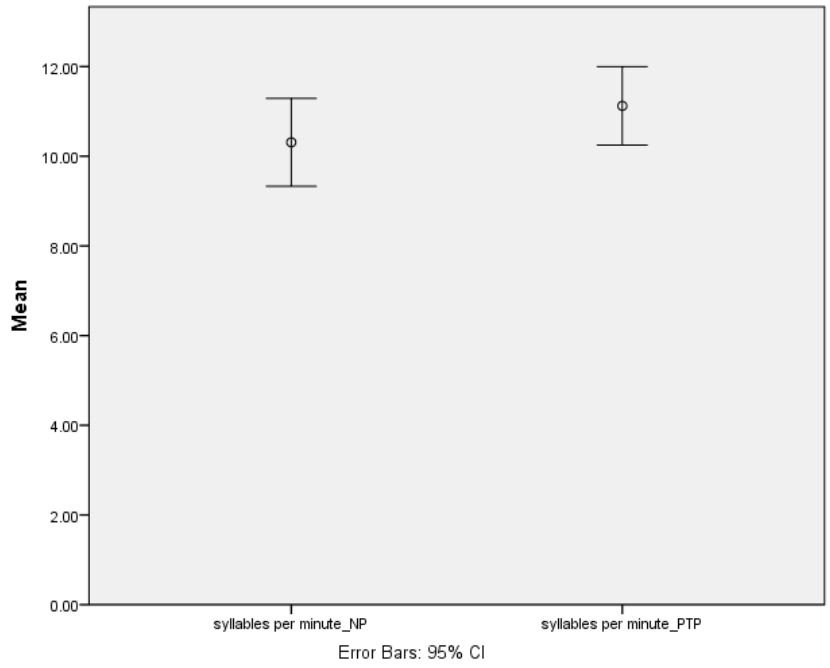


Figure 9. 95% Confidence intervals for syllables per minute measure

The results of Wilcoxon Signed-rank test affirm the findings of descriptive statistics. The test did not indicate any significant difference between NP (*Mdn* = 10.07) and PTP (*Mdn* = 10.80), $T = 7$, $p > .05$, $r = -.26$. This means that the participants produced similar number of syllables under both task conditions. Additionally, it should be noted that the effect size ($r = -.26$) is close to medium level for this fleuncy measure as well.

To sum up, the descriptive statistics and the non-parametric test results show that providing students pre-task planning time does not affect writing fluency. Therefore, the hypothesis one was not affirmed for fluency measure either.

4.5 Writing anxiety

The descriptive statistics table below (Table 10) proposes that there is a wide range in the survey scores with minimum two and maximum 45. It shows that the group of participants is not homogenous in terms of writing anxiety levels. Table 11 presents the correlations between writing anxiety levels of participants and their CAF levels. The correlations presented are calculated for only the task conducted under NP condition, so that the results are comparable with the previous studies.

Table 10. Descriptive Statistics for Writing Anxiety Survey

	Mean	SD	Min.	Max.
Writing Anxiety Survey	20.20	10.2	2	45

Table 11. Spearman`s Correlation Coefficient Results for Writing Anxiety and CAF for NP condition

	Complexity	Accuracy	Fluency
Writing Anxiety	.11	.40*	-.22

Note: *shows a significant correlation between variables

4.5.1 The relationship between writing anxiety and syntactic complexity

The Spearman`s Correlation Coefficient results for writing anxiety and mean length of T-unit revealed no significant relationship. The results were $r_s = .11$, $p > .05$ for NP condition. It can be concluded that writing anxiety levels of participants do not affect the syntactic complexity of their written productions.

4.5.2 The relationship between writing anxiety and accuracy

The Spearman`s Correlation Coefficient was conducted to find out the relationship between writing apprehension and errors per T-unit measurement under NP task condition. The results for NP ($r_s = .40$, $p < .05$) revealed a significant positive correlation between the two. This means that the participant with high level of writing anxiety tend to have more errors per T-unit while participants with lower anxiety tend to have less errors.

4.5.3 The relationship between writing anxiety and fluency

The Spearman`s Correlation Coefficient calculations show that fluency measure (i.e. the number of words per minute) in NP condition has no significant relationship with writing apprehension ($r_s = -22$, $p > .05$). Although the relationship is not significant its direction is negative so it could be stated that there is a tendency that fluency measure might be negatively affected by higher writing apprehension levels of students.

To summarize, writing apprehension levels correlated significantly with the accuracy measure (i.e. errors per T-unit), but there no correlation was calculated between

writing apprehension and the general syntactic complexity measure (i.e. mean length of T-units) and the fluency measure (i.e. the number of words per minute). As a result, the hypothesis two was confirmed only for accuracy but not for complexity and fluency.

CHAPTER 5

DISCUSSION

5.1 Introduction

The present chapter discusses the findings of the written text analysis presented in Chapter 4. Turkish university preparatory class students produced the written narratives under two task conditions (i.e. pre-task planning and no-planning). The findings from the analysis are interpreted in the light of the Cognition Hypothesis and Limited Attentional Capacity Hypothesis, along with previous research and a writing model. In this chapter the following hypotheses will be discussed:

Hypothesis 1. An increase in accuracy and/or syntactic complexity levels of written language performance is expected under PTP condition. However, fluency levels are expected to increase under NP condition (Skehan, 1996, 1998a; Robinson, 2001a, 2011a).

Hypothesis 2. The researcher expects a negative correlation between writing anxiety and complexity as well as fluency under NP condition. However, positive correlations between writing anxiety and accuracy are expected under NP condition because the accuracy measure used in this research is errors per T-unit which is the opposite of error free T-units measure used in the previous studies (Zabihi, 2017; Saeedi & Farnia, 2017; Haddad, 2018).

The discussion is organized according to the hypotheses above. The analysis of complexity, accuracy, and fluency are sometimes difficult to separate and discuss in isolation, so there will be occurrences of overlapping during the discussion.

5.2 Relationship between task condition and L2 written performance

The written performances of participants were assessed with syntactic complexity, accuracy, and fluency measures. In the study, the participants` accuracy levels showed a significant increase when they were given pre-task planning time. Therefore, the results for accuracy are in line with previous studies (Foster & Skehan,1996 – with unguided planning conditions only; Ellis & Yuan, 2004; Mochizuki & Ortega, 2008; Tavares, 2009). However, this result is not compatible with the literature for the spoken language production which revealed mixed results for accuracy. There can be two reasons of improvement in accuracy level under the simpler task condition. First of all, Ellis (2005) states that pre-task planning lessen the burden on learners` working memory which allows learners to attend to form while they are conveying the message (pg.4). Therefore, the learners were able to pay attention to accurate use of language in the PTP condition. Furthermore, the accuracy results are also compatible with the Cognition Hypothesis (Robinson, 2001a, 2011). Robinson (2001a) explains that the increase in accuracy levels when learners are given pre-task planning time can be partially due to learners` behavior in the planning phase. He posits that during planning time, learners only dip into the tried and trusted language forms to use them in the task. Consequently, availability of planning time can result in more accurate language production (Robinson, 2001a). Besides that, the Trade-off Hypothesis (Skehan, 1998) can explain the findings of accuracy in the study. The hypothesis presumes that a difficult task, i.e. no-planning condition for this study, requires learners to pay more attention to meaning rather than form of the language they use. Therefore, learners can produce less accurate and less complex language under NP condition. However, when the task condition (i.e. extra

planning time) lessens the cognitive burden, they can be either more accurate or more complex.

As for syntactic complexity findings, this study`s results are similar to Foster and Skehan (1996) -with unguided planning condition results only-, Tavares (2009), Tabari (2016) and Johnson, Mercado and Acevedo (2012). They also present no effect of pre-task planning on complexity. Hence, the complexity result cannot affirm the Cognition Hypothesis (Robinson, 2001) which expects less complex language production under NP condition, but can affirm the Trade-off Hypothesis (Skehan, 1998) because the hypothesis predicts either an increase in accuracy or complexity so the preference of participants in this study was for accuracy but not for complexity. Moreover, the previous studies on speaking and planning showed contrastive results regarding the complexity results. There can be four reasons for the non-significant differences between two task conditions in the current study. Firstly, Ortega (2005) claims that some learners can be oriented towards form, while others towards meaning. Therefore, it can be stated that learners' personal tendencies play a role in what aspects of language i.e. fluency, accuracy or complexity will be prioritized during the production. Some learners may be more willing to take risks and try complex forms, while others may prioritize error-free performance. It can be that the participants in the current study are more conservative in terms of their language use and prioritize accuracy by forsaking complexity. Also, Skehan and Foster (2001) state that a learner may not allocate his attention to using syntactically complex structures for being on the safe side in terms of accuracy and vice versa. Secondly, the non-significant results could be due to the syntactic complexity measures used in the current study. According to a theoretical model of language

development (Halliday and Mathiessen, 1999), learners` syntactic complexity development proceeds from: (i) first by means of mostly coordination or the sequencing of self-standing words, sentences, and clauses; through (ii) an expansion of expression via subordination, and finally (iii) nominalization and more complex phrases. That is, the syntactic complexity should be assessed through different grammatical units according to learners` proficiency levels. Norris and Ortega (2009) propose that for beginner level learners, the coordination index can have a more predictive power instead of subordination index. However, as I reviewed the literature, I could not find any task complexity studies using coordination index so I decided not to use it in the current study. The subordination ratio used in the study could be more appropriate for higher levels of proficiency and could have failed to capture the syntactic complexity of the participants with low level proficiency in this study. Furthermore, the genre of narrative writing could play a role in the insignificant results. Pallotti (2009) states "It is more logical to expect that a task involving counterfactual reasoning stimulates the productions of subjunctives and conditionals more than a simple narration." Another reason of lower complexity levels can be the nature of the task. According to Pallotti (2009) linguistic complexity increases if it is specifically required by the task and if it is one of the aims of the task. The narrative task used in this study is a pre-organized one. That is, it does not require the participants to order the pictures as they were not given in mixed order, so the task might not have triggered the learners to use more syntactically complex structures as it was not cognitively demanding that much. In other words, the less demanding nature of telling a story based upon the pre-organized pictures may have resulted in less complex language production.

The fluency results are compatible with Foster and Skehan (1996), Mochizuki and Ortega (2008), Ong and Zhang (2010), Johnson, Mercado and Acevedo (2012) which also found no effect of planning on fluency measurements. However, although there was no significant difference, there was a tendency toward PTP condition so this tendency shares similarity with most of previous studies on the written and spoken task planning studies. In the literature, the positive effect of pre-task planning time on fluency was measured by the majority of task complexity studies. On the other hand, following hypothesis one, an increase in fluency under NP condition was expected because Robinson (2011) states that increasing a task`s complexity through resource-dispersing dimensions (i.e. by not giving pre-task planning time) will promote faster access to existing interlanguage resources and vice versa. This might not have happened due to the low proficiency level of students. Following Kellogg's (1996) model of writing, it can be inferred that the translating process of writing may overburden the working memory resource of low proficiency learners (Johnson, et. al., 2012). Therefore, little working memory capacity could be left to the execution of writing and in accordance to that fluency levels did not increase under NP condition as expected by Robinson (2011) and Skehan (1998). Furthermore, the construct of fluency and the measures used can be other causes of this finding. Fluency is a controversial construct for writing because, unlike speaking, the writing process does not occur in real-time (Michel, 2017; pg.13). Due to its nature, it allows planning, executing and monitoring happen recursively (Kellogg, 1996; Johnson, et. al, 2012). That is, planning and text production can occur at the same time while writing (Hayes & Nash, 1996). Actually, defining fluency as the number of words per minute would be appropriate only if all participants wrote for the full 13

minutes non-stop to complete the task. Therefore, it is quite hard to measure the real fluency in the written language production as we cannot measure the dysfluencies.

For explaining the findings of this study regarding fluency and syntactic complexity, Kellogg`s (1996) writing model could be useful. In the model Kellogg proposes that working memory resources are limited, so writers need to prioritize one process (i.e. formulation, execution, and monitoring) over the other. According to De Larios, et al. (2001) low level L2 writers need to concentrate more on the translation/formulation of sentences at the expense of execution and monitoring. As a result, they spend more time on thinking about how to state their ideas in L2 rather than writing them. The planning time provides little relief from the cognitive demands of the process of writing due to their inadequate level of proficiency. Actually, in the preliminary analysis of pre-task planning sketches of students, it came to the researcher`s attention that some students wrote the story in Turkish first and then tried to translate it into English during the given pre-task planning time. Hence, it is highly possible that these lower proficiency level students still use their L1 to conceptualize the story. Then they translate it into English which causes a huge burden on their working memory. Moreover, in terms of task difficulty, narrative writing tasks are considered as challenging tasks for lower level English learners (Foster & Skehan, 1996). That is, due to the limited capacity of working memory, they cannot concentrate on complexity or fluency while producing the narrative writing under NP condition.

These findings are partially in line with the Robinson`s Cognition Hypothesis (2001a, 2011) which posits that the less resource-depleting task condition, in this study PTP one, results in more accurate and complex written production whereas more

resource-depleting one results in decrease in these two dimensions but increase in the fluency dimension. Robinson (2011) claims that fluency is not expected to deteriorate when task complexity is increased along resource-dispersing dimensions (i.e. absence of planning time). He explains that in the absence of planning time, learners can only access to their existing interlanguage resources which must be in real time and fast so their fluency increases while their accuracy and complexity levels decrease (Robinson, 2011, pg. 17). However, his predictions in terms of complexity and fluency were not confirmed through this study`s results as no difference was found for complexity and fluency measures between the task conditions. On the contrary, a decrease in both measures under NP condition, complex condition, was observed although it did not reach significant levels. On the other hand, regarding accuracy, his claim was acknowledged by the findings of this study because accuracy levels dropped under NP condition, more language-dispersing condition.

Similarly, the results support partially the claims of Limited Attentional Capacity Hypothesis (Skehan, 1998). The hypothesis presumes that a difficult task, i.e. no-planning condition for this study, requires learners to pay more attention to meaning rather than form of the language they use. Therefore, learners should produce both less accurate and less syntactically complex but more fluent language in difficult tasks. However, in the study, this was not affirmed because under NP condition, there was a decrease in all dimensions (i.e. CAF). Although this decrease was not significant for fluency and complexity, it was significant for accuracy. As a result, Limited Attentional Capacity Hypothesis was only confirmed for accuracy dimension of written language production.

In a nutshell, this study`s results partially affirm the predictions of the Limited Attentional Capacity (Skehan, 1998, 2009) and the Cognition Hypothesis (Robinson 2001a, 2011). Decreasing task complexity by giving extra planning time before a narrative writing task resulted in more accurate language production.

5.3 The relationship between writing anxiety and CAF

The Hypothesis two was only confirmed for accuracy measure but not for complexity and fluency. For accuracy, as errors per T-unit analysis used, it correlated with WA positively which means as errors increase WA increases. However, in the literature error-free T-unit measure was used and had negative correlation with WA, which means less anxiety results in fewer errors. For complexity and fluency, unlike other studies (Zabihi, 2017; Saeedi and Farnia, 2017; Haddad, 2018), no relationship was found in NP condition.

The study of Zabihi (2017) and Saeedi and Farnia (2017) which investigated the effect of writing anxiety on CAF presented some different results from the study. They have found that writing anxiety is negatively correlated with all measures of L2 production in a significant level. On the other hand, the current study has presented only negative correlation- although seems positive- with accuracy measure. Unlike Haddad (2018) which presented a significant negative correlation with number of subordinate clauses produced, this study failed to show this negative correlation. The differences between the results could be due to the differences between participants` proficiency levels. The current study investigated the WA and CAF levels of beginner and elementary level EFL learners whereas the participants in those studies had upper-intermediate level of English. Actually, this result is not unexpected and shows similarity

with accuracy results in relation to planning time. In the analysis of the effect of PTP time on CAF dimension, the only measure that revealed a significant improvement under PTP condition was accuracy. Therefore, it can be inferred that the students with lower level of proficiency takes `safety first approach` (Skehan & Foster, 2001) and they feel anxious about being correct. Moreover, this tendency could also be related to the educational and socio-cultural context of L2 writing instruction instead of being a cognitive approach to the task of writing. Hyland (2011) states that learning to write is also a social practice pertaining to cultural and institutional contexts in which the written work is produced. Therefore, the task of writing has a cultural purpose which is specific to the present community. In the study conducted by Altınmakas and Bayyurt (2019), an association is revealed between the institutional context regarding the education system in Turkey and undergraduate students` writing practices. One of the results of this qualitative study was that Turkish language teachers seem to value grammatical accuracy and writing format more than development of ideas in L1 writing instruction in Turkish secondary schools. Parallel to that, L2 writing teachers are also interested more in the correct use of taught grammatical structures and vocabulary rather than the content, fluency and organization of ideas. Considering that, observing an increase in accuracy levels under PTP condition and the positive correlation between WA levels and the number of errors are not unexpected results in Turkish educational context.

CHAPTER 6

CONCLUSION

6.1 Summary and conclusions

The main goal of this study was to examine the effects of planning conditions on lower level proficiency L2 learners` written accuracy, syntactic complexity and fluency levels in a narrative writing task. Most of the previous studies on +/- pre-task planning time investigated the effects of availability of pre-task planning time on written language production with highly proficient learners. However, this study aimed to find out whether the reported effects of pre-task planning time on written production in the previous studies can be valid for low level English language learners. Hence, the present study was conducted with beginner and elementary level participants. In addition to that, the design of the study was distinct from the majority of the studies conducted in task complexity research. Although the effects of pre-task planning on written language production have been studied previously in relation to task complexity (Ellis & Yuan, 2004; Ong & Zhang, 2010; Rahmpour & Safari, 2011; Ghavamnia, Tavakoli & Esteki, 2013; Tabari, 2016) very few, if any, studies investigated it with a repeated-measures design (with-in subject design). Following Field (2009) with-in subject design has relatively more statistical power than between groups design since it diminishes the effect of unsystematic variance (i.e. individual differences) and enables researcher to detect any systematic variance more easily (pg. 342). Also, this study is one of the very few studies which attempted to examine the effect of an individual difference variable (i.e. writing eanxiety) in relation to CAF measures.

The findings of the study has revealed that English learners with lower level proficiency produce more accurate narrative writings when they were given 10 minute pre-task planning time compared to when they were not given that extra planning time. However, this result was not compatible with most of the task complexity studies` (Ghavamnia, Tavakoli & Esteki, 2013; Abrams & Byrd, 2016; Ellis and Yuan, 2004; Ortega, 1999; Tavakoli & Skehan, 2005; Yuan & Ellis, 2003) since they found mixed results for accuracy measure. Regarding the syntactic complexity and fluency, the participants of the study did not show any significant difference between the task conditions. That is, their syntactic complexity and fluency levels were similar under both task conditions, although the differences in fluency measures were close to reach a significant level. This finding was contrary to the most of the task complexity research manipulating complexity through pre-task availability. Lastly, these results partially confirmed the predictions of neither Robinson`s Cognition Hypothesis nor Skehan`s Limited Attentional Capacity Hypothesis.

Two pedagogical implications can be drawn from the findings of this study. Firstly, providing lower level students with unguided pre-task planning time can lead learners to pay more attention to grammatical accuracy in a narrative writing task. Therefore, teachers can implement the writing tasks with pre-task planning time to increase the accuracy levels of students with limited proficiency. Secondly, writing anxiety can be a hindrance during language performance for students with lower level English especially when they want to produce error-free writings. Therefore, language teachers should be aware of this and try to alleviate writing anxiety levels of their students by using some educational techniques such as peer feedback (Rollinson, 2005).

Last but not least, a theoretical implication could be deduced from the findings of this study. The task complexity variables presented in theoretical task sequencing frameworks (i.e. Triadic Componential Framework) may not be sensitive enough to observe their effects on language learners with lower level pfociency in L2. Therefore, proficiency could be added as a mediating variable into these frameworks and be taken into consideration while designing a task-based syllabus and while categorizing the tasks. For instance, to improve accuracy in initial levels of proficiency, the tasks could be designed with pre-task planning time.

6.2 Limitations and suggestions for future research

It is worth noting that findings of the current study should be interpreted within certain limitations. One of the limitations is the quantitative design of the study. No qualitative method was utilized to investigate participants` perceptions and feelings about the task conditions (i.e. allocated planning and writing time) or the two picture sets used, the Picnic and the Chase. Skehan and Foster (2001) state that how planning time is used by the learner depends heavily on the individual but not on the task conditions, and they distinguish three uses of pre-task planning time: for rehearsing the task, for recall the relevant language, or for engaging with the content of the task. According to how students use the planning time, it can be inferred to which language dimension they pay most of their attention. For example, if planning time is used for rehearsing the task, increase in accuracy dimension can be expected (Skehan & Foster, 2001). Therefore, future studies can employ more qualitative methods as well as quantitative ones to investigate learner perceptions and preferences about the planning time and the task type.

Further retrospective interviews with participants about how they take advantage of the planning time (i.e. if they pay more attention to grammar, vocabulary, or the organization of the narrative) or how they perceive the difficulty level of the two stories can provide further evidence to interpret the findings of this study.

Second limitation can be the context of the study. The participants (*N*= 20) are adult Turkish beginner and elementary level learners of English so the results cannot be generalized for another group of learners. Therefore, the future studies can be conducted with participants who have different L1s. Also, the same research design could be conducted with intermediate and advanced level students to observe any differences in their CAF levels under two planning conditions. This can be useful to investigate learners` language production preferences regarding to their language level. For example, higher level students may prefer to use more elaborated language.

Third limitation could be related to the validity of measures chosen to evaluate L2 writing performance. For example, participants` syntactic complexity was measured in terms of subordination which is a powerful index of complexity at intermediate and upper-intermediate levels (Norris & Ortega, 2009). However, the participants of this study are at beginner and elementary level. Hence, this measure might not be sensitive enough to investigate complexity of their narrative writings. Moreover, most of the validity and reliability of CAF measures are unknown because most of the research does not report on that issue. Hence, the future research should zero in on the standardization of these metrics (Michel, 2017, pg. 26)

The genre type and learners` having no previous practice with this genre type could be another limitation of the current study. As the participants were novice in

writing a story, they may have had difficulties in organizing the content or finding the right linkers to relate two consequitive pictures.

Another point to consider would be the application of the narrative tasks along with other tests and the questionnaire in 90 minutes. The writing anxiety test and the placement test were applied after or before the tasks, so this could have increased the cognitive load and participants became tired towards end of the data collection. Therefore, this could have affected the results of the study negatively. Moreover, pushing the participants to produce 100 words in just 13 minutes could have increased their anxiety, and therefore, it could have affected the results as well.

All in all, the future research is required to address the proficiency issue along with reliability and validity of linguistic measures issue. The design of study should be repeated with more advanced learners to see the effects of task conditions on CAF dimensions more clearly. Additionally, the future research should focus on the realiability and validity of CAF measures in order to make the findings of different studies comparable so that general trends of the relationship between CAF dimensions and task complexity research could be identified. Furthermore, more qualitative methods should be utilized in task complexity research to acknowledge students` perceptions and feeling about task conditions, task complexity, and the tasks used so that the teachers can be more aware of students` feelings while implementing TBLT syllabuses. Also, the role of IDs such as motivation, anxiety, self-efficacy should be investigated more in relation to task complexity since students` characteristics contribute to how they perceive the task complexity and have an implact on what kind of strategies they use.

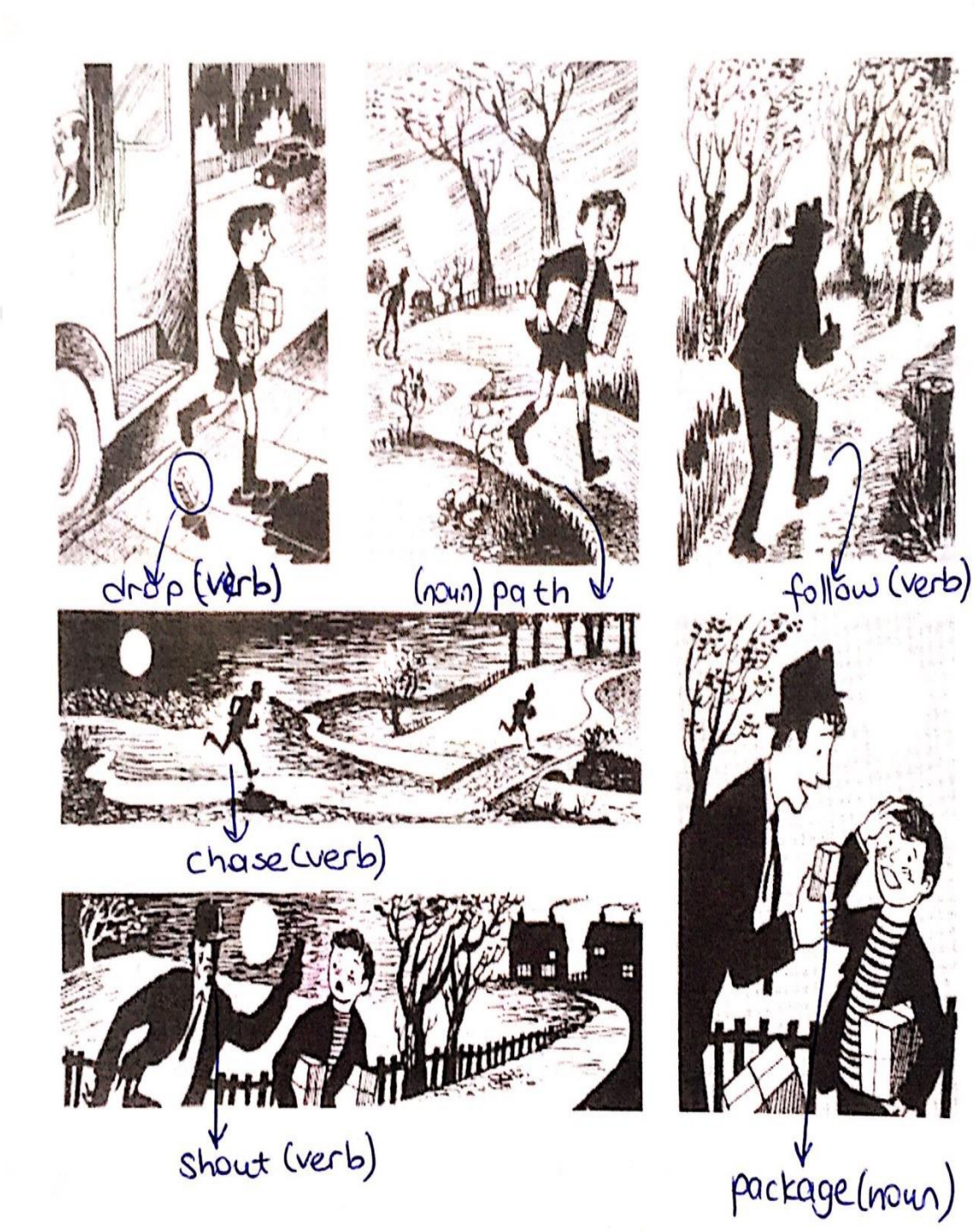
APPENDIX A

PICTURE PROMPTS USED IN THE TASKS

Picture 1: The Picnic

Source: [Heaton, 1975]

Picture 2: The Chase



Source: [Heaton, 1975]

APPENDIX B

BACKGROUND QUESTIONNAIRE

1. İsim & Soyisim: _____
   (Name & Surname)

2. Yaş *(Age)*: _____
   Cinsiyet *(Sex):* ☐ KIZ*(female)* ☐ ERKEK *(male)*

3. Ne zamandır İngilizce öğreniyorsunuz?
   (How long have you been learning English?)
   _____

4. İngilizceyi nere(ler)de öğrediniz?
   (In which institutions have you learnt English?)
   _____

5. Anadiliniz(leriniz) nedir?
   (What is/are your mother tongue(s)?)
_____

6. Hiç yurtdışında yaşadınız mı? Yaşadıysanız ne kadar ve nerede?
(Have you ever lived abroad? If yes, how long and where?)
_____

7. Daha önce İngilizce bir hikaye yazdınız mı?
(Have you ever written a story in English?)
_____

APPENDIX C

WRITING ANXIETY INVENTORY (TURKISH)

Bu anket, katılım formunda belirtilen master çalışmasının bir bölümüdür. Bu anket, bir devlet üniversitesinde hazırlık bölümünde okuyan öğrencilerin İngilizce yazma kaygısı hakkındaki düşüncelerini ve inançlarını daha iyi anlamayı amaçlamaktadır.

Bu anket bir test değildir, dolayısıyla "doğru" veya "yanlış" cevaplar yoktur. Bir yabancı dil öğrencisi olarak sadece kişisel görüşünüzle ilgileniyoruz. Anketin sonuçları yalnızca araştırma amaçları için kullanılacaktır ve cevapladığınız anket kimse ile paylaşılmayacaktır. Bu sebeple cevaplarınızı içtenlikle belirtmeniz bu projenin başarısına katkı sağlayacaktır.

Anketi cevaplandırmayı istediğiniz zaman ve herhangi bir neden belirtmek zorunda kalmadan yarıda bırakabilirsiniz ve bıraktığınız takdirde herhangi bir olumsuzluk ile karşılaşmayacağınızı araştırmacı taahüt eder.

Yardımınız için çok teşekkürler!

Irem Tümer Görgülü

Katılımcının adı ve soyadı:_____

SORULAR

Soruları yanıtlarken aşağıdaki ifadelerine kadar katıldığınızı veya katılmadığınızı 1'den 5 kadar olan numaraları yuvarlak içine alarak ifade ediniz. Yanıtlanmamış soru bırakmayınız.

Kesinlikle Katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
1	2	3	4	5

1. Bir konu hakkında süre verilerek İngilizce yazmam istendiğinde fikirlerimi toplamakta zorluk çekerim.	1	2	3	4	5
2. Bir konu hakkında süre verilerek İngilizce yazmam istendiğinde genellikle panik yaparım.	1	2	3	4	5
3. Bir konu hakkında süre verilerek İngilizce yazmam istendiğinde titrerim ya da terlerim.	1	2	3	4	5
4. Bir konu hakkında süre verilerek İngilizce yazmam istendiğinde kalp atışım hızlanır.	1	2	3	4	5
5. Bir konu hakkında İngilizce yazarken bütün vücudumun gerildiğini hissederim.	1	2	3	4	5

6. Beklemediğim bir anda hoca İngilizce bir paragraf yazmam istediğinde donar kalırım.	1	2	3	4	5
7. İngilizce paragraf yazmaya başladığım anda genellikle beynim durur.	1	2	3	4	5
8. İngilizce paragraf yazmam istendiğinde yazmamak için elimden gelen en iyi mazeretleri bulurum.	1	2	3	4	5
9. İngilizce paragraf yazmak benim için sorun değildir.	1	2	3	4	5
10. Ders dışında İngilizce paragraf yazabileceğim her fırsatı değerlendiririm.	1	2	3	4	5
11. Genellikle düşüncelerimi İngilizce not etmeyi seçerim.	1	2	3	4	5
12. İngilizce paragraf yazmaktan kaçmak için elimden geleni yaparım.	1	2	3	4	5
13. Bir konu hakkında fikrimi yazmam istendiğinde ancak başka seçeneğim yoksa İngilizceyi kullanırım.	1	2	3	4	5
14. İngilizce yazmamı gerektirecek durumlardan kaçınmak için elimden geleni yaparım.	1	2	3	4	5
15. Başka insanların İngilizce yazdığım yazılar hakkında ne düşündüğü benim için hiç önemli değildir.	1	2	3	4	5
16. Yazdığım İngilizce paragrafın düşük not almasından hiç korkmam.	1	2	3	4	5
17. Yazdığım İngilizce paragrafın sınıftaki diğer öğrencilerin yazdıklarından kötü olmasından endişe duymam.	1	2	3	4	5
18. Sınıftaki diğer öğrencilerin benim yazdığım İngilizce bir yazıyla dalga geçmesinden korkarım.	1	2	3	4	5
19. İngilizce yazdığım paragrafın sınıfta tartışmak için örnek olarak kullanılmasından korkarım.	1	2	3	4	5
20. İngilizce yazı yazarken hiç endişeli olmam.	1	2	3	4	5
21. Eğer İngilizce yazdığım paragraf notlandırılacaksa çok düşük not almaktan korkarım.	1	2	3	4	5
22. Eğer İngilizce yazdığım paragraf notlandırılacaksa kendimi gergin ve huzursuz hissederim.	1	2	3	4	5

Translated from: [Cheng, 2004]

APPENDIX D

WRITING ANXIETY INVENTORY (ENGLISH)

This survey is a part of the thesis mentioned in the approval form. This surveys aims to understand thoughts and beliefs of students studying English at preparatory class of a state university.

This survey is not a test, so there are no "right" or "wrong" answer. We are only interested in your personal opinion as a foreign language student. The results of the survey will be used for research purposes only and your survey will not be shared with anyone. Therefore, your sincere answers will contribute to the success of this project.

You can leave the questionnaire at any time and without having to give any reason, and the researcher undertakes that you will not experience any negativity.

Thanks for your help!

Irem Tümer Görgülü

Name & Surname of participant: _____

QUESTIONS

When answering the questions, circle the numbers from 1 to 5 whether you agree or disagree with the following statements. Do not leave any unanswered questions.

I totally disagree	I disagree	I don`t know	I agree	I totally agree
1	2	3	4	5

1. My thoughts become jumbled when I write English compositions under time constraint.	1	2	3	4	5
2. I often feel panic when I write English compositions under time constraint.	1	2	3	4	5
3. I tremble or perspire when I write English compositions under time pressure.	1	2	3	4	5
4. I feel my heart pounding when I write English compositions under time constraint.	1	2	3	4	5

5. I usually feel my whole body rigid and tense when I write English compositions.	1	2	3	4	5
6. I freeze up when unexpectedly asked to write English compositions.	1	2	3	4	5
7. My mind often goes blank when I start to work on an English composition.	1	2	3	4	5
8. I would do my best to excuse myself if asked to write English compositions.	1	2	3	4	5
9. Whenever possible, I would use English to write compositions.	1	2	3	4	5
10. I usually seek every possible chance to write English compositions outside of class.	1	2	3	4	5
11. I often choose to write down my thoughts in English.	1	2	3	4	5
12. I usually do my best to avoid writing English compositions.	1	2	3	4	5
13. Unless I have no choice, I would not use English to write compositions.	1	2	3	4	5
14. I do my best to avoid situations in which I have to write in English.	1	2	3	4	5
15. I don't worry at all about what other people would think of my English compositions.	1	2	3	4	5
16. I'm not afraid at all that my English compositions would be rated as very poor.	1	2	3	4	5
17. I don't worry that my English compositions are a lot worse than others'.	1	2	3	4	5
18. I'm afraid that the other students would deride my English composition if they read it.	1	2	3	4	5
19. I'm afraid of my English composition being chosen as a sample for discussion in class.	1	2	3	4	5
20. While writing in English, I'm not nervous at all.	1	2	3	4	5
21. If my English composition is to be evaluated, I would worry about getting a very poor grade.	1	2	3	4	5
22. While writing English compositions, I feel worried and uneasy if I know they will be evaluated.	1	2	3	4	5

Source: [Cheng, 2004]

APPENDIX E

TASK INSTRUCTIONS (TURKISH)

Plansiz kosul icin (Chase Hikayesi):

Yar. Doç. Dr. Şebnem Yalçın önderliğinde yürütülen bu çalışma yazma becerisini ölçmek üzere tasarlanmıştır.

- Size 1'den 6'ya kadar sıralanmış resimler göstereceğiz. Bu resimler John adında bir çocuğun başına gelen bir olayı anlatmaktadır. Sizden yapmanızı istediğimiz şey bu resimlerde betimlenen hikayeyi detaylı bir şekilde İngilizce olarak anlatmanız.
- Hikayeyi yazmaya başlamadan önce resimleri incelemeniz için süre verilmeyecektir.
- Lütfen hikayeyi bu resimleri hiç görmemiş olan bir arkadaşınıza anlatırmış gibi mümkün olduğu kadar ayrıntılı bir şekilde anlatın.
- Hikayeyi anlatırken geçmiş zaman (past tense) kullanınız.
- Bilmediğiniz kelimeleri hocanıza sorabilirsiniz. Verilen kelimeleri kullanmak zorunda değilsiniz.
- Hikâyeyi anlatırken yalnızca **13 dakikanız** var ve **100 kelime** yazmanız gerekiyor.

Eğer bir sorunuz varsa şimdi sormanız gerekiyor. Resimleri gördüğünüz andan itibaren herhangi bir soru sormamalı (kelimeyle ilgili olanlar haric), süreniz başladığında hikayeyi ara vermeden yazmalısınız. Eğer bir sorunuz yoksa 10 dakikalık planlama sürenizi başlatacağım.

Araştırmacı: *(kağıtları ters cevrili sekilde dağıtır)* Simdi hemen hikayeyi yazmaya başlamalısınız.

Planli kosul icin (Picnic hikayesi):

Bu çalışma Yar. Doç. Dr. Şebnem Yalçın önderliğinde yürütülen yazma becerisini ölçmek üzere tasarlanmış olan bir çalışmadır.

- Size 1'den 6'ya kadar sıralanmış resimler göstereceğiz. Bu resimler Matt ve Sue adında iki çocuğun başına gelen bir olayı anlatmaktadır. Sizden yapmanızı istediğimiz şey bu resimlerde betimlenen hikayeyi detaylı bir şekilde İngilizce olarak anlatmanız.

- Hikayeyi yazmaya başlamadan önce resimleri incelemeniz ve yazacağınız hikayenizi planlamanız için size 10 dakika süre vereceğiz. 10 dakika sonunda sizden hikayeyi yazmanızı isteyeceğiz.
- Bu 10 dakikalık sürede hikayenizi içerik, organizasyon ve dil açısından planlamalısınız.Planlama için size ayrı bir kağıt verilecektir.
- Hikayeyi yazarken aldığınız notlar sizden alınacaktır fakat resimlere bakmaya devam edebileceksiniz.
- Lütfen hikayeyi bu resimleri hiç görmemiş olan bir arkadaşınıza anlatırmış gibi mümkün olduğu kadar ayrıntılı bir şekilde anlatınız.
- Hikayeyi anlatırken geçmiş zaman (past tense) kullanın.
- Planlama zamani sirasinda bilmediğiniz kelimeler icin sözlüğe bakabilirsiniz ya da hocanıza sorabilirsiniz. Verilen kelimeleri kullanmak zorunda değilsiniz.
- Hikayeyi yazmaya başladığınız anda **13 dakikanız** var. **100 kelime** yazmanız gerekiyor.

Eğer bir sorunuz varsa şimdi sormanız gerekiyor. Resimleri gördüğünüz andan itibaren herhangi bir soru sormamalı, istendiğinde hikayeyi ara vermeden yazmalısınız. Eğer bir sorunuz yoksa sürenizi başlatacağım.

(10 dakika sonunda)

Araştırmacı: 10 dakikalık süreniz sona erdi. Şimdi hikayeyi yazmaya başlayabilirsiniz.*(Araştırmacı öğrencilerin notlarını toplar.)*

APPENDIX F

TASK INSTRUCTIONS (ENGLISH)

For No-planning condition (The Chase story):

This study, which was conducted under the leadership of Assist. Prof. Şebnem Yalçın, was designed to measure writing skills.

- We'll show you pictures ordered from 1 to 6. These pictures depict an incident that happened to a child named John. What we want you to do is to write the story in detail in English.
- You will not be given any extra time to review the pictures before you start writing the story.
- Please tell the story as thoroughly as possible as if you were telling a friend who has never seen these pictures.
- Use past tense to tell the story.
- You can ask any unknown words to your instructor. You do not have to use the given words.
- You have only 13 minutes to tell the story and need to write 100 words.
- If you have a question, you have to ask now. You should not ask any questions from the moment you see the pictures (except the ones about vocabulary), you should write the story without interruption when your period starts. If you don't have any questions, I'll start your 10-minute planning.
- Researcher: *(handing out pictures upside down*) You should start writing the story right away.

For Pre-task planning condition (The Picnic story):

- We'll show you pictures ordered from 1 to 6. These pictures depict an incident that happened to two children, Matt and Sue. What we want you to do is to tell the story in detail in English.
- We will give you 10 minutes to review the pictures and plan your story before you start writing the story. After 10 minutes, we'll ask you to write the story.
- You should plan your story in terms of content, organization and language during this 10-minute period. You will be given a separate paper for planning.
- The notes you make during this 10 minute will be collected, but you will be able to continue to look at the pictures.

- Please tell the story as thoroughly as possible as if you were telling a friend who has never seen these pictures.
- Use past tense to tell the story.
- If you do not know the words you can look up the dictionary or ask your teacher during the planning time. You do not have to use the given words.
- You have 13 minutes as soon as you start writing the story. You need to write 100 words.

If you have a question, you have to ask now. You should not ask any questions from the moment you see the pictures, you should plan and write the story without interruption. If you don't have any questions, I will start your time.

(After 10 minutes)

The researcher: Your 10-minute period has ended. Now you can start writing the story (the researcher collects the students' notes.)

APPENDIX G

EXAMPLE T-UNIT AND ERROR CODING

In a sunny afternoon, Matt and Sue decided to go to the picnic. They prepared the picnic basket. They have got a small dog. The dog was very pretty. Then, their mother opened the map and showed. Suddenly, the dog jumped in the picnic basket, but children didn`t notice it. Then, children started to go and they waved to their mum. Children went to a hill and the hill was very high. When children arrived the hill, the dog jumped out of, so the children was surprised. When children looked to in the picnic basket, they can`t see anything. Because the dog ate everything in the picnic basket. The felt surprised and sad, because they haven`t any food, and they felt hungry. Poor they! Then, the dog finded someone. (122 words)

1. In a sunny afternoon, Matt and Sue decided to go to the picnic.
2. They prepared the picnic basket.
3. They have got a small dog.
4. The dog was very pretty.
5. Then, their mother opened the map and showed.
6. Suddenly, the dog jumped in the picnic basket,
7. but children didn`t notice it.
8. Then, children started to go
9. and they waved to their mum.
10. Children went to a hill
11. and the hill was very high.
12. /When children arrived the hill/, /the dog jumped out of,/ (2 clauses)

13. so the children was surprised.
14. /When children looked to in the picnic basket/, /they can`t see anything/ /Because the dog ate everything in the picnic basket/ (3 clauses)
15. /The felt surprised and sad/, /because they haven`t any food/ (2 clauses)
16. /and they felt hungry/. Poor they!
17. Then, the dog finded someone.

(17 T-units/ 21 clauses / 9 MISTAKES)

APPENDIX H

CODING GUIDELINES FOR ACCURACY ANALYSIS

COUNT:

1. Syntactic mistakes related to:
   - ✓ a missing object of a transitive verb;
   - ✓ places of subject verb object,
   - ✓ places of connectors,
   - ✓ tense agreement within the narrative (Ss are toldu to use past tenses)
2. Morphological mistakes related to:
   - ✓ verb tense mistakes – past tense and past cont especially;
   - ✓ singular-plural mistakes;
   - ✓ subject-verb agreement mistakes
3. Lexical mistakes related to:
   - ✓ not semantic mistakes but LEXICAL FORM mistakes such as using *scared* instead of *scaring*;
   - ✓ prepositional mistakes- missing or wrong preposition use
   - ✓ collocation mistakes such as *make a picnic* instead of *have a picnic*.

DON'T COUNT:

1. Semantic mistakes related to words (Students' vocabulary was limited so I want to ignore Wrong Word mistakes. And they looked up the words in dictionary so sometimes they chose unsuitable words)
2. Punctuation mistakes
3. Spelling mistakes
4. Article mistakes

APPENDIX I

CODING GUIDELINES FOR T-UNIT AND CLAUSE ANALYSIS

T-unit is a main clause plus any subordinate clause(s) or non-clausal expression(s) that are attached to or embedded to it. It is a linguistic measurement aiming to measure the smallest word group that can be defined as a grammatical sentence regardless of the punctuation. Any complex or simple sentence would be count as one T-unit. Any compound or compound complex sentence would be consisted of two or more T-units.

- ✓ If sentence has a sentence connector such as however, therefore etc., it is counted as one T-unit.
- ✓ If a sentence has coordinating conjunctions (for, and,yet,but, nor, yet, so) it should be analyzed as one T-unit
- ✓ If a sentence has a subordinator (because, when, while, after, before, since, as, although, if, so that, as if, as soon as, in case, since, until, so that, in that) it is counted as a dependent clause.
- ✓ RELATIVE CLAUSES AND NOUN CLAUSES cannot be counted as one T-unit. They should be labelled as clause.
- ✓ Punctuation of the writer is not important when we slice sentences into T-units.

Clause is defined as a subject or a coordinate subject plus a finite verb or coordinate verbs.

For example: She entered the house and closed the door. *This is one T-unit and it has only one clause because it has coordinate subject.*

Here is an example writing sample from Hunt, 1965:

I like the movie we saw about. Moby Dick the white whale thecaptain said if you can kill

the white whale Moby Dick I will givethis gold to the one that can do it and it is worth

Sixteen dollars they tried and tried but while they were trying they killed a whale and

used the oil for the lamps they almost caught the white whale

Source: [Hunt, 1970]

B) Here the sentences are sliced up so that each T-unit is numbered and begins a new line. The words in slashes are clauses.

1. / I like the movie/ /we saw about Moby Dick the white whale/ (1 T-unit/2 clauses)

2. / the captain said/ /if you can kill the white whale Moby Dick/ /I will give this gold to the one/ /that can do it/ (1 T-unit/4 clauses)

3. and it is worth sixteen dollars (1 T-unit/1 clause)

4. /they tried and tried/ (1 T-unit/1 clause)

5. but /while they were trying/ /they killed a whale and used the oil for the lamps/ (1 T-unit/2 clauses)

6. /they almost caught the white whale/ (1 T-unit/1 clause)

Source: [Hunt, 1965]

REFERENCES

Altınmakas, D. & Bayyurt, Y. (2019). An exploratory study on factors influencing undergraduate students` academic writing practices in Turkey. *Journal of English for Academic Purposes, 37,* 88-103.

Amiri, M. & Saeedi, M. (2017). Relationship Among Writing Apprehension, Writing Self-Efficacy, and Iranian EFL learners` writing performance. *Journal of Studies in Learning and Teaching English,* 6 (2), 1-26.

Baddeley, A. D. (1986). *Working memory.* Oxford: Oxford University Press.

Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition, 11,* 17-34.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Mahwah, NJ: Erlbaum.

Berman, R. (1994). Learners' transfer of writing skills between languages. *TESL Canada Journal, 12* (l), 29-46.

Brown, G., Anderson, A., Shilcock, R., & Yule, G. (1984). *Teaching Talk: Strategies for Production and Assessment.* Cambridge: Cambridge University Press.

Brown, J. S., McDonald, J. L., Brown, T. L., Carr, T. H. (1988). Adapting to processing demands in discourse production: The case of handwriting. *Journal of Experimental Psychology: Human perception and Perfomance, 14,* 45-59.

Candlin, C. (1987). Towards task based language learning. In C. Candlin and D. Murphy (*eds.). Language Learning Tasks*. Englewood Cliffs, NJ: Prentice Hall.

Cheng, Y. S. (2004). A measure of second language writing anxiety: Scale development and preliminary validation. *Journal of Second Language Writing, 13,* 313–335.

Cheng, Y. S. , Honvitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning, 49*, 417-446.

Choi, S. (2014). Language anxiety in second language learning: Is it really a stumbling block? *Second Language Studies, 31*(2), 1-42.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* (2nd ed.). New York: Academic Press.

Cohen, J. (1992). *A power primer. Psychological Bulletin, 112*(1),155–159.

Cooper, T. C. (1976). Measuring written syntactic patterns of second language learners of German. *The Journal of educational research, 69*(5) , 176-183.

Crookes, G. (1989). Planning and interlanguage variability. *Studies in Second Language Acquisition, 11,* 367-383.

Cumming, A. (2001). The difficulty of standards, for example in L2 writing. In T. Silva & P. Matsuda (Eds.), *On second language writing* (pp. 209-229). Mahwah, NJ: Lawrence Erlbaum Associates.

Cumming, A,, Rebuffot, J. & Ledwell, M. (1989). Reading and summarizing challenging texts in first and second languages. *Reading and Writing: An Interdisciplinary Journal, 2* (2), 201 -219.

Daly, J. A. (1977). The Effects of Writing Apprehension on Message Encoding. *Journalism Qauterly, 27,* 566-572.

Daly, J. A. & Miller, M. D. (1975). Apprehension of Writing as a Predictor of Message Intensity. *The Journal of Psychology, 89*(2), 175-177.

Daly, J. (1978). Writing apprehension and writing competency. *The Journal of Educational Research, 72*(1), 10–14.

De Larios, J., Marin, J., & Murphy, L. (2001). A temporal analysis of formulation processes in L1 and L2 writing. *Language Learning, 51*, 497–538.

Dörnyei, Z. (2009). *The psychology of second language acquisition.* Oxford: Oxford University Press.

Dörnyei, Z. (2005). *The psychology of the language learner: individual differences in second language acquisition.* Mahwah, NJ: L. Erlbaum

Dörnyei, Z. & Skehan, P. (2005). Individual differences in second language learning. In C. Doughty & M. Long (eds.) *The handbook of second language acquisition*. Malden, MA: Blackwell, 589–630.

Ellis, R. (1987). Interlanguage variability in narrative discourse: style shifting in the use of the past tense. *Studies in Second Language Acquisition, 9*, 1-20.

Ellis, R. (2004). Individual differences in second language learning. In D. Alan and C. Elder (Eds.) *Handbook of Applied Lingustics*. Blackwell Publishing Ltd, Oxford, 525-551.

Ellis, R., (2005). Planning and task-based performance: Theory and research. In Ellis, R. (*Ed*.), *Planning and task performance in a second language*. John Benjamins Publishing Co: Philadelphia PA, 3–34.

Ellis, R., (2018). *Reflections on task-based language teaching*. Bristol: Multilingual Matters

Ellis, R. & Barkhuizen, G. (2005). *Analyzing learner language.* Oxford: Oxford University Press.

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.

Field, A. (2009). *Discovering statistic using SPSS*. London: SAGE Publications.

Flower, L. & Hayes, J. R. (1980a). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication, 31*(1), 21-32.

Flower, L., & Hayes, J. R. (1980b). The dynamics of composing: Making plans and juggling constraints. In L. Gregg & E. Steinberg (*Eds.*), *Cognitive processes in* writing. Hillsdale, NJ: L. Erlbaum, 31-50.

Flower, L. & Hayes, J. R. (1981a). A Cognitive Process Theory of Writing. *College Composition and Communication 32* (4), 365-87.

Foster, P., and Skehan, P. (1996). The influence of planning on Performance in task-based learning. *Studies in Second Language Acqusition, 18*(3), 299-324.

Gaies, S. J. (1980). T-Unit Analysis in Second Language Research: Applications, Problems and Limitations. *TESOL Quarterly,14*(1), 53-60.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation.* London: Edward Arnold.

Gardner, R.C., Clement, R., Smythe, P. C. & Smythe, C. C. (1979). *Attitudes and motivation test battery.* Research Bulletin: 10th Issue. Ontario. Department of Psychology, University of Western Ontario.

Ghavamnia, M., Tavakoli, M., & Esteki, M. (2013). The effect of pre-task and online planning conditions on complexity, accuracy, and fluency on EFL learners' written production. *Porta Linguarum, 20,* 31-43.

Gilabert, R., Baron, J. & Levkina, M. (2011). Manipulating task complexity across task types and modes. In: P. Robinson (*Ed*.), *Second language task complexity: researching the cognition hypothesis of language learning and performance.* John Benjamins Publishing Co, Philadelphia PA, 105-138.

Grabe, W., & Kaplan, W. (1996). *Theory and practice of writing: An applied linguistic perspective.* London: Longman.

Haddad, M. T. A. (2018). Writing anxiety and the use of subordinate clauses in the written compositions of Jordanian EFL undergraduate students. *Theory and Practice in Language Studies, 8*(4), 404-409.

Halliday, M. A. K. and Matthiessen, C. (1999). *Construing experience through meaning: A language-based approach to cognition.* Norfolk: Biddles Ltd.

Hayes, J. R. (1989). *The complete problem solver* (2nd ed.) Hillsdale, NJ: Lawrence Erlbaum Associates.

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Hayes, J. R. & Flower, L. S. (1980). Identifying the organization of writing processes. In L. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-30). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hayes, J. R. & Nash, J. G. (1996). On the nature of planning in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 29-55). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Heaton, J. (1975). *Beginning composition through pictures.* London: Longman.

Holland, J. G. & Jarmansz, J. (2010). Revisiting confidence intervals for repeated measures designs. *Psychonomic Bulletin & Review 2010, 17* (1), 135-138.

Horwitz, E. K., Horwitz, M. B., & Cope, J. A. (1986). Foreign language classroom anxiety. *Modern Language Journal, 70*, 125–132.

Housen, A, & Kuiken, F. (2009). Complexity, Accuracy, and Fluency in Second Language Acquisition. *Applied Linguistics*, *30*(4), 461–473.

Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research. In A. Housen, F. Kuiken, & I. Vedder (Eds.). *Dimensions of L2 performance and proficiency: complexity, accuracy and fluency in SLA* (pp. 1–20). Amsterdam: John Benjamins.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels.* Urbana, IL: The National Council of Teachers of English. (Research report).

Hunt, K. W. (1970). Syntactic maturity in school children and adults. *Monographs of the Society for Research in Child Development, 35*(1), 1-67.

Hyland, K. (2011). Learning to write: Issues in theory, research, and pedagogy. In R. M. Manchon *(Ed.), Learning-to-write and writing-to-learn in an additional language* (pp. 17-35). Amsterdam: John Benjamins Publishing.

Ishikawa, T. (2006). The effect of manipulating task complexity along the (_Here-and_Now) dimension on L2 written narrative discourse. In C. M.Garcia-Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 136–156). Clevedon, UK: Multilingual Matters.

Iwashita, N., A. Brown, T. McNamara, and S. O'Hagan. 2008. 'Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics 29*, 24–49.

Jackson, D. O., & Suethanapornkul, S. (2013). The Cognition Hypothesis: A synthesis and metaanalysis of research on second language task complexity. *Language Learning, 63(2),* 330–367.

Johnson, M. D. , Mercado, L. & Acevedo, A. (2012). The effect of planning sub-processes On L2 writing fluency, grammatical complexity, and lexical complexity. *Journal of Second Language Writing, 21*, 264-282.

Kellogg, R. T. (1988). Attentional Overload and Writing Performance: Effects of Rough Draft and Outline Strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition,14*(2), 355-365

Kellogg, R. T. (1990). Effectiveness of Prewriting Strategies as a Function of Task Demands. *The American Journal of Psychology, 103*(3), 327-342.

Kellogg, R. T. (1996). A model of working memory in writing. In C. Michael Levy, Sarah Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (57-71). Hillsdale, NJ, US: Lawrence Erlbaum Associates.

Kleinmann, H. H. (1977). Avoidance behavior in adult second language acquisition. *Language Learning*, *27*(1): 93–107.

Kormos, J. (2014). Differences across modalities of performance: An investigation of linguistic and discourse complexity in narrative tasks. In H. Byrnes and R. M. Manchon (*Eds.*). *Insights from and Task-based Language Learning for L2 Writing*, John Benjamins Publishing Co., Amsterdam.

Kuiken, F. & Vedder, I. (2007). Task complexity and measures of linguistic performance in L2 writing. *IRAL 45,* 261–284.

Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied linguistics, 30*(4), 579-589.

Lee, S. Y. (2005). Facilitating and inhibiting factors on EFL writing: A model testing with SEM. *Language Learning, 55,* 335–374.

Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly, 26,* 27–56.

Long, M. (2015). Second Language Acquisition and Task-Based Language Teaching. Oxford, UK: John Wiley and Sons, Inc.

MacIntyre, P. D., & Gardner, R. C. (1991b). Methods and results in the study of anxiety and language learning: A review of the literature. *Language Learning, 41*, 85–117.

Mackey, A., Philp J., Egi, T., Fujii, A. & Tatsumi T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (ed.), *Individual differences and instructed language learning*. Philadelphia, PA: John Benjamins, 181–209.

Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing, 19*, 85–104.

Martin, K., & Ellis, N. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition, 34*(3), 379-413.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition, 20,* 52–83.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition, 21,* 108–148.

Michel, M. C. (2011). Effects of task complexity and interaction on L2 performance. In: P. Robinson (*Ed.*), *Second Language Task Complexity: Researching the Cognition Hypothesis of language learning and performance*. John Benjamins Publishing Co, Philadelphia PA, pp. 141-173.

Michel, M. C., Kuiken, F., and Vedder, I. (2007). The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching*, *45*(3), 241-259.

Michel, M. (2017). Complexity, accuracy and fluency. In S. Loewen & M. Sato (*Eds.), The Routledge Handbook of Instructed Second Language Acquisition.* London: Routledge.

Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research, 12*(1), 11–37.

Monroe, J. H. (1975). Measuring and Enhancing Syntactic Fluency in French. *The French Review, 48(*6), 1023-1031.

Nunan, D. (1989). *Design Tasks for the Communicative Classroom Cambridge.* Cambridge University Press.

Norris, J. M. (2009). Task-based teaching and testing. In M. Long & C. Doughty (Eds.), *Handbook of language teaching* (pp. 578–594). Cambridge: Blackwell.

Norris, J.M., & Ortega, L. (2009). Measurement for understanding: An organic approach to investigating complexity, accuracy, and fluency in SLA. *Applied Linguistics, 30*(4), 555–578.

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*, 218–233.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *SSLA, 21*, 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.

Polio, C. 2001. 'Research methodology in second language writing research: The case of text based studies' in T. Silva and P. K. Matsuda (*Eds*): *On Second Language Writing*. Lawrence, Erlbaum.

Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics, 30*(4), 590–601.

Prabhu, N.S. (1987). *Second language pedagogy*. Oxford: Oxford University Press

Proulx, P. (1991). Anxiety in language learning: Recognition and prevention. *Canadian Journal of Native Education, 18,* 53–64.

Rahmpour, M. & Safarie, M. (2011). The effects of on-line and pre-task planning on descriptive writing of Iranian EFL learners. *International Journal of English Linguistics, 1*(2): 274-280.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning, 45*, 99–140.

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

Robinson, P. (2001b). Task complexity, cognitive resources and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (*Ed.*), *Cognition and second language instruction* (pp. 185-316). New York: Cambridge University Press.

Robinson, P. (2002). Effects of individual differences in intelligence, aptitude and working memory on adult incidental SLA: A replication and extension of Reber, Walkenfeld & Hernstadt (1991). In P. Robinson (ed.), *Individual differences and instructed language learning,* (pp. 211–266). Amsterdam: John Benjamins.

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21*(2), 45-105.

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching, 43*, 1–32.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics in Language Teaching, 45*, 193–213.

Robinson , P. (2011). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In Peter Robinson (Ed.), *Second Language Task Complexity: Researching the Cognition Hypothesis of language learning and performance,* (pp.3-37). Amsterdam, Philadelphia: John Benjamins Publishing Company.

Rollinson, P.(2005). Using peer feedback in the ESL writing class. *ELT Journal, 59*(1), 23-30.

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.

Ruiz-Funes, M. (2014). An Exploration of Task Complexity in Foreign Language Writing at the Intermediate Level and Measures of Linguistic Production. *Journal of Linguistics and Language Teaching, 5*(2), 181-206.

Saeedi, M. & Farnia, M. (2017). Iranian EFL Learners' Perceived Writing Anxiety and the Complexity, Accuracy, and Fluency of their Task-based Writing: Are they correlated? *Journal of Applied Linguistics and Language Research 4*(8), 68-78.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14*, 357-385.

Scovel, T. (1978). The effect of effect on foreign language learning: A review of the anxiety research. *Language Learning, 28*, 129–142.

Sharma, A. (1980). Syntactic maturity: Assessing writing proficiency in a second language. In R. Silverstein (Ed.), *Occasional Papers in Linguistics, No. 6*, (pp. 318-325). Carbondale: Southern Illinois University.

Skehan, P. (1992). Strategies in second language acquisition (Thames Valley University Working Papers in English Language Teaching No.1). London, England: Thames Valley University.

Skehan, P. (1996a). A Framework for the implementation of task-based instruction. *Applied Linguistics 17*(1), 38-62.

Skehan, P. (1996b). Second language acquisition research and task based instruction. In J. Willis and D. Willis (*Eds.*), *The Challenge and Change in Language Teaching*. Oxford: Heinemann.

Skehan, P. (1998a). Task-based instruction. *Annual Review of Applied Linguistics 18*, 268–86.

Skehan, P. (1998b). A rationale for task-based instruction. In Peter Skehan (*Ed*.), *A cognitive approach to language learning,* (pp. 93-120). Hong Kong: Oxford University Press.

Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P. Skehan, & M. Swain, (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 167–185). London, England: Longman.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research, 1*, 185–211.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning 49*(1): 93–120

Skehan, P. & Foster, P. (2001). Cognition and tasks. In P. Robinson (ed.), *Cognition and Second Language Instruction,* 183-205. New York NY: Cambridge University Press.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics,30*(4), 510–532

Skehan, P. (2014). Limited attentional capacity, second language performance, and task-based pedagogy. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 211-260). Amsterdam: John Benjamins.

Spielberger, C. D. (1983). *Manual for the State-Trait Anxieg Inventory STAI (Form Y).* Palo Alto, CA: Consulting Psychologists Press.

Steinberg , F.S. & Horwitz E. K. (1986). The Effect of Induced Anxiety on the Denotative and Interpretive Content of Second Language Speech. *Tesol Quaterly, 20*(1): 131-135.

Tabari, M. A. (2016). The effects of planning time on complexity, accuracy, fluency, and lexical variety in L2 descriptive writing. Asian-Pacific Journal of Second and Foreign Language Education 1(10), 1-15.

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 – 277). Amsterdam: Benjamins.

Tavares, M. G. G. (2009). The relationship among pre-task planning, working memory capacity, and L2 speech performance: A pilot study. *Linguagem & Ensino, 12*, 165–194.

Truitt, S. N. (1995). Anxiety and beliefs about language learning: A study of Korean university students learning English. Unpublished doctoral dissertation, University of Texas, Austin.

Tuzcu, A. (2018). Task Type and Online Planning as Mediating Factors in the Oral Performance of L2 Learners (Unpublished Master Thesis). Bogazici University, Istanbul, Turkey.

VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition, 72*, 287-301.

Vedder, I. (1 999). Linguistic and pragma-rhetorical characteristics or argumentative discourse in L2 and Ll. In R. van Eemeren, R. Grootendorst & J. Blair (Eds.), *Proceedings of the Fourth International Conference of The International Study of Argumentation* (pp. 813-819).Amsterdam: Society for the Study of Argumentation.

Wendel, J. (1997). Planning and second language narrative production. Unpublished doctoral dissertation., Temple University Japan, Tokyo.

Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse, *Language Testing, 14*, 85-106.

Wolfe-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: Measures of fluency, accuracy and complexity.* Honolulu, Hawai'i, University of Hawai'i at Manoa.

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