TASK TYPE AND ONLINE PLANNING AS MEDIATING FACTORS IN THE ORAL PERFORMANCE OF L2 LEARNERS

AYŞEN TUZCU

BOĞAZİÇİ UNIVERSITY

2018

TASK TYPE AND ONLINE PLANNING AS MEDIATING FACTORS IN THE ORAL PERFORMANCE OF L2 LEARNERS

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

Ayşen Tuzcu

Boğaziçi University

2018

Task Type and Online Planning

as Mediating Factors in the Oral Performance of L2 Learners

The thesis of Ayşen Tuzcu

has been approved by:

Assist. Prof. Şebnem Yalçın _____
(Thesis Advisor)

Assist. Prof. İnci Ayhan _____
(Co-Advisor)

Prof. Yasemin Bayyurt _____

Prof. Gülcan Erçetin _____

Assist. Prof. Sevdeğer Çeçen _____
(External Member)

May 2018

DECLARATION OF ORIGINALITY

I, Ayşen Tuzcu, 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 A.Tuzcu

Date 07/06/2018

ABSTRACT

Task Type and Online Planning

as Mediating Factors in the Oral Performance of L2 Learners

This study examines the separate and combined effects of manipulating task complexity along two dimensions in learners' accuracy and syntactic complexity levels and reports the results in the light of two influential hypotheses: the Cognition Hypothesis and the Limited Attentional Capacity Hypothesis. The former supports the presence of multiple attentional resources in humans, thus it predicts a concurrent increase in accuracy and syntactic complexity levels when the complexity of tasks is increased. The latter, on the other hand, argues for a limit in attentional resources and claims that paying simultaneous attention to both accuracy and syntactic complexity is not probable. In the light of these two hypotheses, this study aims to investigate how manipulating complexity along task type (Here-and-Now/There-and-Then) and online planning (Pressured/Unpressured) influences learners' syntactic complexity and accuracy levels. The study had an experimental design with sixty-four undergraduate students at a state university in Turkey. Participants completed two picture-based narrative tasks. Syntactic complexity was measured with global complexity, complexity by subordination, and phrasal complexity measures. Accuracy was measured as errors per 100 words, the ratio of error-free AS units, and the ratio of error-free clauses. The results showed that increased task complexity along task type resulted in higher accuracy levels, but it did not affect syntactic complexity scores significantly. Moreover, the availability or absence of online planning did not affect the syntactic complexity and accuracy levels.

ÖZET

Ödev Türü ve Çevrimiçi Planlama Olanaklarının Yabancı Dildeki Sözlü Performans Üzerindeki Etkileri

Bu çalışma bilişsel ödev zorluğunun İngilizceyi yabancı dil olarak öğrenmiş olan öğrencilerin konuşma performanslarındaki sözdizimsel karmaşıklığı ve sözcük doğruluğunu ne şekilde etkilediğini incelemekte ve sonuçları iki modelin varsayımlarını göz önünde bulundurarak rapor etmektedir. Bu modeller Biliş Hipotezi ve Sınırlı Dikkat Kapasitesi Modelidir. Biliş Hipotezi insanlarda çoklu dikkat kaynaklarının olduğunu savunmakta ve bu sayede bilişsel ödev zorluğunun arttırılmasının öğrencilerin performanslarındaki sözcük doğruluğunu ve sözdizimsel karmaşıklığı eş zamanlı olarak arttıracağını varsaymaktadır. Sınırlı Dikkat Kapasitesi Modeli ise insanlardaki dikkat kapasitesinin sınırlı olduğunu iddia etmekte ve bu nedenle bilişsel zorluğu arttırılmış ödevlerde dikkatin sözcük doğruluğuna ve sözdizimsel karmaşıklığa yönlendirilemeyeceğini savunmaktadır. Bu modellerden yola çıkarak bu çalışmada bilişsel ödev zorluğu ödev türü ve çevrimiçi planlama yönlerinden değiştirilmiştir. Çalışmaya Türkiye'deki bir devlet üniversitesinde okuyan altmış dört öğrenci katılmış ve bu katılımcılardan kendilerine verilen resimlerden yola çıkarak iki adet hikaye anlatmaları istenmiştir. Bulgular bilişsel ödev zorluğunun ödev türü çerçevesinde arttırılmasının katılımcıların performanslarının doğruluğunu arttırdığını ancak sözdizimsel karmaşıklık üzerinde bir etkisinin olmadığını göstermiştir. Çevrimiçi planlamanın ise katılımcıların performanslarının doğruluğu ve sözdizimsel karmaşıklığı üzerinde herhangi bir etkisi olmamıştır.

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 and the entire graduate education.

I would like to extend my gratitude to my thesis co-advisor Assist. Prof. İnci Ayhan for her invaluable suggestions. I would also like to thank my thesis committee members Prof. Yasemin Bayyurt, Prof. Gülcan Erçetin, and Assist. Prof. Sevdeğer Çeçen for their insightful comments and encouraging feedback.

I would like to offer my further thanks to The Scientific and Technological Research Council of Turkey (TÜBİTAK) for the financial support within the scope of the 2210-A National Scholarship Programme for Master's Students.

My deepest thanks and gratitude go to my dear friend Fazilet Sönmez for her invaluable help and support during the data collection and data coding processes of the thesis.

I also thank my friends who have always believed in me and supported me. I am really grateful for their precious friendship.

Above all, I would like to thank my parents, Sevcan and Vasfi Tuzcu, and my sister Derya Tuzcu Eken for their unconditional love, endless support and never-ending encouragement.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION……..…….. 1

1.1.Background to the study……..…….. 1

1.2.Aims of the study……..……. 3

1.3.Organization of the thesis……..…… 4

CHAPTER 2: LITERATURE REVIEW…….…... 5

2.1. Introduction……...……...…….…… 5

2.2. First and second language speech production models……..…..……... 5

2.3. Measures of L2 production……..…….…… 7

2.4. Tasks in TBLT……...……...…….…….. 10

2.5. Task complexity in TBLT……...…… 15

2.6. Empirical research on task complexity……..…….…… 31

CHAPTER 3: METHODOLOGY……..……... 53

3.1. Introduction……...…….…...…….. 53

3.2. Participants……...……... 53

3.3. Research design……..…… 53

3.4. Materials……..……... 54

3.5. Procedure……..…….. 57

3.6. Operationalization of task complexity……..…….. 58

3.7. Measures……..……... 59

3.8. Data analysis…….…..…… 63

CHAPTER 4: RESULTS……...…… 66

4.1. Introduction……...…….. 66

4.2. The perceived difficulty of tasks……..….…….. 66

4.3. The length of time…..……..…….. 68

4.4. Dependent variables……...…….…… 70

CHAPTER 5: DISCUSSION……...…….. 80

5.1. Introduction……...…….. 80

5.2. The relationship between time perspective and L2 oral performance.... 80

5.3. The relationship between online planning and L2 oral performance…. 84

5.4. The combined effects of time perspective and online planning……. 86

CHAPTER 6: CONCLUSION……...……..... 88

6.1. Summary and conclusion……...…….…… 88

6.2. Limitations and suggestions for further research……...……...….. 90

APPENDIX A: PICTURES USED IN TASKS……..…….. 93

APPENDIX B: BACKGROUND QUESTIONNAIRE……..……..… 95

APPENDIX C: LIKERT SCALE FOR PERCEIVED TASK DIFFICULTY...….. 96

APPENDIX D: TASK INSTRUCTIONS IN ENGLISH…..……. 97

APPENDIX E: TASK INSTRUCTIONS IN TURKISH……...…..…..…99

APPENDIX F: EXAMPLE AS-UNIT CODING……..……..…..... 101

APPENDIX G: CODING GUIDELINES FOR SECOND CODER……..……. 104

APPENDIX H: INTERCODER RELIABILITY SCORES……..…….. 109

REFERENCES……..…….. 110

LIST OF TABLES

Table 1. Example Definitions of Task in TBLT……..…..…… 13

Table 2. Skehan's Three-Way Distinction for Task Sequencing……...…...……. 19

Table 3. The Triadic Componential Framework for Pedagogic L2 Task Classification…….. 24

Table 4. Example Dual Tasks with Different and Dame Resource Pools..…..……. 27

Table 5. Summary of Studies Manipulating Tasks along +/- Here-and-Now Variable…….. 36

Table 6. Summary of Studies Manipulating Tasks along +/- Online Planning Variable…….. 44

Table 7. The Design of the Study…...……..……. 54

Table 8. The Features of the Picture Strips……..…...……... 55

Table 9. Difficulty Perception of Two Tasks……...…….. 66

Table 10. Statistics for the Length of Time (Seconds) Spent on Task...……..……. 69

Table 11. Descriptive Statistics for Complexity Measures…….……... 70

Table 12. Descriptive Statistics for Accuracy Measures…..……...……... 74

LIST OF FIGURES

Figure 1. Theorizing dimensions of performance…….. 21

Figure 2. Types of planning……...…… 41

Figure 3. Task difficulty ratings of Pressured Online Planning group for two tasks with 95% confidence intervals……..…… 67

Figure 4. Task difficulty ratings of Unpressured Online Planning group for two tasks with 95% confidence intervals……...…… 68

Figure 5. Total time spent on task by two group with 95% confidence intervals..... 69

Figure 6. Overall complexity scores of two groups in two tasks with 95% confidence intervals……...……... 71

Figure 7. Subordination scores of two groups in two tasks with 95% confidence intervals……...……... 72

Figure 8. Phrasal complexity scores of two groups in two tasks with 95% confidence intervals……...……... 73

Figure 9. Errors per 100 words with 95% confidence intervals……..…….. 75

Figure 10. The ratio of error-free AS units with 95% confidence intervals…..…..... 76

Figure 11. The ratio of error-free clauses with 95% confidence intervals…...…….. 77

CHAPTER 1

INTRODUCTION

1.1 Background to the study

Tasks provide "real communication and learner-centeredness in the classroom by respecting learners' interests and attending to interlanguage development" (Norris, 2009, p. 580). Therefore, they have been widely used in Second Language Acquisition (SLA) field, especially in Task Based Language Teaching (TBLT) over the last twenty years. In TBLT, the syllabus is composed of pedagogic tasks that have been created based on the real tasks that learners are expected to achieve in their lives. Thus, designing a task-based syllabus requires designers to decide on what learners should be able to do upon completing all the tasks. Moreover, the sequence of the tasks in a syllabus is argued to be as important as the selection of tasks since grading of tasks affect learners' performances differently (Skehan, 1996).

The sequencing of tasks in a task-based syllabus includes many criteria such as valency and criticality (i.e. the communicative value), frequency, learnability, and complexity (Long, 2015). Long (2015) defines complexity in TBLT as "inherent, unchanging qualities of a task that make it more or less challenging than another task at a given moment in time" (p. 232). Tasks with different complexity levels are predicted to impact learners' L2 performances and their proficiency levels (Robinson, 2001a, 2005, 2007a, 2011a; Skehan, 1996, 1998, 2009).

L2 performance is considered to include multiple components namely complexity, accuracy, and fluency (CAF) rather than being "a unitary construct" (Housen, Kuiken, 2009; Housen, Kuiken & Vedder, 2012; Skehan, 2009). Complexity refers to "the extent to which learners produce elaborated language"

(Ellis & Barkhuizen, 2005, p. 139). Thus, it is achieved through the acquisition of new L2 components. In the light of the newly acquired L2 elements, nonnative like parts of learners' interlanguage undergo changes leading to accuracy. Therefore, accuracy is related to the "capacity to handle whatever level of interlanguage complexity a learner has currently attained" (Skehan, 1996, p. 46). Finally, fluency is concerned with "the production of language in real time without undue pausing or hesitation (Ellis & Barkhuizen, 2005, p. 139). Although each CAF construct taps into a different part of the L2 system, they also interact with each other. Therefore, changes in one construct may affect the changes in the other one.

There are two influential hypotheses predicting the effects of task complexity on L2 performance. These are the Limited Attentional Capacity Hypothesis and the Cognition Hypothesis proposed by Skehan (1998) and Robinson (2001a) respectively. These two hypotheses have opposite predictions for the effects of increased task complexity on learners' L2 performance. The former suggests that there is a limit in learners' attentional resources and that learners cannot pay attention to form because they prioritize meaning. Therefore, more complex tasks result in lower accuracy and syntactic complexity levels. In addition, Skehan states that even if learners attend to form, simultaneous attention to accuracy and complexity is not probable if not impossible. The latter, on the other hand, predicts the existence of multiple attentional resources which allows for simultaneous attention to different L2 constructs. Therefore, according to the Cognition Hypothesis, more complex tasks lead to higher accuracy and syntactic complexity levels whereas fluency is decreased.

1.2 The aims of the study

This study aims to investigate the effects of task complexity in L2 oral performance of highly proficient English learners. The tasks are manipulated along two features, namely task type (Here-and-Now/ There-and-Then) and online planning (Unpressured/ Pressured). Although there are several studies which examined the impacts of these features in learners' L2 performance, there are few, if any, studies which manipulated the tasks along these two features simultaneously. Manipulating tasks along two dimensions are argued to result in some combined effects (Robinson, 2005, 2011a). Therefore, this study examines both separate and combined effects of manipulating tasks in terms of task type and online planning dimensions in learners' accuracy and syntactic complexity levels. Sixty-four English learners with high proficiency participated in the study. They were randomly assigned to two groups (unpressured online planning/ pressured online planning) and they were asked to complete one simple and one complex task based on a series of pictures given to them. While Unpressured Online Planning group was given unlimited time to complete the tasks, participants in Pressured Online Planning group were given only 90 seconds to tell their stories. Accuracy was measured as Errors per 100 words, the ratio of error-free clauses, and the ratio of error-free Analysis of Speech units (AS-units). Syntactic complexity, on the other hand, was measured as overall complexity, complexity by subordination, and phrasal complexity as suggested by Norris and Ortega (2009).

The study investigated the following research questions:

1. How does manipulating task complexity along Here-and-Now/ There-and-Then affect the syntactic complexity levels of L2 learners' oral performance?

2. How does manipulating task complexity along Here-and-Now/ There-and-Then affect the accuracy levels of L2 learners' oral performance?
3. How does manipulating task complexity along pressured/ unpressured online planning affect the syntactic complexity levels of L2 learners' oral performance?
4. How does manipulating task complexity along pressured/ unpressured online planning affect the accuracy levels of L2 learners' oral performance?
5. Are there any interaction effects between task type and online planning in relation to syntactic complexity and accuracy levels of L2 learners' performance?

1.3 Organization of the thesis

This thesis is composed of six chapters. Chapter 1 is an introductory chapter and provides background information for the study as well as the aims of the study. Chapter 2 reviews the literature for first and second language speech production models, the measures of L2 performance, task complexity and previous studies manipulating tasks along +/- Here-and-Now and +/- Online Planning conditions. Chapter 3 gives information about the methodology and design of the study. Chapter 4 reports the findings for accuracy and syntactic complexity levels of participants in two tasks. Chapter 5 discusses the results of the study in the light of the two hypotheses, namely the Limited Attentional Capacity Hypothesis and the Cognition Hypothesis, first and second language speech production models and previous studies. Finally, Chapter 6 summarizes the study and presents the limitations and suggestions for further research.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter aims to, first, investigate how speech is produced in first language and in second language and to give a brief description of Levelt's speech production model for first language (1989, 1999) and Kormos' bilingual speech production model (2006, 2011) to have an idea for the speech production in oral tasks. Secondly, it examines the constructs that constitute second language performance, namely, fluency, accuracy, and complexity. Finally, it provides an overview of tasks and task complexity research and focuses on the two mainstream hypotheses explaining the effects of task complexity in L2 performance, Skehan's (1998) Limited Attentional Capacity Hypothesis and Robinson's (2001a) Cognition Hypothesis.

2.2 First and second language speech production models

Levelt's speech production model (1989, 1999) is one of the most cited first language (L1) speech production models. This model proposes speech production to be modular consisting of a number of processing components. According to this speech production model, there are three basic mechanisms in the speech system. These are conceptualization, formulation, and articulation. Conceptualization involves conceptual preparation for speech. During the conceptualization stage communicative goals, topic, focus and content of a speech are determined through macroplanning and microplanning stages. Whereas macroplanning includes determining the communicative goals, microplanning is related to the propositional content of the message and involves deciding on the perspective and the focus of the

speech. Since determining communicative intentions require attention, conceptualization stage is “highly controlled” (Levelt, 1989, p.21). The conceptualization stage creates the pre-verbal plan and it includes all necessary information for the linguistic message. Later on, in the formulation stage these pre-verbal messages are converted into linguistic messages retrieving lexical items from the mental lexicon. After composing the surface structure of a speech according to the meaning and syntax of lexical items, phonological and morphological features are incorporated into this structure and phonetic/ articulatory plan, in other words, internal speech, is generated. During the articulation stage, internal linguistic messages are articulated as overt speech. Unlike conceptualization, formulation and articulation are automatic. The model also includes self-monitoring in different phases of speech production. Processing of all mechanisms in the model is incremental. For example, when a syllable of a word is encoded in the formulator, it can be articulated without waiting for the whole sentence to be completed. Moreover, all mechanisms work in parallel.

de Bot (1992) argues that “many aspects of speaking are the same for monolingual and bilingual speakers, and a single model to describe both types of speaker is to be preferred over two separate models for different types”. Therefore, incorporating de Bot’s (1992) research, Kormos (2006, 2011) proposes a bilingual speech production model based on Levelt’s speaking model for monolinguals. This model is similar to Levelt’s model regarding that it also consists of three mechanisms, namely, conceptualizer, formulator, and articulator and they work incrementally and in parallel as in Levelt's Model. However, differently from monolingual speech production, parallel processing can only be achieved after a certain proficiency level. In Levelt’s L1 model, formulating and articulating stages

are asserted to be autonomous, whereas conceptualizing and monitoring are controlled. In the bilingual model, however, when L2 proficiency is low, formulating and even articulating stages as well as conceptualizing and monitoring need conscious attentional control. To explain this controlled stages of formulation and articulation, Kormos argues for the existence of a knowledge store which accumulates lexical and phonological rules in declarative knowledge form and states that this store allows for unautomated phonological and syntactic rule formation. Only when proficiency increases, the stages of formulation and articulation become more automatized. Apart from several additions to L1 speech production model, Kormos asserts L1 and L2 speech production to be similar.

In summary, both L1 and L2 speech production are suggested to be the part of the same system which works differently in L1 and L2. This system is composed of a conceptualizer, a formulator, and an articulator. Differently from L1, during the early stages of L2 acquisition, a speaker needs to attend to all these three mechanisms consciously. Therefore, L2 learners need to be allocated sufficient time for their speech.

So far, the widely accepted monolingual and bilingual speech production models have been reviewed. The following section will focus on the assessment of speech produced in the system explained above and it will present an overview of the constructs that are used to measure L2 oral production.

2.3 Measures of L2 production

L2 proficiency has attracted the interest of numerous researchers for many years. Especially in recent years, many researchers as well as educators have started to consider L2 proficiency as a multi-dimensional construct which is composed of three

components; complexity, accuracy, and fluency (CAF) (Ellis & Barkhuizen, 2005; Housen & Kuiken, 2009; Housen et al., 2012; Skehan, 1998). Besides, Skehan (2009) has incorporated lexis to these constructs.

Fluency is defined as "the capacity to produce speech at normal rate and without interruption" (Skehan, 2009, p. 510). Although, historically it has been used for assessing global L2 proficiency of learners (Housen & Kuiken, 2009; Housen et al., 2012), fluency as a CAF construct constitutes only one part of L2 performance. In addition, it is treated as a multi-componential construct consisting of breakdown/pausing fluency, speed fluency, and repair fluency (Housen & Kuiken, 2009; Skehan, 2009; Tavakoli & Skehan, 2005). Speed fluency is indicated as syllables or words per minute and is generally measured with speech rate (e.g. number of syllables per minute). Breakdown fluency is related to pausing behavior and it is measured as length and number of unfilled pauses, filled pauses, and total amount of silence. Lastly, repetitions, reformulations, false starts, and replacements in a speech are measured for repair fluency.

Accuracy is regarded as the oldest, simplest and most internally consistent construct which is based on the errors in learners' performances (Housen & Kuiken, 2009; Pallotti, 2009). Housen et al. defines accuracy as "the extent to which an L2 learner's performance (and the L2 system that underlies this performance) deviates from a norm" (p. 4). While assessing accuracy, most researchers have referred to native-speaker norms and they have regarded the differences as errors in learners' performances. Accuracy is measured with various measures. Some measures focus on the accuracy of the performance (e.g. the ratio of error-free units or the ratio of error-free clauses) whereas some examine the errors available in the performance (e.g. the mean number of errors per units and errors per 100 words). In addition to

general measures, accuracy is also measured with some task-specific measures such as the correct use of specific grammatical structures (e.g. articles and plural –s).

The final construct of the triad, complexity, is stated to be the "most complex and enigmatic" one of all three constructs (Housen et al., 2012, p.10) since the same term is used in various areas of instructed SLA (e.g. linguistic complexity, cognitive complexity). As a learner language construct, Housen et al.(2012) defines complexity as "the ability to use a wide and varied range of sophisticated structures and vocabulary in the L2" so in CAF triad complexity refers to linguistic complexity of L2 performance and proficiency (p.2). According to this definition, complexity consists of syntactic complexity and lexical complexity.

Syntactic complexity is generally measured as the amount of subordination in most studies (e.g. Gilabert, 2007a, Kuiken & Vedder, 2011, Michel, 2011, Robinson, 2007b). However, Norris and Ortega (2009) argue that measuring complexity as the amount of subordination is too limited to capture complexity as a dynamic, multidimensional construct. Therefore, they suggest employing some measures for the amount of coordination (i.e. the overall complexity) and the mean length of an utterance (i.e. phrasal complexity) as well as the amount of subordination for measuring complexity. Moreover, task specific measures such as structural variety (Ahmadian & Tavakoli, 2010; Yuan & Ellis, 2003) can also be used in studies since these measures "capture the variety, sophistication, and acquisitional timing of grammatical forms used in production" (Norris & Ortega, 2009, p.558).

Lexical complexity is generally measured with type-token ratio (TTR) which is calculated by dividing the number of word types to all word tokens. However, this measure is influenced by the length of a text (Skehan, 2009), therefore it is not the most reliable measure of lexical variety. Most of the recent studies employ D-value

and Guiraud's Index of Lexical Richness which use random sampling of the tokens and compensate for text length (Gilabert et al., 2011; Michel et al., 2007). Skehan (2009) points out the importance of examining lexical sophistication in addition to lexical variety measures since these measures may not be very effective in showing task effects in some contexts. Therefore, he suggests employing Lambda, which is calculated based on the distribution of numbers of difficult words, as a lexical sophistication measure.

On the basis of the definitions and measures for CAF constructs given above, they seem to be independent constructs. However, in addition to being distinct, CAF constructs also complement each other (Norris & Ortega, 2009), so all of them should be taken into account when considering L2 proficiency and performance of learners. Additionally, Norris and Ortega propose to have more "organic" and "sustainable" measures for CAF constructs so that these constructs can capture the changes in learner performance in various learning contexts over time and lead to generalized understandings related to L2 learning (p. 556).

In the current section, the constructs used to measure learners' L2 performance have been examined. Since these constructs are operationalized through various tasks in Task-Based Language Teaching (TBLT), the focus of the following section will be tasks used in L2 teaching and learning as well as in the field of instructed SLA.

2.4 Tasks in task-based language teaching

Tasks used in promoting second language learning have always been a great interest not only for teachers but also for researchers in instructed SLA field. This interest can be attributed to the influential nature of tasks in learning an L2.

Firstly, tasks play an important role in the allocation of attention. In the Noticing Hypothesis, Schmidt (1990, 2001) states that noticing, in other words conscious attention to input, is important for learning since it converts input to intake. Schmidt (1990) discusses several factors that affect noticing. These are a) the frequency of the form, b) perceptual saliency, c) instruction, d) individual differences in processing ability, e) readiness to notice, and f) task demands. For explaining how task demands affect noticing, Schmidt discusses that tasks with high demands may decrease the chances of noticing some grammatical structures, but tasks with low demands may provide better opportunities for noticing these structures. Based on Schmidt's (1990) arguments Skehan (1998) exemplifies the effects of task demands as follows:

> …tasks based on familiar information with clear discourse structure, for example a pair of students giving one another instructions to get to their respective homes, will probably have low task demands, while a task requiring imagination and abstraction, and a complex outcome, such as agreeing on the solution to a moral problem, will probably make much higher ones (p. 51).

Secondly, tasks have a potential to lead learners to use language communicatively (Long, 2015; Norris, 2009). Therefore, tasks can provide learners conditions to negotiate for meaning and to get feedback. As stated in Long's Interaction Hypothesis (1996), these negotiations of meaning direct learners' attention to emerging linguistic problems during the communication and increase the likelihood of noticing new forms and creating new meaning-form mappings.

Thirdly, since tasks require learners to achieve various outcomes, they encourage learners to produce language. According to the Output Hypothesis (Swain, 1993, 1995; Swain & Lapkin, 1995, 2002) when learners produce language, they can notice the gaps and problems in their interlanguage (IL) and modify them.

All these theories in instructed SLA and facilitative roles of tasks led to the development of Task-Based Language Teaching (TBLT) approach. In TBLT, tasks rather than grammatical forms play a key role in the learning of a new language and broadening the existing knowledge. Tasks are created according to learners' needs and used as units of analysis in design, implementation and evaluation of a syllabus.

TBLT aims to enable learners achieve real-world tasks (i.e. target tasks) that have been found to be necessary for them in the needs analysis. Therefore, some pedagogic tasks which prepare learners for the real-world tasks are employed in TBLT classrooms. After selecting the pedagogic tasks to be used, these tasks are sequenced to create a task-based syllabus (Long, 1985, 2015). The term "task" used in the rest of this thesis will refer to "pedagogic tasks" rather than "target tasks".

Although tasks have been widely employed in L2 teaching and learning (Nunan, 1989, 2004; Prabhu, 1987; Willis, 1996) and in many TBLT studies (Foster & Skehan, 1996; Gilabert, 2007a, 2007b; Robinson, 1995, 2007b; Robinson, Cadierno, & Shirai, 2009; Skehan & Foster, 1997; Tavakoli & Foster, 2008), there is not an agreement for the definition of the term "task" among researchers. Table 1 indicates some examples for different descriptions of tasks in TBLT research.

Table 1. Example Definitions of task in TBLT

Long (1985)	A piece of work undertaken for oneself or for others, freely or for some reward. Thus, examples of tasks include painting a fence, dressing a child, filling out a form, buying a pair of shoes, making an airline reservation, borrowing a library book, taking a driving test, typing a letter... In other words, by "task" is meant the hundred and one things people do in everyday life, at work, at play, and in between. Tasks are the things people will tell you to do if you ask them and they are not applied linguists (p. 89).
Prabhu (1987)	An activity which required learners to arrive at an outcome from given information through some process of thought, and which allowed teachers to control and regulate that process, was regarded as a "task" (p. 24).
Willis (1996)	An activity where the target language is used by the learner for a communicative purpose (goal) in order to achieve an outcome (p. 23).
Skehan (1998)	An activity in which meaning is primary, there is some communication problem to solve, there is some sort of relationship to comparable real-world activities, task completion has some priority, and the assessment of the task is in terms of outcome (p. 95)
Nunan (2004)	A piece of classroom work that involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is focused on mobilizing their grammatical knowledge in order to express meaning, and in which the intention is to convey meaning rather than to manipulate form. The task should also have a sense of completeness, being able to stand alone as a communicative act in its own right with a beginning, a middle and an end (p.4)

Long's (1985) definition for tasks includes various real world actions and activities that people do in their everyday lives and some of these activities do not even need language use (e.g. painting a fence, dressing a child). Although this definition is very broad, it emphasizes that tasks should be similar to real-life activities.

Skehan (1998) also focuses on the resemblance of tasks to real-world situations. Besides, he states that meaning rather than form should be a priority in tasks and that tasks should achieve an outcome. Similar to Skehan, Nunan (2004) also points out the focus of meaning in tasks. In his definition a task is presented as "a piece of classroom work that involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is focused on mobilizing their grammatical knowledge in order to express meaning, and in which the intention is to convey meaning rather than to manipulate form" (Nunan, 2004, p.4).

Similar to Skehan (1998), Prabhu (1987) and Willis (1996) focus on the importance of achieving an outcome at the end of a task. Furthermore, Willis (1996) asserts that the use of the target language has a communicative purpose. In other words, a task should be goal-oriented.

Although the definitions of task in TBLT vary, their common idea is that a task is a real-life like meaningful communicative activity with a specific goal that needs to be achieved at the end of the task. According to the definitions given above, a task in TBLT includes the following common features: a) a task resembles real-world situations, b) a task has an outcome that needs to be achieved at the end of the task, c) a task involves communicative language use and emphasizes meaning rather than form since learners should, first, convey the meaning to complete the task or achieve the outcome.

Willis and Willis (2001) states that a focus on meaning and outcome in tasks distinguish them from grammatical exercises because in tasks learners are not given specific language structures to use as in grammatical exercises, instead they are free to choose the structures that they will use during the task to achieve the outcomes of the task.

So far, various definitions of tasks, their use in TBLT, and the importance of tasks in L2 learning and teaching has been reviewed in a general sense. In TBLT, sequencing and grading of tasks are as important as the tasks themselves. Therefore, the next section will present various attempts that have been made for grading and sequencing tasks according to their complexity levels.

2.5 Task complexity in TBLT

In TBLT, grading and sequencing of the pedagogic tasks utilized in language learning and teaching as well as having a rationale for the chosen units of analysis is important (Long, 2015). In particular, grading and sequencing pedagogic tasks in terms of their complexity (i.e. difficulty) have attracted the attention of many researchers. Here, complexity refers to the complexity of tasks themselves rather than the linguistic features (e.g. having easy or difficult grammatical structures) of the content. Long (2015) defines task complexity as "inherent, unchanging qualities of a task that make it more or less challenging than another task at a given moment in time" (p. 232). Task complexity is argued to be important since changing the complexity of tasks may lead to the desired pedagogical outcomes (Skehan, 1998).

2.5.1 Initial approaches to task complexity

Tasks have been attempted to be categorized as "simple" and "complex" since 1980s. In 1984, Brown, Anderson, and Yule investigated the effects of task complexity through different speaking task designs. They suggested tasks to be sequenced in terms of the type of information contained and the relationships among the elements of tasks. According to them static tasks which require the description of a never changing relationship between objects such as describing a diagram are the simplest tasks. Dynamic tasks, on the other hand, where the learners need to describe a sequence of actions or activities such as telling a story about a car crash, are more complex tasks. Finally, abstract tasks such as expressing opinions on a topic are regarded as the most complex tasks since they include decontextualized elements and require argumentation or justification about abstract rather than concrete elements.

Furthermore, in the Bangalore Communicational Teaching Project, Prabhu (1987) attempted to teach language based on task rather than form. This project was motivated by a necessity to change the traditional, structural syllabi with the new ones which created conditions to handle meaning in the classroom. The project was conducted in a number of English classes in India for a few years. During the project, Prabhu stated a preference for challenging classroom activities which engaged learners in reasoning, inferring or problem solving. He suggested that the most satisfying activities allowed for negotiation and their "demand on thinking was just above the level which learners could meet without help" (p. 24). Therefore, he proposed a hierarchy based on thinking demands of tasks in which the opinion-gap tasks were the most difficult ones for learners because of their levels of unpredictability. They posed learners a greater challenge than reasoning-gap activities, and reasoning-gap activities posed a greater challenge than information-gap activities. Prabhu proposed using reasoning-gap tasks in classrooms since they required learners to derive new information from the information given in the task and allowed for assessing whether the outcome was right or wrong. He claimed that opinion-gap tasks were too difficult to complete for learners because they were more decontextualized and did not lead to a common correct answer. He also suggested using information-gap tasks as a pre-task for reasoning-gap tasks since they were easier for learners to perform and helpful in preparing them for the main task.

Differently from previous researchers, Candlin (1987, as cited in Robinson, 2001a) established a more extensive criteria for sequencing and grading tasks. He identified six features for classifying tasks: cognitive load (the general cognitive complexity of the task), communicative stress (e.g. the existence of various interlocutors), particularity and generalizability (the relative familiarity or novelty of

tasks), code complexity and interpretative density (the linguistic and argumentative complexity of texts used in tasks), process continuity (learners' choice about sequencing tasks) and content continuity (resemblance to real world tasks).

Similarly, Nunan (1989, 2004) suggested three factors that had an effect in grading and sequencing of tasks. Input factors referred to aspects related to the content such as grammatical difficulty, the length of the text, the propositional density, the amount of low frequency vocabulary, the speed of spoken texts and the number of speakers involved, the explicitness of the information, the discourse structure and the clarity with which it is signaled. He also considered learners factors such as background knowledge, confidence, motivation, learning pace, observed ability in language skills, cultural knowledge/awareness, linguistic knowledge as factors affecting the completion of a task. Finally, he discussed the activity factors (i.e. procedural factors) which referred to the operations that learners needed to perform in the task. He stated that relevance, complexity, amount of context provided prior to the task, processability of language of the task, amount of help available to the learner, degree of grammatical complexity, time available to the learner, and follow-up affected the grading of tasks.

In line with previous research, Long & Crookes (1992) asserted that task sequencing should be based on the cognitive complexity of tasks in addition to the linguistic features. They suggested sequencing tasks from simple to complex according to their cognitive complexity levels since this sequence could help learners build on their existing knowledge and achieve the target tasks in L2. They provided several examples that had an impact on the cognitive complexity of tasks:

> 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 a pedagogic task poses are just a few of the potential grading and sequencing criteria that have been proposed (p.45).

These early notions for categorizing and sequencing tasks involve various task features that differed in terms of their openness to manipulation. Whereas some features such as input factors and task factors can be manipulated easily by a teacher or a researcher, other features such as learner factors are brought to the task externally by the learner and these factors have an altering effect on the complexity of a task. Therefore, Robinson (2001a) criticizes these frameworks for not differentiating task complexity and task difficulty which are two distinct areas in current TBLT research.

Regardless of their shortcomings, these previous discussions of task complexity have provided a basis for the development of two most influential models in TBLT. These are Skehan's (1998) three-way distinction 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 complexity on task completion. However, they have different theoretical foundations, so that they predict the role of cognitive complexity in learners' performances and L2 learning differently. In the next two sections, Skehan's and Robinson's models will be reviewed.

2.5.2 Skehan's three way distinction and the Limited Attentional Capacity Hypothesis

According to Skehan (1996, 1998) and Skehan and Foster (2001) the complexity of tasks can be manipulated in various dimensions. Following Candlin (1987, as cited in Skehan, 1998) and Nunan (1989), Skehan (1996, 1998) divides the features affecting task complexity into three areas; Code complexity, cognitive complexity, and communicative stress. Table 2 illustrates these three areas and their components.

Table 2. 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 – sufficiency of information given	– time pressure – scale – number of participants – length of texts used – modality – stakes – opportunity for control

Source: [Skehan, 1998]

Code complexity is concerned with the difficulty and density of the syntactic and lexical elements used in a task. A task with difficult structures such as complex tenses and advanced vocabulary requires the use of more linguistic resources for the completion of the task. Therefore, these tasks with more advanced structures and vocabulary are more complex than the ones with simple structures and basic vocabulary.

Cognitive complexity, on the other hand, refers to the content of task performance, rather than linguistic aspects, and the ways for manipulating it. Skehan (1998) categorizes cognitive complexity into two: cognitive processing and cognitive familiarity. Cognitive processing is about the online processing demands that a task imposes on learners and it includes the amount of online computation, information

organization, clarity and sufficiency of information given. Cognitive familiarity, however, is related to the task- related knowledge that learners hold and it covers the familiarity of a topic, a discourse genre and a task to a learner. Thus, when the cognitive processing demands of a task increase, the complexity of that task increases. However, an increase in cognitive familiarity results in lower task complexity levels.

Finally, communicative stress involves the factors affecting the completion of a task in addition to its linguistic code and meaning. These factors are time pressure (the time period for completing a task), modality (e.g. speaking vs. writing), scale (e.g. the number of participants, the number of relationships), stakes (the importance of completing a task correctly), and control (the amount of control that participants have on the implementation of a task).

For discussing the effects of manipulating tasks along these three dimensions stated above, Skehan (1996, 1998, 2009, 2014) proposes the Limited Attentional Capacity Hypothesis. He argues that humans' attentional resources and working memory capacities are limited so that they cannot pay concurrent attention to content and form. He supports his hypothesis with a study conducted by VanPatten (1990). VanPatten investigated a group of Spanish L2 learners 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 (the definite article or a particular morphological marker). At the end of the study, the researcher found learners to prioritize content unless they are instructed to pay attention to form. In addition, attending to grammar resulted in less comprehension. In other words, learners had difficulty in directing attention to the content and the form of the input

simultaneously and when they were not asked to pay attention to a grammatical structure, focus on meaning was primary for them.

Building upon VanPatten's (1990) study, Skehan (1998) asserts that humans have limited attentional capacity. Moreover, when dealing with a high complexity task, content will take up most of the attention since these tasks require extensive online processing. Therefore, there will be less attention left for form. This means that 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). Moreover, focus on form can only be achieved when "it is necessary for the recovery of meaning" (Skehan, 1996, p.45) and when meaning is "executed automatically" without much attention (Skehan & Foster, 2001, p.189). Figure 1 indicates the relationship among different performance dimensions.

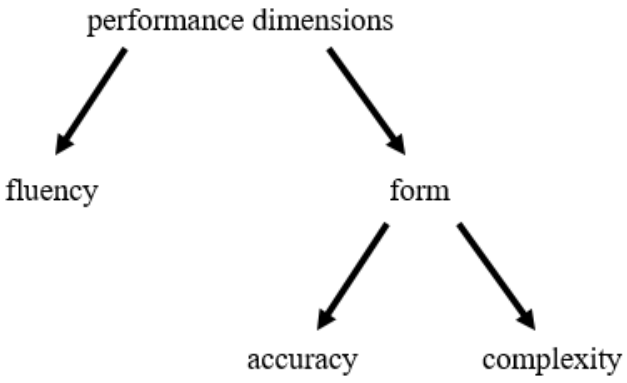


Figure 1. Theorizing dimensions of performance
Source: [Skehan & Foster, 2001]

According to Skehan (1996) when learners are able to direct their attention to form, it is not possible to attend to all three linguistic aspects of language, namely fluency, accuracy, and complexity, at the same time because attending to one aspect takes up the attention from the others. Skehan (1998, 2009) argues for a particular tension between complexity and accuracy and advocates trade-off effects between these two

aspects. This hypothesis is supported in a study conducted by Foster & Skehan (1996). In that study, the researchers employed tasks with various cognitive complexity levels. At the end of the study, different tasks were found to be improving different language performance aspects. For example; the Personal Task produced a great degree of accuracy at the expense of complexity while in the Narrative Task students achieved higher degree of complexity, but their accuracy level was the lowest among all tasks. Therefore, the researchers argued that there were trade-off effects between accuracy and complexity and that achieving both greater complexity and greater accuracy was not possible.

Skehan & Foster (2001) argue that this limit in attentional capacity resources affects non-native speakers of a language more than native speakers. Since native speakers' knowledge of the language forms is proceduralized, these forms can be attended automatically without control. However, for non-native speakers, this process needs attentional control. For non-native speakers, Skehan and Foster suggest two different strategies of focusing on form and discuss the effects of more cognitively demanding tasks in these words:

> … tasks which are cognitively demanding in their content are likely to draw attentional resources away from language forms, encouraging learners to avoid more attention-demanding structures in favor of simpler language for which they have already developed automatic processing (the "safety first approach"). Conversely, very cognitively demanding content might result in learners paying insufficient attention to language forms which still require controlled processing and which are therefore likely to be poorly done (the "accuracy last approach") (p. 189).

Taking the limited attentional capacity and working memory of humans, Skehan (1998) suggests that tasks used in language learning and teaching should be chosen carefully to achieve a balanced interlanguage development in terms of fluency,

accuracy, and complexity. Besides, he states that the difficulty of tasks should be appropriate for the target L2 learners.

In summary, Skehan (1998) argues for a limit in attentional resources and working memory capacity in humans. Therefore, he states that attending to all performance constructs, especially to accuracy and complexity, simultaneously is not possible for L2 learners in complex tasks. The following section will introduce an opposing view held by Robinson (2001a) and give an overview for his framework which categorizes tasks and for the Cognition Hypothesis which makes predictions for the effects of different tasks in L2 performance.

2.5.3 Robinson's Triadic Componential Framework and the Cognition Hypothesis

Robinson (2001a, 2001b, 2005, 2007a, 2007b, 2011a, 2011b) proposes a theoretically motivated taxonomy for classifying pedagogic tasks according to their complexity levels. Unlike previous researchers and Skehan's three-way distinction, Robinson differentiates the terms task complexity and task difficulty. In his Triadic Componential Framework, shown in Table 3, he lists three areas for task classification. These are task complexity, task condition, and task difficulty.

All the three factors presented in the framework tap into a different aspect of task complexity affecting the task performance and each factor contains various subcomponents.

Table 3. The Triadic Componential Framework for Pedagogic L2 Task Classification

Task Complexity (Cognitive Factors)		Task Condition (Interactive Factors)		Task Difficulty (Learner Factors)	
Resource-directing	Resource-dispersing	Participation	Participant	Ability	Affective
+/- Here-and-Now	+/- Planning time	+/- Open solution	+/- Same proficiency	H/L Working memory	H/L Openness
+/- Few elements	+/- Prior knowledge	+/- One way flow	+/- Same gender	H/L Reasoning	H/L Control of emotion
+/- Spatial reasoning	+/- Single task	+/- Convergent solution	+/- Familiar	H/L Task-switching	H/L Task Motivation
+/- Causal reasoning	+/- Task structure	+/- Few participants	+/- Shared content knowledge	H/L Aptitude	H/L Anxiety
+/- Intentional reasoning	+/- Few steps	+/- Few contributions needed	+/- Equal status and role	H/L Field independence	H/L Willingness to communicate
+/- Perspective-taking	+/- Independency of steps	+/- Negotiation not needed	+/- Shared cultural knowledge	H/L Mind-reading	H/L Self-efficacy

Source: [Robinson, 2011]

Task complexity is concerned with the attention and memory related processing demands of tasks. Therefore, it can explain learner differences in performing two tasks such as a simple task and a complex task (Robinson, 2001a; Robinson & Gilabert, 2007). According to Robinson, task complexity is comprised of resource-directing (cognitive/ conceptual demands) and resource-dispersing (performative/ procedural demands) dimensions.

Increasing the complexity of tasks along resource-directing dimension enables learners to direct their attention to specific L2 structures and forms and linguistic aspects such as accuracy and syntactic complexity since they are necessary for completing the task correctly (Robinson, 2001a, 2001b, 2005, 2007a, 2011a). For example a task which requires learners to support their ideas while giving reasons

can help them to direct their attention to specific L2 structures such as subordinators (e.g. because, thus, etc.) (Robinson, 2005). In his framework, Robinson (2011a) includes various subcomponents for resource-directing dimension of task complexity. Tasks can be manipulated through +/- here-and-now, +/- few elements, +/- spatial reasoning, +/- causal reasoning, +/- intentional reasoning, +/- perspective-taking. Increasing the complexity of tasks along these areas leads to the analysis of specific linguistic items in L2 system and "promotes second language acquisition" (Robinson, 2011a, p.14).

Resource-dispersing dimension, on the other hand, includes +/- planning time, +/- prior knowledge, +/- single task, +/- task structure, +/- few steps, +/- independency of steps. This dimension manipulates the non-linguistic features of tasks so it disperses learners' attention rather than directing it on linguistic aspects and enabling analysis of linguistic items. For instance, when learners are asked to complete a task, they direct all their attention to the completion of this task. However, if they are asked to complete a second task simultaneously, this task diverts their attention and memory resources away from linguistic aspects to other processing activities necessary to complete the two tasks at the same time. Therefore, increasing tasks along this dimension promotes "real time access" and control on already available linguistic knowledge (Robinson, 2005, p.7).

The factors related to interactional demands of tasks are classified under task condition. The interactional demands are distinguished as participation variables and participant variables. Participation variables include whether tasks are open (has multiple answers) or closed (has only limited number of answers), whether they are one-way or two-way tasks (whether the task needs to be completed individually or in pairs/groups), whether they are convergent or divergent tasks (whether learners need

to decide on one answer or not), whether they require few or more participants, whether few or more new contributions needed, and whether tasks require negotiation. Participant variables, on the other side, consist of proficiency, gender, familiarity, equal status, shared content knowledge, and shared cultural knowledge.

Task difficulty involves individual differences varying from one learner to another and these differences affect how learners perceive a task. Robinson (2007a) classifies individual differences into two: ability variables (aptitude, working memory, reasoning, task-switching, field independence, mind-reading) and affective variables (openness, control of emotion, task motivation, anxiety, willingness to communicate, self-efficacy). The factors listed under task difficulty addresses between-learner differences when completing a task (Robinson, 2001b; Robinson & Gilabert, 2007). Robinson (2007b) emphasizes the importance of taking individual differences into consideration since they "increasingly differentiate learners' speech production, and interaction and uptake, as tasks increase in complexity" (p.196). According to Robinson (2005, 2011a), when the effects of learner factors on learners' performances in different tasks are known, learners can be supported when their abilities to perform a specific task are low. Moreover, they can be matched to tasks and dimensions to increase their chances of learning.

Robinson (2001a, 2001b) asserts that in a task-based syllabus pedagogic tasks should be sequenced from simple to more complex according to their cognitive complexity levels. Moreover, he states that task condition and task difficulty can only have an impact for online decisions made about the implementation of these tasks in classroom (Robinson, 2005).

Differently from Skehan's (1996, 1998) Limited Attentional Capacity Model, Robinson (2011a) promotes the Multiple Attentional Capacity Model which claims

the presence of various attentional resources that allow humans to direct their attention to more than one aspect of their performances. His hypothesis is underpinned by Wickens' (2002, 2007) Multiple Resource Model. Wickens' model explains the dual task interference with the presence of three dimensions for allocating attention. These are processing modality (visual or auditory), processing code (verbal/linguistic or spatial/analog) and processing stage (perception or response). According to Wickens, performing two tasks which require the use of opposite resource pools poses less burden than two tasks drawing attention from the same pool. The tasks using the same resources interfere with each other, so it is more likely to have a competition among these tasks. For example, performing a task which requires the use of visual and auditory resources (e.g. reading a book while listening to a teacher) is easier than two tasks with the visual input (e.g. reading a book while watching a movie). Table 4 presents more examples for easy and difficult tasks according to Wickens' multiple resource model.

Table 4. Example Dual Tasks with Different and Same Resource Pools

	Processing Modality	Processing Code	Processing Stage
Easy	Reading a book (visual) while listening to a teacher (auditory)	Listening to a song (verbal) while driving (spatial)	Listening to a teacher (perception) while taking notes (response)
More difficult	Reading a book (visual) while watching a movie (visual)	Listening to a song (verbal) while reading (verbal)	Listening to a teacher (perception) while reading a book (perception)

On the basis of Wicken's (2002, 2007) model, Robinson (2011a) argues that humans have multiple attentional resources, so they can pay concurrent attention to multiple aspects of the language. In other words, he claims that different language aspects, namely complexity, accuracy, and fluency draw on different attentional resources so that there is not a competition between them.

For explaining how complex tasks affect L2 learner performance, Robinson (2001a, 2001b, 2005, 2007a, 2007b, 2011a) offers the Cognition Hypothesis. He argues that when the complexity of tasks are increased along resource-directing dimension, these tasks require higher cognitive and conceptual processing. For this reason in more complex tasks, learners are encouraged to attend to complexity and accuracy of their language so that their performance will be more complex and accurate, whereas their fluency will decrease. On the contrary, increasing the complexity in resource-dispersing dimensions such as the absence of planning time requires higher performative and procedural demands from learners. These performative demands decrease the complexity, accuracy, and fluency of learners' performances.

Moreover, Robinson (2011a) proposes the existence of combined effects of manipulating tasks along both resource-directing and resource-dispersing dimensions on L2 performance (Robinson, 2011a, p. 21). In other words, manipulating the complexity of tasks along resource-dispersing dimensions in addition to resource-directing dimensions may have an enhancing or weakening effect on learners' performances.

In short, unlike Skehan (1998, 2009), Robinson (2001a, 2005, 2007a, 2011a, 2011b) claims that cognitively more complex tasks require greater functional/communicational demands from learners and they enable learners direct their attention to linguistic features so that learners can attend to both accuracy and complexity at the same time in their performances. Since the hypotheses of these researchers have aroused great interest in the field of TBLT, the next section provides a comparison of their hypotheses.

2.5.4 The comparison of Skehan's Limited Attentional Capacity Hypothesis and Robinson's Cognition Hypothesis

Both Skehan's (1998) and Robinson's (2001a) frameworks and theories have been pioneers in task complexity. The main difference between Skehan's Limited Attentional Capacity Hypothesis and Robinson's Cognition Hypothesis stems from these researchers' opinions regarding the effects of high cognitive task demands in complexity and accuracy levels in L2 performance (Skehan, 2009). Skehan (1998, 2009) argues for trade-off effects between these two aspects because of the limited attentional resources and working memory capacity in humans and suggests that learners prioritize one over the other. This interpretation contrasts with that of Robinson (2001, 2005, 2007a, 2011a, 2011b) who claims that these constructs draw different attentional resources in mind so that they can be attended simultaneously in cognitively complex tasks along resource-directing dimension.

The studies examining task complexity present inconclusive findings for the effects of cognitively complex tasks in complexity, accuracy, and fluency levels. In a meta-analysis, Jackson and Suethanapornkul (2013) reports only small positive effects for accuracy and small negative effects for fluency whereas no effects for syntactic complexity in tasks with greater cognitive complexity, unlike the predictions of the Cognition Hypothesis. Yet still, Skehan (2009) mentions a few studies which reported a simultaneous increase in complexity and accuracy of learners' performances.

The first study conducted by Foster and Skehan (1999) investigated the effects of various pre-task planning conditions and they reported higher accuracy and complexity scores for learners in teacher-led group. Tavakoli and Skehan and colleagues (2005, 2008) carried out a series of studies investigating the effects of

task structure and storyline complexity. These studies reported simultaneous increase in complexity and accuracy in tasks with a clear structure and requiring the integration of background information in the story. Another study conducted by Foster and Skehan (2013) examined the effects of transcribing learners' own utterances as a post task activity and reported more accurate and more complex performances for learners who were post task group.

Although all these studies reported a simultaneous increase in accuracy and complexity of learners' performances in certain tasks and conditions, Skehan (2009) states that these results were not due to an increase in cognitive complexity, rather they were achieved through pedagogic interventions, the structure and linguistic demands of tasks, and learners' expectations of a post task. In other words, none of these combined increases resulted from particularly greater cognitive task complexity as proposed by Robinson. Skehan (2009) demonstrates a need for considering other factors affecting task performance, such as pedagogic interventions and learners' expectations, in addition to task complexity.

In short, although there is a great number of studies investigating the effects of task complexity, they only partially confirm the hypotheses proposed by Skehan (1998) and Robinson (2001a). Firstly, in contrast to the Limited Attentional Capacity Hypothesis which predicts decreased accuracy and complexity levels in more complex tasks, most of the studies present increases in either of these constructs. Secondly, unlike the predictions of the Cognition Hypothesis, most studies do not present simultaneous increase in accuracy and complexity levels and when they do, these increased effects do not stem from an increase in cognitive complexity of the tasks only.

This section provided a comparison for the predictions of Skehan (1998) and Robinson (2001a) about the effects of increased task complexity in L2 performance. The next section will, first, give an overview of empirical research on task complexity, and then focus on the two selected variables: +/- Here-and-Now and +/- Online Planning.

2.6 Empirical research on task complexity

As it has been discussed in previous sections, task complexity has been the focus of many researchers as well as teachers and material developers for its effects on L2 performance and L2 learning. The predictions of Skehan (1998) and Robinson (2001a) for the impacts of task complexity on performance constructs, namely, fluency, accuracy, and complexity have attracted the most attention. To recap, Skehan argues for an inability for allocating simultaneous attention to multiple language constructs. Therefore, he states that since cognitively complex tasks require lots of attention to conceptualization of ideas, little attention is left for the linguistic elements. In addition, paying attention to both complexity and accuracy at the same time is not possible. Robinson, on the other hand, supports the existence of multiple attentional resources in humans. Thus, he states that increased cognitive task complexity requires more elaboration from the learners and it results in greater accuracy and complexity. However, this situation causes a decrease in fluency. This decrease in fluency levels can be compensated with manipulating the performative aspects of the task such as providing pre-task planning time.

These contradictory predictions have been investigated by manipulating various variables in tasks. Many researchers have investigated the effects of increased cognitive task complexity (i.e. increasing the complexity along resource-

directing dimensions). For instance, they have manipulated the elements in tasks (Kuiken & Vedder, 2007, 2008, 2011; Michel, 2011; Michel Kuiken & Vedder, 2007; Revesz, 2011; Robinson, 2001b; Sasayama & Izumi, 2012), the reasoning demands of tasks (Gilabert, Baron, & Levkina, 2011; Kim, 2009; Kim & Tracy-Venture, 2011; Robinson, 2007b), and the perspectives taken in tasks (Iwashita, McNamara, & Elder, 2001). Some researchers examined the effects of resource-dispersing factors on L2 performance and they investigated many aspects such as pre-task planning (Ellis, 1987; Foster & Skehan, 1996; Gilabert, 2007a; Iwashita et al., 2001; Ortega, 1999; Sasayama & Izumi, 2012; Yuan & Ellis, 2003), the presence of a post-task (Foster & Skehan, 2013; Skehan and Foster, 1997), task repetition (Ahmadian & Tavakoli, 2010; Wang, 2014), and task structure (Ahmadian, Tavakoli, & Dastjerdi, 2015; Tavakoli & Foster, 2008).

The current study investigates the effects of manipulating tasks along task type (Here-and-Now/ There-and-Then) and online planning (Unpressured/ Pressured) variables. Therefore, the following two sections will provide a critical overview for the effects of two selected variables.

2.6.1 Here-and-Now/ There-and-Then studies

+/- Here-and-Now condition manipulates tasks in terms of time reference. In other words, it is related to "a task's location in time and space" (Long, 2015, p. 232). In Here-and-Now tasks, learners complete simpler and cognitively less demanding tasks than There-and-Then tasks since they require the use of present, context-supported references. For example, talking about a car crash with another eyewitness at the crash scene is simpler than talking about the same event with a friend who did not witness it. In the former case, context is available to both people whereas in the latter

there is not a context available for the hearer. Therefore, in There-and-Then condition, the speaker needs to explain the event in a more detailed way since the hearer does not have any idea about the event (Long, 2015). In Here-and-Now tasks, performers only need to use appropriate linguistic resources from context. There-and-Then tasks, on the other hand, refer to past events. This displacement in time and place requires performers to be more accurate and syntactically more complex (e.g. giving background information, using more sequence markers such as before, after, next) since their interlocutors cannot benefit from an available context. In other words, to be understood correctly, they need to ensure that all elements are reflected in their speech (Robinson, 1995). For this reason, There-and-Then tasks present more challenge to learners, so they are more complex than Here-and-Now tasks. In addition, in Here-and-Now tasks learners only need to describe a sequence of events before their eyes. However, There-and-Then tasks involve the use of memory for retrieving the event to be narrated. Finally, learners use present tense in Here-and-Now tasks whereas they use past tense in There-and-Then tasks which require the narration of an event displaced in time. The acquisition of present references are stated to be earlier than the acquisition of past references in L1 and L2 (Robinson, 1995). This situation poses another challenge for learners with low proficiency since learners need to be familiar with specific tenses and aspectual systems to refer to past events.

Most studies which changed the levels of task complexity along +/- Here-and-Now condition used narrative tasks (Gilabert, 2007a, 2007b; Gilabert, Baron & Llanes, 2009; Ishikawa, 2007; Iwashita et al., 2001; Robinson, 1995; Wang & Skehan, 2014). These narrative tasks require learners to tell a story in the presence

and absence of a context (e.g. a series of pictures, a video). Table 5 indicates a summary of studies which manipulated tasks along +/- Here-and-Now dimension.

In a study, Robinson (1995) used three similar picture strips depicting a character named Mr. Brown. 12 intermediate level English learners studying at a university were asked to narrate these pictures either in Here-and-Now or in There-and-Then conditions. In both conditions, participants were given one minute to understand the pictures before narrating the story. In Here-and-Now task, they were instructed to narrate the story in present tense while looking at the pictures. Differently, in There-and-Then task they were required to tell the story in past tense without seeing the pictures. In each condition, participants were provided with prompts either in present tense or in past tense according to the condition. Since There-and-Then condition was regarded as more complex than Here-and-Now condition, Robinson expected to find less fluent but more accurate and complex speech in these tasks. His hypothesis was confirmed for fluency, accuracy and lexical complexity. Participants' performances were less fluent but more accurate and lexically more complex under There-and-Then condition. However, there was not a difference between two conditions in terms of syntactic complexity.

Rahimpour (1997, cited in Robinson, 2001a) reported similar findings. He manipulated the tasks in +/- Here-and-Now and open/ closed task dimensions. The most complex task in his study resulted in less fluency but more accuracy with regard to error-free units. Target-like use of articles and structural and lexical complexity levels were not affected by the increase in complexity levels.

Following Skehan's (1998) predictions, Iwashita et al. (2001) examined the task complexity from a language testing perspective. Their aim was to discover a potential relation between task characteristics and performance conditions and

different levels of fluency, complexity, or accuracy of learners' performance. 193 adult English learners participated in study and they performed narrative tasks. These tasks were manipulated in four dimensions: a) +/-perspective, telling the story as if it happened to the participant or to someone else; b) +/-immediacy, telling the story either in Here-and-Now or in There-and-Then condition ; c) +/-adequacy, telling the story with complete or missing picture sets; d) +/-planning, whether participants had 0.5 or 3.5 minutes before telling the story. All participants were required to complete eight tasks, two conditions for each dimension. At the end of the study, the only significant effect was found for accuracy in the immediacy dimension. This finding confirmed Robinson's (1995) finding who reported an increase in accuracy levels in There-and-Then condition. There were no significant effects of performance conditions on fluency and complexity of participants' performances in any task dimension.

Table 5. Summary of Studies Manipulating Tasks along +/- Here-and-Now Variable

	L2	N	Participant	Proficiency	Tasks	Mode	A	F	C	L
Robinson (1995)	English	12	Adult	Intermediate	Narrative (picture-based)	Oral	+	-	=	+
Rahimpour (1997, cited in Robinson, 2001a)	English	20	Adult	Not reported	Narrative (picture-based)	Oral	+/-	-	=	X
Iwashita, McNamara & Elder (2001)	English	193	Adult	M = 493.1 institutional TOEFL	Narrative (picture-based)	Oral	+	=	=	X
Gilabert (2007a)	English	48	Adult	Lower intermediate	Narrative (picture-based)	Oral	+	-	=	-
Gilabert (2007b)	English	42	Adult	Lower-intermediate / upper intermediate	Narrative (picture-based)	Oral	+	X	X	X
Ishikawa (2007)	English	54	High school students	50 – 88 on Michigan English Placement Test	Narrative (picture-based)	Written	+/-	+	+/-	=
Wang & Skehan (2014)	English	72	Adult	M = 38.18/ M = 38.72 on TOEFL listening test	Narrative (video-based)	Oral	=	+/-	+	+

Note: L2 = second language; N = number of participants; A = accuracy, F = fluency, C = complexity, L = lexical complexity/lexis; + shows that complex task had a positive effect, - = shows that complex task had a negative effect, = shows a neutral effect, X = not investigated; +/- = shows mixed effects.

Gilabert (2007a) manipulated the tasks in two dimensions, +/- Here-and-Now and +/- planning, and examined their separate and combined effects on learners' performances. The operationalization of Here-and-Now and There-and-Then conditions were similar to Robinson (1995). 48 lower-intermediate level English learners performed four narrative tasks based on similar picture sequences under four different conditions. Following the Cognition Hypothesis, he hypothesized that There-and-Then conditions would result in dysfluency but greater accuracy and lexical and syntactic complexity. At the end of the study, researcher's hypothesis was partially confirmed. Here-and-now conditions led to more fluent performances regardless of planning time. Moreover, the accuracy levels were higher in There-and-Then task. However, manipulating task complexity along There-and-Then condition decreased the lexical complexity and did not affect the syntactic complexity. The predicted simultaneous increase in accuracy and complexity levels in There-and-Then task was achieved only when learners were provided with pre-task planning time.

The same year, Gilabert (2007b) conducted another study where he investigated the self-repair behaviors of learners. He used self-repair behaviors as a measure for accuracy and attempted to find out if learners have more self-repairs in cognitively complex tasks, if different task types have different impacts on self-repair behaviors, and if proficiency has an effect on self-repair behaviors of learners. 42 English learners from two universities in Spain participated the study. They had either low or high proficiency in English. Three different types of tasks, a narrative, an instruction-giving, and a decision making task was employed in the study. The narrative task was manipulated along Here-and-Now and There-and-Then, the instruction-giving task was manipulated along +/- elements, and the decision-making

task was manipulated along +/- reasoning. Results showed that more complex versions of the narrative task and the instruction-giving task led to higher levels of self-repair behaviors whereas for decision-making task self-repair behaviors did not show a significant difference. In addition, narrative task led to greater number of self-repairs than instruction-giving and decision making tasks. Finally, there was not a difference between low proficiency and high proficiency level participants in terms of their self-repair behaviors.

Ishikawa (2007) also examined the effects of changing tasks along +/- Here-and-Now conditions. Differently from previous studies, she employed written narrative tasks. 54 Japanese high school students were asked to write a story based on cartoon strips. Participants were divided into two groups; Here-and-Now group and There-and-Then group. The participants in the former group were asked to tell their stories in present tense and they were allowed to see the pictures while writing. The latter group, on the other hand, wrote their stories in past tense and they did not see the pictures during the writing phase. She expected to find greater accuracy, complexity, and writing fluency in the performances performed under There-and-Then condition. Except for lexical complexity, researcher's hypotheses for accuracy, syntactic complexity, and writing fluency were confirmed.

Wang & Skehan (2014) also examined the effects of task conditions on participants' oral performances. 72 English learners participated in the study. However, differently from previous studies, they used four silent videos rather than picture sequences. These tasks were manipulated in three dimensions: a) task structure, whether stories have a problem-solution structure or untied events; b) vocabulary difficulty; (c) time perspective, in Here-and-Now and There-and Then conditions. In Here-and-Now tasks, participants were instructed to narrate the story

in present tense while watching the video whereas in There-and-Then tasks they were told to narrate the story in past tense after watching the video. Researchers reported more complex language for tasks performed under There-and-Then condition. In addition, participants used less frequent words and had fewer pauses at clause boundaries in these tasks. Additionally, when There-and-Then tasks had a tight structure, they resulted in more accurate and complex performance than unstructured tasks.

In summary, the most consistent results across different studies were achieved by accuracy. All studies except for Wang & Skehan (2014) reported increased accuracy in There-and-Then tasks with at least one accuracy measure. Only in Wang and Skehan's study accuracy was found to be similar in simple and complex tasks. Furthermore, most of the studies manipulating the tasks in Here-and-Now/ There-and-Then dimensions indicated less fluent participant performances under There-and-Then conditions except for Iwashita et al. (2001) and Ishikawa (2007). In the former study, researchers reported no significant difference between simple and complex tasks in terms of fluency. In the latter study, greater task complexity led to higher fluency. However, in Ishikawa's study fluency was operationalized as written fluency which included different measures. In addition, Wang and Skehan reported mixed results for fluency. For syntactic complexity, on the other hand, only Ishikawa (2007) and Wang and Skehan (2014) reported greater scores under There-and-Then condition. However, there are two points that should be kept in mind: a) Ishikawa's study required participants to complete written narrative tasks; and b) Wang and Skehan's operationalization of Here-and-Now/ There-and-Then tasks was different from previous studies. Therefore, the results should be interpreted cautiously. Finally, except for Ishikawa's study which employed written

tasks, none of the studies investigating the oral L2 performance reported simultaneous increase in accuracy and complexity levels in There-and-Then tasks. Thus, all studies only partially confirmed the Cognition Hypothesis. Moreover, the absence of simultaneous increase in accuracy and complexity levels seem to support the arguments of Skehan (1998) who predicts the existence of trade-off effects between these two constructs.

In the current section, +/- Here-and-Now condition was considered as a variable affecting the cognitive demands of tasks and several studies examining its effects of on L2 performance were presented. The variables affecting the performative demands of tasks are claimed to provide more consistent results than variables which are related to cognitive demands of tasks (Skehan, 2016). Therefore, the next section will focus on online planning and provide information for some studies investigating this variable.

2.6.2 Online planning studies

Regardless of the modality of a task (written or oral), all tasks even the ones that can be completed automatically requires planning. People need to plan what to say or to write before saying or writing it. Moreover, people require more time to plan their L2 productions than their L1 productions due to lower levels of automaticity in L2 (Kormos, 2011). Therefore, planning phase is important for L2 learners especially in their oral productions.

Ellis (2005) divides planning into pre-task planning and online (within-task) planning. Figure 2 indicates the types of planning proposed by Ellis. Pre-task planning refers to the time given to learners before performing a task. Ellis further

divides pre-task planning into rehearsal (performing the same task before the main task) and strategic planning (providing learners planning time before the main task).

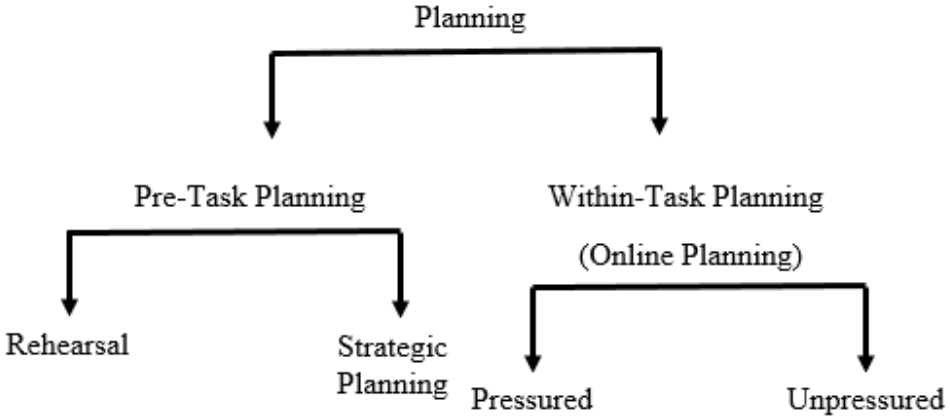


Figure 2. Types of planning
Source: [Ellis, 2005]

Skehan (1998) and Robinson (2011a) have similar predictions for the effects of pre-task planning time on learner performances. Based on the distinction that he makes on an exemplar-based system (i.e. learning as accumulation of chunks) and a rule-based system (i.e. learning is based on analysis and rules), Skehan asserts that pre-task planning time enables learners to get access to their rule-based system. Therefore, when provided with pre-task planning, the linguistic complexity in L2 performance increases. Robinson categorizes pre-task planning under resource-dispersing dimensions of a task and states that absence of planning time leads to greater fluency but lower accuracy and complexity in L2 performance. Similar to Skehan, Robinson also asserts an increase in learners' performances when provided with pre-task planning time.

Online planning, on the other hand, is related to the time given to learners to complete a task. Yuan and Ellis (2003) explains unpressured online planning as "the process by which speakers attend carefully to the formulation stage during speech planning and engage in pre-production and post-production monitoring of their

speech" (p.6). Neither Skehan (1998) nor Robinson (2011a) makes predictions regarding presence or absence of online planning. However, Ellis (2005) states online planning to be similar to pre-task planning in terms of its effects on learner performance.

Ellis and Yuan (2005) differentiate unpressured (careful) online planning from pressured online planning. In unpressured online planning learners are provided with sufficient time that allows them to pay attention to conceptualizing, formulating, articulating, and monitoring their speech. Attending to conceptualization stage may result in higher syntactic complexity levels since learners will have more time to create their pre-verbal messages. Besides, attending to formulation stage can increase learners' accuracy scores since they have more time to activate the lemmas and syntactic information for creating the surface structure of the message. However, allocating more attention to these stages may have a detrimental effect on fluency. Therefore, unpressured online planning may result in higher accuracy and complexity but less fluency. Pressured online planning, however, provides learners a limited time to produce their speech, so it puts learners under pressure. Therefore, learners may not allocate enough attention to conceptualization, formulation, and monitoring stages. Unlike unpressured online planning, pressured online planning does not improve accuracy and complexity levels, but it leads to more fluent speech.

While various studies have examined the effects pre-task planning time on learners' oral and written performances in terms of accuracy, fluency, and complexity (Ellis, 1987; Foster & Skehan, 1996; Gilabert, 2007a; Iwashita et al. ,2001; Mochizuki & Ortega, 2008; Ortega, 1999; Skehan & Foster,1997; Yuan & Ellis, 2003) only a few studies controlled for the online planning (Ahmadian & Tavakoli, 2010; Ahmadian, Tavakoli, & Dastjerdi, 2015; Ellis & Yuan, 2005; Wang,

2014; Yuan & Ellis, 2003). Table 6 shows the summary of studies which manipulated the online planning variable of tasks.

Hulstjin and Hulstjin (1984) investigated the correct use of Dutch word-order rules of L2 Dutch learners in a story retelling task manipulated along two variables: time pressure (present/absent) and focus of attention (information/grammar). 32 Dutch learners participated in the study and they retold 68 listening texts under four conditions: IF (information/fast), IS (information/slow), GF (grammar/fast), and GS (grammar/slow). At the end of the study, participants produced shorter responses under time pressure. In addition, time pressure had more effect when the attention was on grammar. Therefore, this study showed that unpressured online planning can be helpful only when attention is directed to form.

Table 6. Summary of Studies Manipulating Tasks along +/- Online Planning Variable

	L2	N	Participant	Proficiency	Tasks	Mode	A	F	C	L
Hulstjin and Hulstjin (1984)	Dutch	32	Adult	Not reported	Narrative based on listening	Oral	– shorter responses – more effect on grammar			
Ellis (1987)	English	17	Adult	Early intermediate	Narrative (picture-based)	Oral/ Written	– more accurate use of regular past tense verbs.			
Yuan and Ellis (2003)	English	42	Adult	100–120 (max. 150) on HEB Examination	Narrative (picture-based)	Oral	+	=	+*	=
Ellis and Yuan (2005)	English	42	Adult	100 – 120 (max. 150) on HEB Examination	Narrative (picture-based)	Oral	+	=	+	=
Ellis and Yuan (2005)	English	42	Adult	100 – 120 (max. 150) on HEB Examination	Narrative (picture-based)	Written	+	=	+	=
Ahmadian and Tavakoli (2010)	English	60	Adult	Intermediate	Narrative (video-based)	Oral	+	-	+	X
Wang (2014) (Online planning)	English	77	Adult	540 - 630 on TOEFL and 6 - 7.5 on IELTS	Narrative (video-based)	Oral	=	=	=	=
Wang (2014) (Watched online planning)	English	77	Adult	540 - 630 on TOEFL and 6 - 7.5 on IELTS	Narrative (video-based)	Oral	+	-	+	=
Ahmadian, Tavakoli and Dastjerdi (2015)	English	60	Adult	Intermediate	Narrative (video-based)	Oral	+	-	+	X

Note: L2 = second language; N = number of participants; A = accuracy, F = fluency, C = complexity, L = lexical complexity/lexis; + shows that unpressured online planning had a positive effect, - = shows that unpressured online planning had a negative effect, = shows a neutral effect, X = not investigated.

* Careful online planning had a positive effect only for syntactic complexity not for syntactic variety.

Ellis (1987) investigated the accurate use of past simple forms in oral and written narrative tasks differing in online planning and pre-task planning. 17 English learners with intermediate level proficiency participated in the study and they completed both written and oral narrative tasks under three conditions: a) planned writing, b) planned speech, and c) unplanned speech. Task 1 was a written task in which participants wrote a story based on a picture set in an hour (+pre-task planning and +online planning); Task 2 required learners to retell the story used in Task 1 orally (+pre-task planning and -online planning); and in Task 3 learners were instructed to tell a story based on a different set of pictures after seeing them for 2 minutes (-pre-task planning and -online planning). In Task 1, which provided both pre-task planning and unpressured online planning opportunities, participants used regular past tense verbs more accurately than in other tasks. Moreover, there was no difference between Task 2 and Task 3. Therefore, unpressured online planning was stated to have an effect on the correct use of regular past tense verbs. This confirms Skehan's (1998) predictions for the positive effects of planning on the rule-based system. Since the structure of regular verbs are based on rules, participants used them more accurately in the presence of planning. However, the use of irregular verbs can be stated to be exemplar-based so their use is not affected by planning. However, the findings of this study should be treated carefully since the difference in modality between Task 1 and tasks 2 and 3 may have influenced the results. While the first task required participants to write a story, in tasks 2 and 3 participants narrated their stories orally. Crookes (1989) also points out that difference between Task 1 and the other two tasks stating that in this study the difference might have been caused either from the presence of planning or from the difference in modalities.

Yuan and Ellis (2003) compared how pre-task planning and unpressured online planning affects L2 learners' fluency, accuracy, syntactic complexity, and lexical complexity levels. 42 English learners involved in the study and they were divided into three groups: a) no planning (NP), b) pre-task planning (PTP), and c) unpressured online planning (UOP). They were asked to tell a story based on pictures. In pre-task planning condition, learners were given 10 minutes before telling their stories and they were required to produce four sentences for each picture and to complete their stories in 5 minutes. In unpressured online planning, participants started their stories immediately after 50 seconds however they were given unlimited time for the completion. In no planning condition, participants started their stories after 50 seconds and they were given only 5 minutes to complete. When compared to the NP group, UOP led to more accurate and grammatically complex speech. Participants in UOP group were less fluent than NP group but this difference was not significant. Besides, the two groups were similar in terms of syntactic variety and lexical variety. Finally, UOP group were significantly less fluent than PTP group. Two years later Ellis and Yuan (2005) investigated the effects of unpressured and pressured online planning in writing and speaking tasks. The study included 42 English learners divided into three groups. The first group completed an oral task with pressured online planning and a writing task with unpressured online planning. The second group completed an oral task with unpressured online planning and the third group completed a writing task with pressured online planning. Under all conditions, participants were given 50 seconds to understand the pictures. In unpressured online planning conditions, participants were provided with unlimited time to complete their stories whereas in pressured online planning conditions, they were given 5 minutes for oral task and 17 minutes

for writing task. At the end of the study, there was no significant differences between unpressured and pressured online planning groups in terms of fluency and lexical variety regardless of the mode of the task. However, greater syntactic complexity and accuracy were achieved under unpressured online planning condition in both oral and writing tasks.

In another study, Ahmadian and Tavakoli (2010) investigated the effects of manipulating tasks along unpressured/pressured online planning and task repetition on learners' performance. Four conditions were included in the study: unpressured online planning without task repetition; unpressured online planning with task repetition; pressured online planning without task repetition; and pressured online planning with task repetition. 60 participants were assigned to these four groups randomly. Participants in each group were instructed to a narrate story based on a silent video. Unpressured/pressured online planning variable was operationalized following Yuan and Ellis (2003). For task repetition variable, learners repeated the same task with one week interval. In line with their hypotheses, the researchers reported significantly higher accuracy and syntactic complexity and less fluency for unpressured online planning groups. Additionally, this decrease in fluency was compensated when tasks were repeated. In other words, simultaneous manipulation of unpressured online planning and task repetition led to more accurate, syntactically complex, and fluent speech. Five years later, Ahmadian, Tavakoli, and Dastjerdi (2015) conducted a similar study examining the effects of unpressured/pressured online planning. This time, they also investigated the effects of task structure in addition to online planning. 60 intermediate level English learners were distributed to four groups: structured task and unpressured online planning (UOLP + ST), structured task and pressured online planning (POLP + ST), unstructured task and

unpressured online planning group (UOLP + UST), and unstructured task and pressured online planning group (POLP + UST). Similarly to previous study, participants told stories based on silent videos. This study supported the findings of the previous study for accuracy, syntactic complexity, and fluency. Unpressured online planning increased accuracy and syntactic complexity of learners' performance whereas it decreased their fluency levels. Moreover, structured tasks improved the decreased fluency levels caused by the pressured online planning condition.

Wang (2014) conducted a study in which the tasks were manipulated along pre-task planning, online planning and task repetition. 77 English learners participated in the study and they were assigned to a control group and five experimental groups. The experimental groups included two groups with two pre-task planning conditions (watched/ watched strategic planning), two online planning conditions (online planning/ watched online planning), and task repetition. All participants told a story based on a video which was the same for all groups. Online planning group watched the slowed version of the video and they narrated their stories while watching it. Watched online planning group, on the other hand, first watched the video and then narrated the story while watching the slowed version of it. Slowed version of the videos provided extra time for online planning to participants. Therefore, these groups had unpressured online planning during the experiment. At the end of the study, all experimental groups were compared to the control group. Online planning group did not differ significantly from the control group in terms of complexity, accuracy, fluency, and lexical complexity. However, watched online planning group had greater accuracy and complexity but less fluency than the control group.

In short, almost all studies manipulating the online planning variable reported increased accuracy and syntactic complexity under unpressured online planning conditions (Ahmadian & Tavakoli, 2010; Ahmadian, Tavakoli, & Dastjerdi, 2015; Ellis, 1987; Ellis & Yuan, 2005; Yuan & Ellis, 2003). Although Wang (2014) also reported more accurate and syntactically complex speech in Watched Online Planning condition whose operationalization of online planning is similar to unpressured online planning in previous studies, Online Planning group which performed the storytelling task while watching the slowed version of the video did not present a significant difference for these constructs. This can be due to the impacts of different operationalization on L2 speech production system. While the former operationalization of online planning (e.g. Watched Online Planning group in Wang's study) taps into both conceptualization and formulation stages, the latter (e.g. Online Planning group in Wang's study) provides attention only to formulation stage (Wang, 2014). Therefore, it can be said that the positive effects of online planning can only be observed when learners are provided with an opportunity to understand the story before telling it. In other words, the operationalization of unpressured online planning should enable focusing on both conceptualization and formulation stages in L2 speech production to indicate positive effects on L2 performance. Moreover, the studies presented mixed results for how unpressured online planning affect fluency. Three studies (Ahmadian & Tavakoli, 2010; Ahmadian, Tavakoli, & Dastjerdi; and Wang, 2014) reported a decrease in fluency levels whereas two studies (Ellis & Yuan, 2005; Yuan & Ellis, 2003) reported insignificant differences. Finally, none of the studies reported a difference in terms of lexical complexity.

Overall, as in line with Skehan's (2016) arguments, the performative demands of tasks (i.e. resource-directing features), in this case +/- Online Planning, present more compatible findings for their effects on L2 performance, particularly on accuracy and complexity. The studies increasing the cognitive complexity, however, provide somewhat inconsistent findings except for accuracy. These studies do not confirm or partially confirm the predictions of Robinson's Cognition Hypothesis (2001a).

To investigate Skehan's (2016) claims regarding the impact of resource-directing and resource-dispersing variables in tasks and to examine the probable combined effects of these two-way manipulations suggested by Robinson (2011a) the current study manipulates the two variables of task complexity. These are +/- Here-and-Now and +/- Online Planning. Firstly, it aims to examine how the manipulation of these two variables affect L2 oral performance, especially accuracy and syntactic complexity levels of L2 English learners. Secondly, differently from previous studies, the participants in the current study includes advanced English learners who have taken some English communication courses which allowed them to improve their English speaking skills. Therefore, the study examines if the positive effects of Unpressured Online Planning and There-And-Then conditions reported in previous studies can be observed with high proficient English learners. Finally, it tries to find empirical support for either of the two distinct hypotheses, namely the Limited Attentional Capacity Hypothesis and the Cognition Hypothesis, regarding the effects of task complexity.

In the light of what has been discussed above the current study attempts to answer the following research questions:

1. How does manipulating task complexity along Here-and-Now/ There-and-Then affect the syntactic complexity levels of L2 learners' oral performance?
2. How does manipulating task complexity along Here-and-Now/ There-and-Then affect the accuracy levels of L2 learners' oral performance?
3. How does manipulating task complexity along pressured/ unpressured online planning affect the syntactic complexity levels of L2 learners' oral performance?
4. How does manipulating task complexity along pressured/ unpressured online planning affect the accuracy levels of L2 learners' oral performance?
5. Are there any interaction effects between task type and online planning in relation to syntactic complexity and accuracy levels of L2 learners' performance?

The corresponding hypotheses based on previous research and theories for each research question are given below:

1. The Cognition Hypothesis predicts an increase in syntactic complexity levels in more complex tasks along resource-directing dimension. Therefore, in this study the syntactic complexity levels is expected to increase significantly in There-and-Then task.
2. In his Cognition Hypothesis, Robinson (2011) argues for an increase in the accuracy levels of learners in There-and-Then tasks. Thus, the participants are expected to be more accurate in There-and-Then task.
3. Robinson (2011) claims that tasks which are simple along resource-dispersing dimension results in higher syntactic complexity levels.

Therefore, participants in unpressured online planning group are predicted to have higher syntactic complexity scores than the ones pressured online planning group.

4. Tasks which are simple along resource-dispersing dimension are predicted to lead to more accurate performances (Robinson, 2011). Therefore, participants in unpressured online planning group are expected to have higher accuracy scores than participants in pressured online planning group.
5. Robinson (2005) argues that learners perform better in tasks with increased cognitive but decreased performative demands. Therefore, the participants in the unpressured online planning group performing the There-and-Then task is expected to have the highest syntactic complexity and accuracy scores. In addition, drawing on Gilabert's (2007a) predictions, the lowest performance scores are expected in There-and-Then task performed by the participants in pressured online planning group since simultaneous increases in cognitive and performative demands of tasks are claimed to have negative effects on learners' performances.

CHAPTER 3

METHODOLOGY

3.1 Introduction

This chapter describes the data collection methods and procedure. It provides information about the participants, research design, materials, procedure, and the measures used in assessing learner performance.

3.2 Participants

Sixty-four participants majoring in English Language Education (ELT) in an English-medium public university in Turkey participated in the study. Their ages ranged from 19 to 24. They had similar backgrounds in terms of their exposure to English. They all started learning English when they were in fourth grade in primary school and all of them studied English in their high schools. None of the participants lived in an English-speaking country before. They all had a pass mark from an institution-based English proficiency exam. In this test students complete two reading, two listening, and two writing tests and the minimum point accepted as successful is equivalent to 550 in TOEFL PBT (Paper-based), 79 in TOEFL IBT (Internet-based), and 6.5 in Academic IELTS. Moreover, after passing the institutional proficiency exam, in their first year these participants have taken several courses for improving their oral and written skills.

3.3 Research design

This study has a 2x2 mixed design with task type (Here-and-Now/ There-and-Then) as within-participant conditions and online planning (unpressured online

planning/pressured online planning) as between-participant conditions. The task order and the pictures used were counterbalanced in each condition to prevent task order and picture effects, respectively. Table 7 shows the design of the study.

Table 7. The Design of the Study

	Resource-Dispersing (Between-participants)	
	Unpressured Online Planning Group	Pressured Online Planning Group
Resource-Directing (Within-Participants)	Here-and-Now	Here-and-Now
	There-and-Then	There-and-Then

Note: Total N = 64. Participants completed two tasks either under unpressured online planning condition, or under pressured online planning condition.

3.4 Materials

3.4.1 Tasks

Each participant engaged in two tasks, one simple (Here-and-Now) and one complex task (There-and-Then) and all participants were required to complete the tasks orally. There were two main reasons for choosing oral tasks: a) most of the previous studies manipulating the tasks along Here-and-Now and There-and-Then dimensions employed oral tasks so it is easier to compare the results of the study with the results of previous studies; b) modality is a factor affecting learners' performances and speaking is claimed to put learners under more pressure than writing (Ellis, 1987) therefore the possible effects of the presence or absence of online planning can be observed better with oral tasks.

Monologic narrative tasks were considered to be more suitable for the aims of the present study for the following reasons: a) these tasks require more precise language so they lead to more complex performance than personal tasks or decision-making tasks (Foster & Skehan, 1996), b) monologic narrative tasks do not trigger clarification requests and confirmation checks interrupting the performance as in

interactive tasks (Yuan & Ellis, 2003) so that they do not lower the syntactic complexity of learners' performance (Robinson, 2001a, 2007a, 2011a), c) most participants are familiar with narrating a story in their first language therefore narrating a story in English will not be an unfamiliar task for them.

All participants narrated a story based on a series of six picture frames (Heaton, 1966, 1975) (See Appendix A). The first picture sequence depicts a boy and a girl going on a picnic without knowing that their dog was hiding in the picnic basket and eating their food (the Picnic task). The second picture is about a boy carrying several boxes without knowing that he actually dropped one of the boxes. In the story, he is chased by a man who wants to give him his box back (the Chase task). Table 8 shows the features of two picture strips.

Table 8. The Features of the Picture Strips

Picture	Characters	Storyline	Storyline Complexity	Story structure	Code complexity
The Picnic	4	A girl and a boy decide to go on a picnic. They prepare their picnic basket and their mother helps them. While they are looking at a map with their mother, their dog gets into the picnic basket. They do not realize it until they arrive the picnic area. When they open their basket, they see that their dog has eaten everything in their basket.	Foreground and background events	Tight structure	Easy vocabulary related to real life activities with cause and effect (girl, boy, mother, dog, picnic, basket, picnic area, food, eat)
The Chase	2	A boy gets off a bus and he carries several boxes under his arms. While walking, he drops one of the boxes. He does not realize it and keeps walking. A man takes the box and runs after the boy. The boy gets afraid of the man at first and starts to run. When the man catches up with the boy, the boy understands that the man is following him to give his box back.	Foreground and background events	Tight structure	Easy vocabulary related to real life activities with cause and effect (boy, man, bus, get off, box, run, catch)

Two picture strips were matched in terms of their storyline complexity and task structure. Storyline complexity is characterized in terms of foreground and background events. According to Tavakoli and Foster (2008) foreground events are those "that generally move time forward, whereas background elements are there to elaborate on, explain, or evaluate the events in the foreground" (p. 444). Stories that involve both foreground and background events are argued to lead to more subordination and complex language because English speakers use some subordination conjunctions such as because, when, before, after, as, and while to explain what is happening in the background (Tavakoli & Foster, 2008). Moreover, task structure plays an important role on the accuracy and fluency levels. Stories with a tight structure, where changing the order of the pictures and events are not possible, lead to more accurate and fluent speech since tight structure "eases the processing burden" (Tavakoli & Foster, 2008, p. 445). Therefore, both picture strips used in the present study included both foreground and background events and they had a tight structure.

Moreover, specific vocabulary related to a subject is claimed to affect learners' performances negatively if learners are not familiar with these words (Sasayama, 2015). However, the code complexity of both pictures used in the current study were relatively easy. They both included words related to real life events (e.g. boy, girl, mum, picnic, basket, dog, hill, sun for the Picnic task; boy, man, bus, get off, moon, dark, chase for the Chase task).

The pictures differed only in terms of the number of characters in the story. The Picnic task included more characters than the Chase task. Although +/- few elements is categorized as a factor affecting task complexity (Robinson, 2001a), a study conducted by Sasayama (2015) reports that in picture-based narrative tasks

learners' performances as well as their perceptions of task complexity depend more on the clarity of the storyline, the code complexity, and the similarities between characters (e.g. two similar characters doing similar actions such as fishing) rather than the number of the characters that a story includes.

3.4.2 Questionnaires

Participants were given a background questionnaire for assessing their language background, experiences and opinions related to English use (See Appendix B). In addition, after completing each task (one simple and one complex), participants were given a question where they stated their perceived difficulty of the tasks by choosing from a 6-point Likert scale (See Appendix C). The aim of this question was to observe whether participants perceived There-and-Then task to be more difficult than Here-and-Now task as suggested by Robinson (1995, 2001a, 2007a, 2011a).

3.5 Procedure

Data collection took place in a single individual session. Before the experiment, participants were informed that they would be recorded while narrating their stories. They were told that the recordings would be anonymous and confidential. Firstly, they completed the background questionnaire and the consent form. Then, participants completed two oral narrative tasks, one simple and one complex task.

After seeing the picture strips for one minute, depending on their group, they were given either unlimited time (unpressured online planning group) or limited time (pressured online planning group) to tell their stories. After completing each task, students were asked to choose the most suitable number on a 6-point Likert scale showing the difficulty of the task.

All instructions were provided in participants' L1 to ensure their understanding of what they need to do in the tasks. They were also given additional instructions and explanations if needed before the beginning of each task.

3.6 Operationalization of task complexity

Each participant was instructed to narrate two stories, one simple and one complex. The complexity of tasks was manipulated along Here-and-Now/There-and-Then condition. This condition has been operationalized following Robinson (1995). In Here-and-Now condition, participants were shown each picture strip for one minute which was enough to understand the storyline. After one-minute, participants narrated the story in present tense. In addition, they were allowed to see the pictures while telling the story. In There-and-Then condition, participants were shown the picture strip for one minute. Differently from Here-and-Now condition, they were instructed to tell the story in past tense without seeing the pictures during the storytelling phase. In both conditions, they were asked to narrate the stories as detailed as possible, as if they were telling them to one of their friends who had never seen these pictures.

Unpressured online planning and pressured online planning have been operationalized following Yuan and Ellis (2003). Under unpressured online planning condition participants started their stories immediately after seeing the pictures for one minute and they were given unlimited time to formulate and monitor their speech while completing the tasks. Participants performing the tasks under pressured online planning were also shown the picture strips for one minute. However, as in previous online-planning studies (Ahmadian & Tavakoli, 2010; Ellis &Yuan, 2005; Yuan & Ellis, 2003), participants were given three minutes to narrate their stories.

Prior to main study, a pilot study was conducted to investigate the operationalization of conditions as well as the instructions given to the participants in the study. The following section provides some details about the pilot study.

3.6.1 The pilot study

The materials were pilot tested in late spring 2017 with 14 first year students studying at the same English-medium public university in Turkey. The participants completed two tasks with one week interval. Two researchers conducted the study in different classrooms at the same time and the participants were randomly assigned to two groups, unpressured online planning group and pressured online planning group.

At the end of the pilot sessions, some modifications were made in the operationalization of tasks and conditions as well as in the task instructions (see Appendix D and E for task instructions). The three minute time given to do the participants in pressured online planning group was decreased to one minute and thirty seconds since most participants spent less than two minutes on task. Finally, since each participant used different names for the characters in the pictures, we gave names to the main characters so that participants in the real study used these names when telling their stories. Moreover, to increase the time pressure on participants, participants in pressured online planning group were instructed to check their remaining time on a chronometer provided to them while telling their stories.

3.7 Measures

The recordings of the participants in two tasks were first transcribed by the researcher and her colleague who is also an MA student in English Language Education. The transcribed data were assessed in terms of accuracy and syntactic

complexity after coding the data into units of analysis. In the following sections, the units of analysis and the measures of accuracy and syntactic complexity are explained in detail.

3.7.1 The units of analysis

Following Foster, Tonkyn, and Wigglesworth (2000), the data were divided into the Analysis of Speech units (AS- unit) which are reported to be more reliable and consistent indicators of spoken data. Foster et al. (2000) defines an AS-unit as "a single speakers' utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clauses(s) associated with either" (p. 365). In AS-unit, an independent clause refers to a clause with a finite verb and a subordinate clause includes clauses either with a finite verb or with a non-finite verb and at least one other clausal element. Although there are other units of analysis such as T-unit (mainly syntactic unit) and c-unit (mainly semantic unit), AS unit is suggested to be better for oral data since when coding the data into AS- units, intonation and pause information as well as syntactic criterion are taken into consideration. For example, some subordinate conjunctions such as because may function as an ellipted version of an independent clause (e.g. I say this because…). While such an ellipted version is always coded as a dependent clause in T-units, it can be regarded as an independent clause in AS-units (See Appendix E for an example AS-unit coding).

3.7.2 The length of time

To ensure that participants followed the instructions given to them, total time that each participant spent for task completion was measured and the two groups, namely unpressured online planning and pressured online planning groups, were compared.

This measure was similar to the one that Yuan and Ellis (2003) employed in their studies. Participants in unpressured online planning group were expected to spend more time on tasks since they were given unlimited time to tell their stories. Participants in pressured online planning group, on the other hand, was predicted to perform the tasks in a shorter time since they were given only 90 seconds for each task.

3.7.3 Dependent measures

In the current study, learners' performances were assessed through various syntactic complexity and accuracy measures. All measures were calculated from the pruned narratives. In other words, only the final version of all repetitions, false starts, and self-corrections were taken into account when measuring both the syntactic complexity and accuracy of learners' performances. The following two sections provide detailed information about the syntactic complexity and accuracy measures employed in the study.

3.7.3.1 Syntactic complexity

Following Norris and Ortega (2009), syntactic complexity was operationalized as a multi-componential construct composed of several sub constructs. Therefore, three distinct yet complementary measures, overall complexity, subordination, and phrasal complexity, were employed to assess the syntactic complexity of participants' performances.

Overall complexity was defined as the mean length of AS-unit and it was calculated by dividing the total number of words to total number of AS-units. This measure reflects the global syntactic complexity of learners' performances.

Subordination was defined as the ratio of clauses to AS-units and it was calculated by dividing the number of clauses to AS-units in each narrative. Subordination is the most widely used measure for assessing the syntactic complexity in many previous studies (Ellis & Yuan, 2005; Gilabert, 2007a; Iwashita et al., 2001; Kuiken & Vedder, 2007; Michel et al., 2007).

Phrasal complexity was defined as the mean length of clause and it was calculated by dividing the total number of words to the total number of clauses. This measure is especially useful for assessing the performance of advanced learners since they rely less on subordination but more on modifications and nominalizations in their performances (Norris & Ortega, 2009).

3.7.3.2 Accuracy measures

Accuracy was assessed with three general measures used in previous studies. These measures are errors per 100 words and the ratio of error-free AS-units.

Errors per 100 words was calculated by dividing the total number of errors by total number of words, and then multiplying by 100. This measure is argued to be more precise than measures which explores the error-free units of performance since it considers the errors in learners' performances (Gilabert, Baron, & Levkina, 2011).

The ratio of error-free AS units was calculated by dividing the total number of error-free AS-units to total number of AS-units. This measure is widely used in task complexity studies (Kuiken & Vedder, 2007; Michel et al. 2007; Revesz, 2011).

The ratio of error-free clauses was calculated by dividing the total number of error-free clauses to total number of clauses. Similar to error-free AS units, this measure has also been frequently employed in previous task complexity studies (Albert, 2011; Kormos & Trebits, 2011, 2012).

3.8 Data analysis

The data was transcribed by the researcher and her colleague who is an MA student in English Language Education. Some participants commented on the events in the story or signaled the ending at the end of their stories. These parts were regarded to be unnecessary or irrelevant and excluded from the data analysis. The following sentences are examples for the parts that were not included in the data coding process:

> Example 1: I do not understand why the man is going so far because I think still he lives in outside the city.
>
> Example 2: I guess it is all I can say.

The data was coded by the researcher and a second coder, independently. The second coder who was also an MA student in English Language Education coded 50% of the data. Before coding, she completed a training session with the researcher where she was given information about the coding guidelines (see Appendix F for the detailed coding guidelines). Two coders decided not to count the formulaic phrases such as "I think", "I mean", "You know", "I guess" as clauses since most participants used these phrases as fillers in their speeches. An inter-coder reliability analysis was conducted for the 50% of the data. The number of clauses, the number of AS-units, syntactic complexity measures, and accuracy measures were checked for their reliability with a Pearson Correlation and all measures reached more then %85 reliability (see Appendix G for the detailed list). After checking the reliability, the data was entered to SPSS version 21.0.

Before conducting any analysis, the data were checked for the assumptions through some statistical tests and graphics such as skewness and kurtosis values, normality tests, histograms, Q-Q plots, and box plots. At the end of the checking

process, some variables for separate groups showed non-normal distribution. When the data are not normally distributed, non-parametric test are suggested to be used (Field, 2009). However, these tests decrease the statistical power of the test and results in a higher possibility of making a Type II error (Huck, 2004). Therefore, the data was transformed with the log 10 transformation. According to the Kolmogorov-Smirnov test, transformed variables showed normal distribution. In addition, some outliers were detected in the data however these outliers were kept in the analyses for two main reasons. Firstly, these outliers were not errors. They showed the real data gathered from the sample. Secondly, some preliminary analyses were conducted with and without outliers and these analyses showed no great differences.

The data was examined with descriptive statistics, confidence intervals, effect sizes, and inferential statistics. For inferential statistics, a series of two-factor mixed design ANOVAs were conducted with online planning (unpressured and pressured) as a between-subject and task type (Here-and-Now and There-and-Then) as a within-subject variable. Corrected Bonferroni was applied to p values to avoid the possibility of making a Type 1 error.

The p value calculated in inferential statistics has a dichotomous nature, in other words, it only interprets if the effect of a variable is statistically significant or not (Cumming, 2012; Larson- Hall & Plonsky, 2015; Plonsky & Oswald, 2014). However, Cumming (2012) states that "small [p-value] does not guarantee that an effect is large or important. A tiny effect can be highly statistically significant if the experiment is sufficiently large, or a small experiment can find a large effect that's not statistically significant" (p. 28-29). Therefore, Cumming (2012) and Larson-Hall and Plonsky (2015) recommend reporting the confidence intervals and effects sizes as well as p values since they provide more information regarding the data.

According to Field (2009) "a confidence interval for the mean is a range of scores constructed such that the population mean will fall within this range in 95% of samples" (p. 48). In addition, he argues that using 95% confidence interval graphs for each condition is very beneficial for comparing the mean scores. If the mean scores of different conditions do not overlap, and the mean score of one condition exceeds the upper 95% confidence interval of the other condition, it is likely to have a significant difference between these conditions.

Effects sizes, on the other hand, indicate the practical significance of a difference or a relationship (Plonsky & Oswald, 2014). This study employs Pearson's r for calculating the effect size for main effects and interaction effects of tasks and groups. Following Field (2009), the effect size is calculated as:

$$r = \sqrt{\frac{F(1,\ df_R)}{F(1,\ df_R) + df_R}}$$

Although Cohen (1988, cited in Plonsky and Oswald, 2014) considers .10 as small, .30 as medium, and .50 as large correlation coefficients, Plonsky and Oswald (2014) state that these criteria are not suitable for L2 research. They propose new benchmarks for Pearson's r and they consider .25 as small, .40 as medium, and .60 as large effect sizes.

CHAPTER 4

RESULTS

4.1 Introduction

The data collected from two groups were examined with descriptive statistics (i.e. means and standard deviations), confidence intervals, and effects sizes. After the testing assumptions and transforming the non-normally distributed data, a series of two-factor mixed ANOVAs were conducted. This chapter firstly presents the results for how participants perceived the complexity of tasks, then it reports how much time each group spent on two tasks in total. Finally, it indicates the descriptive statistics, confidence intervals, effect sizes and inferential statistics for complexity and accuracy measures.

4.2 The perceived difficulty of tasks

After completing each task, students in both groups were instructed to answer a question regarding the perceived difficulty of the task. They were provided with a 6-point Likert-scale, 1= very easy and 6 = very difficult and were asked to choose the number that shows the difficulty of the task. The results for perceived task difficulty are shown in Table 9.

Table 9. Difficulty Perception of Two Tasks

	Simple Task (Here-and-Now)					Complex Task (There-and-Then)				
	N	Min.	Max.	Mean	SD	N	Min.	Max.	Mean	SD
Unpressured	32	1.00	4.00	1.91	0.86	32	1.00	4.00	2.34	0.90
Pressured	32	1.00	5.00	2.38	1.13	32	1.00	5.00	2.44	0.98

As can be seen in the Table 9, all participants in both groups perceived There-and-Then task to be more difficult than Here-and-Now task. The difference between tasks

were larger for unpressured online planning (UOP) group. This difference was rather marginal for pressured online planning (POP) group. Figures 3 and 4 shows the 95% confidence intervals for two groups in two tasks. According to these graphics, there is an overlap between the confidence intervals of two tasks in POP group. Moreover, the effect size for Here-and-Now and There-and-Then tasks is r = -.02. These close confidence intervals and the very small effect size indicate that the participants in POP group perceived the difficulty of Here-and-Now and There-and-Then tasks to be similar.

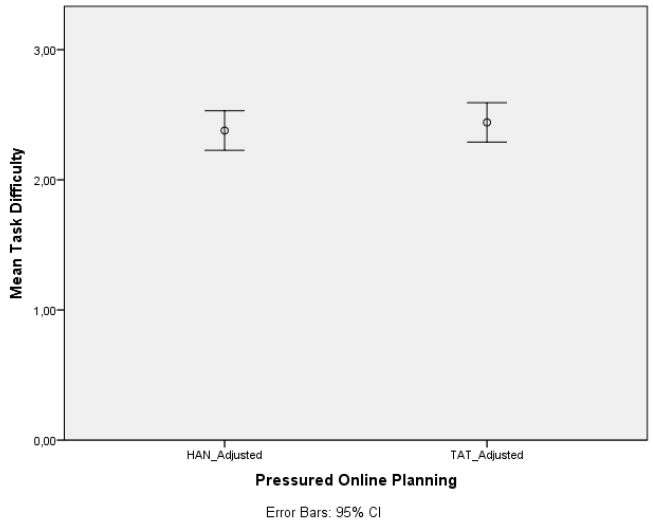


Figure 3. Task difficulty ratings of Pressured Online Planning group for two tasks with 95% confidence intervals.

In Figure 4, on the other hand, the mean task difficulty of There-and-Then task exceeds the upper 95% confidence interval of Here-and-Now task. The Pearson's r is found to be r = -.23. Therefore, UOP group can be said to perceive There-and-Then task to be more difficult than Here-and-Now task, although the effect size is small.

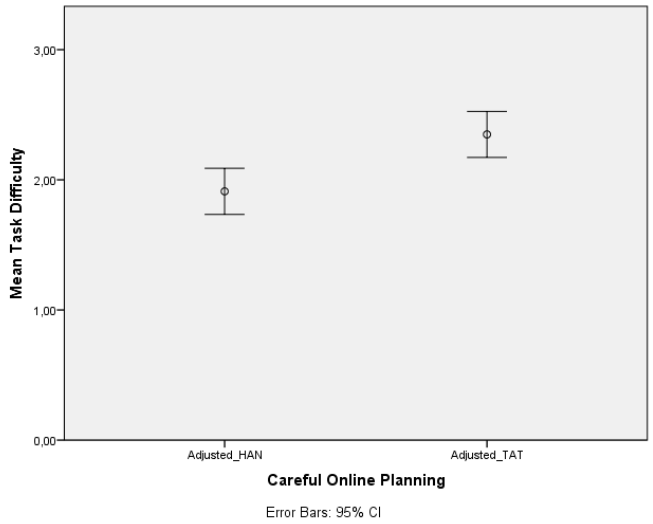


Figure 4. Task difficulty ratings of Unpressured Online Planning group for two tasks with 95% confidence intervals.

In addition to the difference between two tasks, what Table 9 and Figures 3 and 4 show is that regardless of the groups, most participants considered both tasks to be easy or moderately difficult. In both tasks, the minimum value was 1.00 which referred to a very easy task. In addition, the complex task was not regarded as very difficult by any of the participants.

4.3 The length of time

As it has been stated in the methodology section, UOP group was allowed to take as much time as they needed to complete their stories. POP group, on the other hand, were given only 90 seconds to formulate and monitor their stories. To see whether two groups behaved according to the instructions provided to them, the length of time spent on two tasks by two groups were checked. According to the results, UOP group (M = 164.31, SD = 72.17) spent more time on tasks then POP group (M = 148.22, SD =35.48). However, when the 95% confidence intervals were checked, the confidence intervals for two groups seemed to overlap and their mean scores

were close to each other (see Figure 5). Besides, the Pearson's r was found to be *r* = .14. Therefore, the difference between groups had a small effect size.

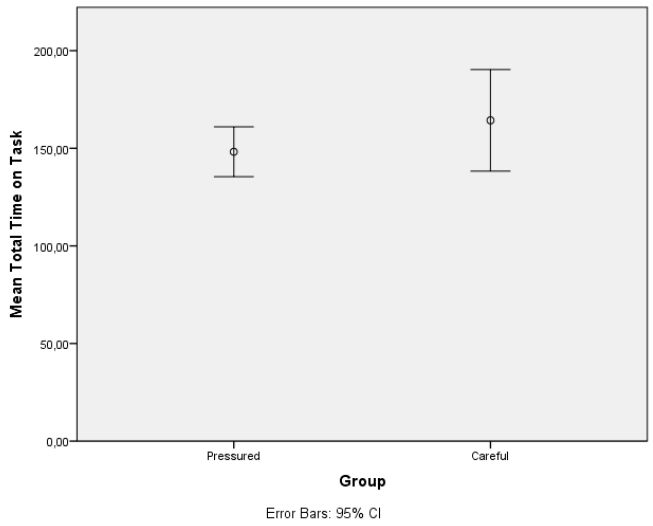


Figure 5. Total time spent on task by two group with 95% confidence intervals.

To see whether this small difference between two means were statistically significant, an independent samples t-test was conducted. See Table 10 for the test results. The t-test showed that the difference between two groups was not statistically significant, $t(53.57) = -.702$, $p = .49$. This means that although the two groups were given different instructions in terms of online planning, UOP group did not spend significantly more time on tasks than POP group.

Table 10. Statistics for the Length of Time (Seconds) Spent on Task

Length of Time (seconds)								
	N	Min.	Max.	Mean	SD	t	df	Sig.
Unpressured	32	80.00	385.00	164.31	72.17	-.702	53.57	.486
Pressured	32	83.00	230.00	148.22	35.48			

4.4 Dependent variables

4.4.1 Syntactic complexity

The current study investigated the syntactic complexity in terms of overall complexity, complexity by subordination, and phrasal complexity. Table 11 indicates the descriptive statistics of the two groups in two tasks.

Table 11. Descriptive Statistics for Complexity Measures

Complexity	Unpressured						Pressured					
	Here-and-Now			There-and-Then			Here-and-Now			There-and-Then		
	Mean (SD)	Min.	Max.	Mean (SD)	Min.	Max.	Mean (SD)	Min.	Max.	Mean (SD)	Min.	Max.
Overall Complexity	9.34 (1.90)	6.60	12.63	9.11 (1.66)	6.70	13.50	9.95 (2.69)	5.45	21	9.86 (2.35)	5.86	15.43
Subordination	1.62 (0.38)	1.00	2.89	1.61 (0.29)	1.00	2.13	1.71 (0.37)	1.06	2.50	1.70 (0.38)	1.17	2.60
Phrasal Complexity	5.91 (0.86)	4.35	8.25	5.76 (1.10)	3.88	8.77	5.82 (0.91)	4.38	8.40	5.79 (0.65)	4.82	7.33

4.4.1.1 Overall complexity

In terms of overall complexity, UOP group had slightly higher scores in Here-and-Now task than There-and-Then task. Similarly, POP group had higher overall complexity in Here-and-Now task than There-and-Then task. When the two groups were compared, POP group had greater overall complexity than UOP in Here-and-Now task. In addition, the overall complexity of POP group was higher than UOP group in There-and-Then task. The 95% confidence intervals in Figure 6 showed overlapping patterns for two tasks by both groups. In addition, errors bars of POP group exceeded UOP group only slightly in both tasks.

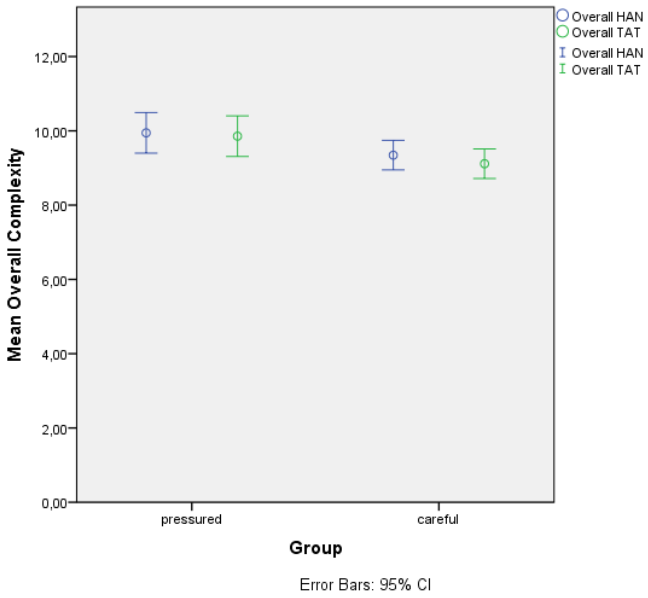


Figure 6. Overall complexity scores of two groups in two tasks with 95% confidence intervals.

A two-factor mixed ANOVA was conducted to investigate the statistical significance of the mean differences. It did not indicate any significant main effect for overall complexity, $F(1,62) = .169$, $p = .68$. Therefore, there was no statistically significant difference between simple and complex tasks in terms of overall complexity. Moreover, there was no significant main effect for group, $F(1,62) = 1.831$, $p = .18$ and it had a small effect size, $r = .16$. In other words, the groups did not differ significantly from each other in terms of overall complexity. In addition, there were no significant interaction effects between overall complexity and group, $F(1,62) = .065$, $p = .80$. The manipulation of online planning time did not have an impact on overall complexity levels of the participants.

4.4.1.2 Complexity by subordination

UOP group had similar mean scores in terms of subordination in both simple and complex tasks. POP group also had nearly equivalent subordination scores in both simple and complex tasks. According to the 95% confidence interval, POP had

slightly higher subordination scores than UOP group in both tasks (see Figure 7). However, error bars for two tasks were overlapping in both groups.

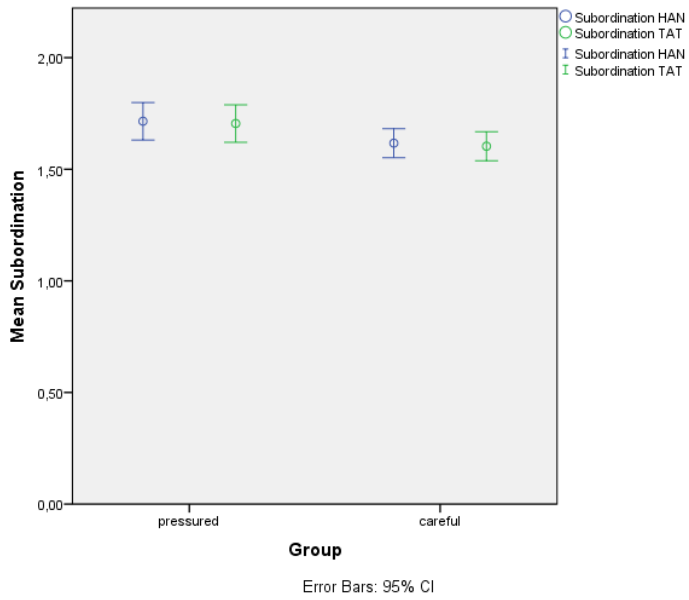


Figure 7. Subordination scores of two groups in two tasks with 95% confidence intervals.

A two-factor mixed ANOVA was conducted for subordination scores. There was no main subordination effect in participants' performances, $F(1,62) = .013$, $p = .91$. Participants had similar subordination scores in both Here-and-Now and There-and-Then tasks. Moreover, there was not a significant main effect of group, $F(1,62) = 1.663$, $p = .20$ and the effect size of group was small, the Pearson's r was .16. This means that the absence of online planning did not result in lower subordination scores as expected. Finally, there was no interaction effects between the subordination scores and the group, $F(1,62) = .010$, $p = .92$.

4.4.1.3 Phrasal complexity

Phrasal complexity was operationalized as the ratio of the total number of words to the total number of clauses. When the mean scores were examined, both UOP and POP groups indicated higher phrasal complexity levels in Here-and-Now task than

There-and-Then task. Moreover, in Here-and-Now task, UOP group had slightly higher scores than POP group whereas in There-and-Then task, the mean was higher in POP group. However, 95% confidence intervals, indicated in Figure 8, showed overlapping error bars across groups and tasks.

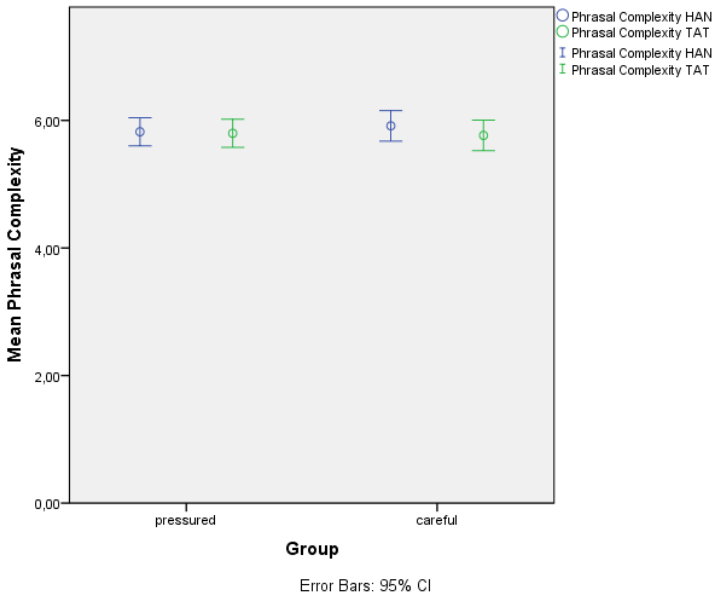


Figure 8. Phrasal complexity scores of two groups in two tasks with 95% confidence intervals.

A two-way mixed ANOVA was conducted to investigate the main effects for phrasal complexity level and group as well as their interaction effects (see Table 14). Neither phrasal complexity ($F(1, 62) = .341$, $p = .56$) nor group ($F(1,62) = .000$, $p = .99$) showed a significant main effect on learners' performances. In other words, participants in both UOP and POP groups had similar phrasal complexity levels in both simple and complex tasks. Moreover, there was no significant interaction between two variables, $F(1,62) = .398$, $p > .05$. The effect sizes of the main effects of phrasal complexity and group and the interaction effects were very small, Pearson's *r*s were .07, .00, and .08, respectively.

In summary, manipulating tasks along +/- Here-and-Now conditions did not show any statistically significant impact on syntactic complexity levels of participants' performances and the practical significance levels were also very small.

In line with previous +/- Here-and-Now studies, and unlike the predictions of the Cognition Hypothesis, There-and-Then task did not present greater syntactic complexity scores. Therefore, Hypothesis 1 was not confirmed. Differently from previous studies investigating the effects of online planning, the participants in UOP group in the current study did not have greater syntactic complexity levels than the ones in POP group and online planning had only small effect sizes for three syntactic complexity measures. Therefore, Hypothesis 3 was not confirmed. Finally, there was not any interaction effects between syntactic complexity and online planning so Hypothesis 5 was not confirmed for syntactic complexity.

4.4.2 Accuracy

Accuracy was measured as the percentage of errors in 100 words, the ratio of error-free AS units, and the ratio of error-free clauses. Table 15 shows the means, standard deviations, minimum values, and maximum values of accuracy measures of two groups in two tasks.

Table 12. Descriptive Statistics for Accuracy Measures

Accuracy	Unpressured						Pressured					
	Here-and-Now			There-and-Then			Here-and-Now			There-and-Then		
	Mean (SD)	Min.	Max.	Mean (SD)	Min.	Max.	Mean (SD)	Min.	Max.	Mean (SD)	Min.	Max.
Errors per 100 words	7.79 (3.05)	2.00	14.81	5.52 (2.03)	2.02	10.00	8.23 (3.14)	3.68	15.08	5.94 (2.40)	2.53	11.48
Error-free AS units	0.47 (0.18)	0.13	0.79	0.58 (0.16)	0.17	0.81	0.47 (0.18)	0.00	0.71	0.56 (0.18)	0.00	0.82
Error-free Clauses	0.61 (0.13)	0.29	0.83	0.71 (0.12)	0.44	0.90	0.60 (0.16)	0.13	0.82	0.69 (0.14)	0.40	0.87

4.4.2.1 Errors per 100 words

The table 12 shows that UOP group had more errors in the simple task than the complex task. Similarly, POP group had more errors in the simple task than the

complex task. In addition, POP group had higher number of errors than UOP in both simple and complex tasks. When 95% confidence interval was examined, there was a great difference between the error bars of simple and complex tasks in both groups (see Figure 9). However, the error bars of two groups were quite close to each other.

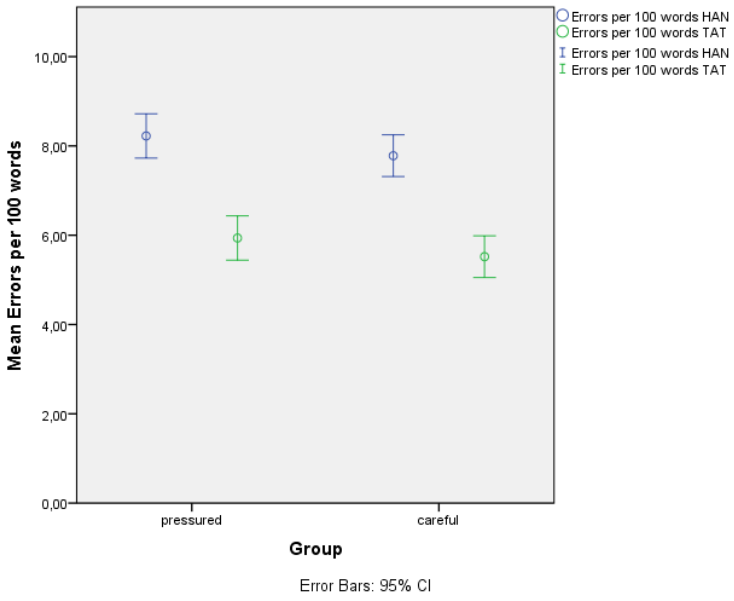


Figure 9. Errors per 100 words with 95% confidence intervals.

A two-way mixed ANOVA was conducted for investigating the main effects of errors per 100 words and group as well as interaction effects of these variables. It indicated a significant main effect of the percentage of errors in 100 words, $F(1,62) = 41.437$, $p = .000$. The Pearson's r was found to be $r = .63$, so task type had a large effect size. In other words, completing the task in There-and-Then condition had a large effect on decreasing the numbers of errors that learners made in their performances. However, there was not any significant main effect of the group, $F(1,62) = .628$, $p = .43$ and Pearson's r was found to be $r = .10$. This means that availability of online planning only had a small effect size on the numbers of errors that learners had in their performances. Besides, the two variables, namely the errors per 100 words and the group, did not interact significantly, $F(1,62) = .011$, $p = .92$.

4.4.2.2 The ratio of error-free AS units

The second accuracy measure, the ratio of error-free AS units, showed that the participants in UOP group were more accurate in the complex task than in the simple task. Likewise, in POP group the complex task led to higher error-free AS units than in the simple task. According to the 95% confidence intervals, shown in Figure 10, the mean score of There-and-Then task exceed the upper 95% confidence interval of Here-and-Now task in both groups. This shows that there is likely to be a significant difference between Here-and-Now and There-and-Then tasks in terms of the ratio of error-free AS units. However, the confidence intervals of two groups for each task are overlapping which shows that the two groups behaved similarly in both Here-and-Now and There-and-Then tasks.

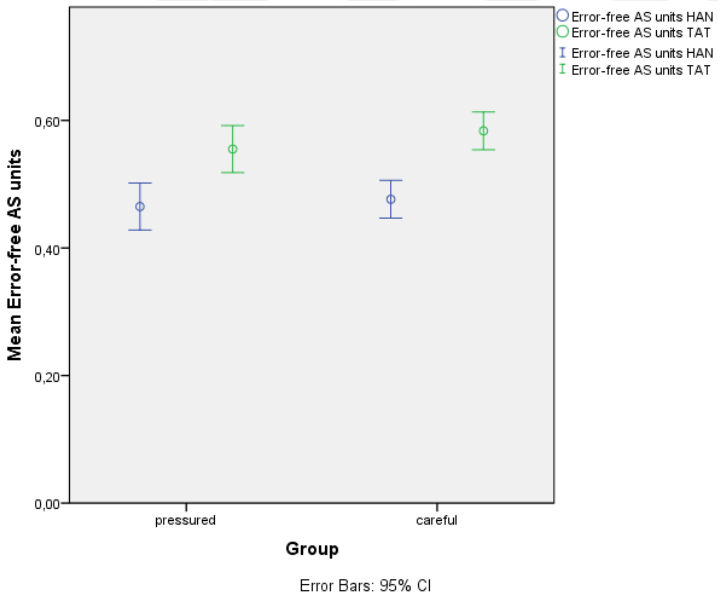


Figure 10. The ratio of error-free AS units with 95% confidence intervals.

A two-way mixed design ANOVA was carried out for investigating the statistical significance of the differences between mean scores. It presented a main effect of error-free AS units, $F(1,62) = 9.305$, $p = .003$. Moreover, Pearson's r was .36 which can be regarded as an effect size very close medium level. Therefore, There-and-Then task can be stated to lead to significantly greater accuracy levels than Here-and-

Now task and to have a medium effect size in learners' production of error-free AS units. However, there was no main effect of group, $F(1, 62) = .191$, $p = .66$ and it had a very small effect size, $r = .06$. In other words, the two groups had similar accuracy scores in both tasks. Additionally, there was not a significant interaction effect of the error-free AS units and the group, $F(1, 62) = .115$, $p = 74$.

4.4.2.3 The ratio of error-free clauses

The third measure, the ratio of error-free clauses indicated UOP group to be more accurate in There-and-Then task than in Here-and-Now task. POP group showed a similar pattern. Participants in POP also had higher accuracy in There-and-Then task than Here-and-Now task. For both groups, the mean scores of There-and-Then tasks surpass the 95% confidence interval of Here-and-Now tasks (See Figure 11). Moreover, when the two groups were compared, UOP group were more accurate than POP group in both tasks. However, according to 95% confidence intervals, the scores of both groups were overlapping in both tasks.

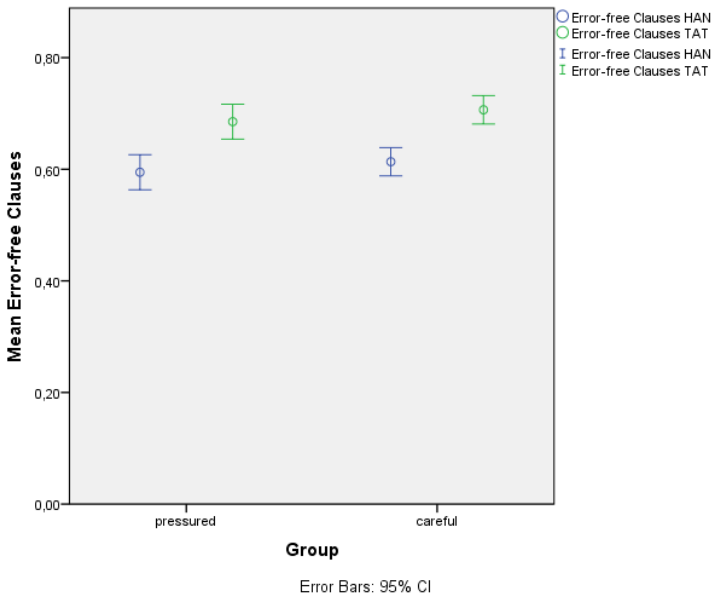


Figure 11. The ratio of error-free clauses with 95% confidence intervals

A two-way mixed design ANOVA was conducted to investigate whether the differences that were observed through descriptive statistics and confidence intervals were statistically significant. Firstly, there was a statistically significant difference between two tasks in terms of the ratio of error-free clauses, $F(1,62) = 9.496$, $p = .003$, $r = .36$. Participants showed significantly more accurate performances in There-and-Then task than in Here-and-Now task with a close to medium effect size. Secondly, there were no significant main effect of group, $F(1,62) = 2.948$, $p = .09$. In other words, availability or absence of online planning did not affect participants' performances significantly. However, this measure had a small effect size, r = .20. This means that although the difference between groups is not statistically significant, the availability of online planning still has an effect on the ratio of error-free clauses of participants' performances. Finally, there was no interaction effect of error-free clauses and group, $F(1,62) = .017$, $p = .90$.

In general, participants had more errors in Here-and-Now task regardless of their group. Specifically, manipulating tasks along +/- Here-and-Now condition had a large effect size on the numbers of errors that participants had in their performances. In addition, they had more accurate performances in There-and-Then tasks in terms of the error-free AS units and error-free clauses. Therefore, Hypothesis 2 was confirmed. Participants were more accurate in There-and-Then task than in Here-and-Now task. However, Hypothesis 4 regarding the effects of online planning was not confirmed. There was no difference between UOP and POP groups in terms of their accuracy levels. Group variable had a small effect size only in terms of the ratio of error-free clauses whereas for errors per 100 words and the ratio of error-free AS-units, the effect sizes were very small. Finally, there was no significant

interaction effects between groups. Hypothesis 5 was not confirmed for accuracy, either.

CHAPTER 5

DISCUSSION

5.1 Introduction

This chapter discusses the oral performance results in the light of the Limited Attentional Capacity Hypothesis, the Cognition Hypothesis, and First and Second Language Speech Production models. It focuses on the effects of attentional resources, the operationalization of tasks, and the proficiency levels of participants. The results showed an increase in accuracy scores in There-and-Then tasks. However, these tasks did not result in increased syntactic complexity levels. Moreover, there were no significant differences between pressured and unpressured online planning groups in terms of accuracy and syntactic complexity levels.

5.2 The relationship between task type and L2 oral performance

The participants' oral performances were assessed through syntactic complexity and accuracy measures. As in line with the Cognition Hypothesis (Robinson, 2001a, 2011a) and previous studies (Gilabert, 2007a, 2007b; Ishikawa, 2007; Iwashita, et al., 2001; Rahimpour, 1997; Robinson, 1995), There-and-Then task resulted in more accurate performances than Here-and-Now task. For explaining this difference in There-and-Then task, Robinson (1995) argues that the absence of a context in There-and-Then tasks leads learners to use a more accurate language since in these tasks learners do not have a context-support which can compensate for the gaps in their language. In other words, learners pay more attention to the use of correct language form to be understood correctly by their interlocutors. Furthermore, although There-and-Then tasks are categorized as more complex tasks in Robinson's Triadic

Componential Framework (2007a, 2011a), these tasks are "less input-dominated and more negotiable" since participants do not have pictures in front of them (Skehan, 2009, p. 526). This means that while formulating their speech, they are more independent and they can choose the vocabulary or syntactic structures more freely. This ease on the formulator of learners' speech production mechanism leads to increased accuracy (Skehan, 2009; Wang, 2014). The results of the current study support these arguments since There-and-Then tasks led to higher ratios of error-free AS units and clauses. Moreover, learners had less errors in these tasks. This improvement in learners' accuracy levels show that learners can direct their attention to form-meaning mappings while performing a more complex task. This finding confirms the Cognition Hypothesis (Robinson, 2001a, 2011a) which predicts higher accuracy levels in more complex tasks. However, it is not compatible with the Limited Attentional Capacity Hypothesis (Skehan, 1998) which presumes that more complex tasks require learners to pay more attention to meaning rather than form and that these complex tasks decrease the accuracy and syntactic complexity levels of learners.

However, similar to the results of the previous studies which investigated the effects of +/- Here-and-Now in L2 oral performance (Gilabert, 2007a; Iwashita, et al., 2001; Rahimpour, 1997; Robinson, 1995), none of the syntactic complexity measures showed a significant difference between two tasks. Unlike the Cognition Hypothesis (Robinson, 2011a), which predicts higher syntactic complexity in more complex tasks, this finding indicates that There-and-Then tasks do not necessarily result in syntactically complex language. There can be two reasons for the lack of significant differences between two tasks in the current study.

Firstly, in narrative There-and-Then tasks learners need to memorize the characters and the storyline in a story. Therefore, learners try to keep the events on their minds and tell their stories simultaneously. This situation may create a burden on their attentional resources so that they may not be able to attend to the syntactic complexity of their speech. In addition, syntactic complexity of L2 performance is affected by the pressure on the conceptualization stage and There-and-Then tasks increase the conceptual operations since they require learners to do two operations simultaneously (i.e. keeping the storyline in the memory while narrating the story) (Skehan, 2009). Despite this pressure, two studies conducted by Gilabert (2007a) and Ishikawa (2007) reported increased syntactic complexity in There-and-Then tasks. However, in Gilabert's study this improvement in complexity levels was achieved when participants were provided with pre-task planning time. Similarly, in Ishikawa's study participants had five-minute pre-task planning time and they completed the tasks in written mode which allows for planning. In these studies, the availability of pre-task planning time eased the burden on learners' attentional resources as well as on the conceptualizer in their bilingual speech production mechanisms. Therefore, unlike Robinson's predictions, these two studies showed that There-and-Then condition alone was not sufficient to direct learners' attention to syntactic complexity of their speech. The pressure caused by a necessity of memorization should be mitigated with providing pre-task planning time. Since the participants in the current study completed the tasks in oral mode and they were given only 60-second time which was sufficient for understanding the pictures but not for planning their speeches, they may not have had an opportunity to direct their attention to the syntactic complexity of their performances.

Secondly, based on Wickens's (1989, as cited in Kormos, 2000) arguments, Kormos (2000) argues that "upon processing their speech, L2 learners need to rely on the same verbal resource pool, therefore the various phases of speech production need to compete with each other for attentional resources". In other words, there may be trade-off effects between accuracy and complexity due to a competition for attention between them as also suggested by Skehan (1998). Because of the absence of a context to support their speeches, in There-and-Then tasks learners may prioritize accuracy over syntactic complexity for the sake of being understood correctly by their interlocutors. Thus, instead of using new and more complex structures, they may have adopted a "safety first approach" and they used more stable and automated L2 structures in their speeches (Skehan & Foster, 2001). The findings related to the syntactic complexity are more compatible with the Limited Attentional Capacity Hypothesis than the Cognition Hypothesis as the former one predicts a competition between accuracy and complexity in learners' performances. The latter one, on the other hand, expects a concurrent improvement in learners' accuracy and syntactic complexity levels in more complex tasks.

Overall, the results of this study partially confirmed the predictions of the Cognition Hypothesis (Robinson, 2001a, 2005, 2007a, 2011a) and the Limited Attentional Capacity Hypothesis (Skehan, 1998, 2009). Increased task complexity along +/- Here-and-Now resulted in more accurate performance as predicted by the Cognition Hypothesis. However, there seemed to be trade-off effects between accuracy and syntactic complexity since there were no differences between participants' syntactic complexity levels in two tasks as presumed by the Limited Attentional Capacity Hypothesis.

5.3 The relationship between online planning and L2 oral performance

Following the previous online planning studies (Ahmadian & Tavakoli, 2010; Ahmadian, Tavakoli, & Dastjerdi, 2015; Ellis & Yuan, 2005; Yuan & Ellis, 2003), availability of online planning was expected to have a positive effect on participants syntactic complexity and accuracy levels. Thus, unpressured online planning group was predicted to have higher syntactic complexity levels than pressured online planning group. However, this hypothesis was not confirmed by any of the three syntactic complexity measures. Similarly, the absence of online planning in this study did not result in significantly lower accuracy scores and higher numbers of errors as hypothesized. Instead, both groups had similar accuracy scores across two tasks. There can be several reasons for this discrepancy between the current study and the previous studies.

First reason for this could be related to the proficiency levels of participants. The participants in this study were studying at English Language Education Department and they had high scores in the institution based English Proficiency Test. In addition, during their first year in the program, they had taken courses to develop their communicative skills in English. Therefore, the participants can be argued to have high proficiency in English. Kormos (2006) points out that as learners' L2 proficiency increases, the morphological, syntactic, and phonological rules that had been learned in the declarative form become proceduralized. Therefore, the retrieval of knowledge that once required controlled processing becomes more automatic. This assumed automatization in the formulator of the speech production system can be the reason for insignificant differences between the groups in the current study. As a result of their automatization in the formulation stage, time pressure might not have affected the performances of participants in

pressured online planning group. In addition, Wang (2014) argues that time pressure affects the retrieval of "newly learned but not yet automatized linguistic structures" in L2 (p.32). The pictures used in the current study required the use of rather frequent, daily vocabulary. In addition, participants used either present or past tenses in their stories. This use of simple vocabulary and grammatical structures can also be another reason for the lack of difference between groups. In other words, both tasks employed in the study may be too easy for the participants.

Secondly, the operationalization of online planning was different from previous studies (Ahmadian & Tavakoli, 2010; Ahmadian, Tavakoli, & Dastjerdi, 2015; Ellis & Yuan, 2005; Yuan & Ellis, 2003). In previous studies, the researchers operationalized online planning through time pressure (e.g. three minutes for completing the task) and the amount of sentences that learners needed to produce at the given time (e.g. four sentences for each task) to increase the pressure (Ahmadian & Tavakoli, 2010; Ahmadian, Tavakoli, & Dastjerdi, 2015; Ellis & Yuan, 2005). Thus, the participants in pressured online planning group not only had to complete their tasks in a limited time, they also had to reach the minimum amount of sentences per each picture. Since this study manipulated the tasks along +/- Here-and-Now condition, it was not possible to give participants a sentence limit for each picture and the pressure was only set through the time limit given to the participants in the pressured online planning group. Therefore, it can be said that time pressure does not necessarily lead to lower syntactic complexity and accuracy scores in learners' performances since giving a limited time for completing the task may not put high proficient learners under enough pressure.

These findings in relation to the effects of online planning do not support the Cognition Hypothesis and the Limited Attentional Capacity Hypothesis since both of

them predict decreased syntactic complexity and accuracy levels in tasks with increased resource-dispersing demands. Time pressure alone may be said to have no significant impacts on the accuracy and syntactic complexity levels of the high proficient learners especially with tasks that involve the use of vocabulary and grammatical structures that learners have already proceduralized in their memories.

5.4 The combined effects of task type and online planning

Robinson (2011a) argues for the existence of combined effects when tasks are manipulated along both resource-directing and resource-dispersing features. In the light of this argument, this study investigated the interaction effects of task type and online planning for syntactic complexity and accuracy measures.

A study conducted by Gilabert (2007a) reported positive combined effects of tasks that are complex along resource-directing dimension (i.e. There-and-Then) and simple along resource-dispersing dimension (i.e. + pre-task planning-time). However, these positive effects were only significant for accuracy but not for syntactic complexity. Similar to Gilabert's finding, this study also failed to report improved syntactic complexity levels in There-and-Then task in unpressured online planning group. This means that, even when participants completed the tasks in There-and-Then condition and had online planning opportunity, this optimal condition, as suggested by Robinson (2011a), did not result in syntactically more complex language. However, differently from Gilabert (2007a), there was no significant improvements in accuracy levels, either. Although the lowest mean for the numbers of errors per 100 words and the highest means of the ratio of error-free AS units and clauses were found in There-and-Then task completed by unpressured

online planning group, these scores were not significantly different from other conditions.

Moreover, Gilabert (2007a) predicted negative effects of task complexity when the complexity was increased along both resource-directing and resource-dispersing dimensions. In the light of this prediction, the lowest mean scores for accuracy were expected for There-and-Then task completed by POP group. However, the lowest mean scores for accuracy in all three measures belonged to POP group in Here-and-Now task.

These discrepancies between Gilabert's (2007a) study and the current study can be due to the use of distinct accuracy measures. While Gilabert investigated accuracy with the percentage of self-repairs, this study investigated the errors with errors per 100 words and the accuracy levels with the ratio of error-free AS-units and clauses. Secondly, the participants in Gilabert's study had lower-intermediate proficiency in English while the participants in the current study had higher levels of proficiency. Therefore, the combined effects of resource-directing and resource-dispersing dimensions may depend on the proficiency levels of the learners.

Overall, these results show that online planning time and task type do not have any positive or negative combined effects on the accuracy and syntactic complexity levels of highly proficient L2 learners.

CHAPTER 6

CONCLUSION

6.1 Summary and conclusions

The main aim of this study was to examine the effects of task type and online planning on learners' accuracy and syntactic complexity levels. Although the effects of both features have been studied previously (Ahmadian & Tavakoli, 2010; Ahmadian, Tavakoli, & Dastjerdi, 2015; Ellis & Yuan, 2005; Gilabert, 2007a, 2007b; Gilabert, Baron & Llanes, 2009; Ishikawa, 2007; Iwashita et al., 2001; Robinson, 1995; Wang & Skehan, 2014; Wang, 2014; Yuan & Ellis, 2003), very few, if any, studies investigated the combined effects of these features. This study is one of the first studies which attempted to investigate the separate as well as combined effects of manipulating tasks along task type and online planning variables.

Moreover, differently from previous studies, most of which used subordination measures for measuring syntactic complexity, following the recommendations of Norris and Ortega (2009), this study employed three different syntactic complexity measures which are argued to tap into different aspects of performance, namely overall complexity, complexity by subordination, and phrasal complexity. Accuracy was also measured with more than one measure. In addition to measures focusing on the correct use of language (i.e. the ratios of error-free AS units and clauses), this study also employed a measure investigating the errors in participants' performances (i.e. errors per 100 words). This measure is argued to be more precise than measures examining the correct language use since it determines the number of errors in each narrative (Gilabert, Baron, & Levkina, 2011).

In addition, most of the previous studies on +/- Here-and-Now and +/- Online Planning investigated the effects of task complexity manipulation on lower or upper intermediate level English learners. Differently from previous studies, one of the main purposes of this study was to observe whether the promoting effects of +/- Here-and-Now and +/- Online Planning reported in previous studies can be true for highly proficient English learners. Therefore, the current study was conducted with relatively proficient participants.

The results of the study show that highly proficient English learners evaluated There-and-Then tasks as a little more complex than Here-and-Now tasks regardless of the availability of online planning however this difference was not as much as the ones reported by the previous studies. Even if the perceived difficulty mean score was higher for There-and-Then task than Here-and-Now task, the mean score was still regarded to be relatively easy. In other words, none of the participants evaluated There-and-Then tasks as very difficult.

According to the production results, There-and-Then tasks resulted in more accurate speech than Here-and-Now tasks but the two tasks did not significantly differ from each other in terms of their impacts on L2 learners' syntactic complexity levels. Pressured online planning, on the other hand, did not have any significant effects on accuracy and syntactic complexity scores of high proficient English learners. Finally, task type and online planning did not have any combined effects in L2 performance of learners with high proficiency unlike the predictions of Robinson (2011a) and Gilabert (2007a).

Overall, this study has some pedagogical implications for the effects of different task types in language classrooms. Regardless of task complexity, picture narration tasks are easy tasks for highly proficient English learners. Manipulating

these tasks along +/- Here-and-Now condition does not direct attention to syntactic complexity levels, yet they can be good tasks for attending to accuracy. Therefore, these tasks can be mostly used for practicing the existing knowledge of the participants. Besides, they can be good tasks for increasing the automaticity in the language. Moreover, although changing online planning availability has some effects on lower intermediate or intermediate level learners, it does not influence the oral performance of highly proficient L2 learners greatly. In other words, online planning availability does not affect allocation of attentional resources of proficient L2 learners. All of these findings should be considered when designing tasks for highly proficient learners.

Finally, this study has some theoretical implications. It shows that all task complexity variables presented in the Triadic Componential Framework (Robinson, 2007, 2011a) may not have the predicted effects on all learners. Proficiency could be one of the determinant factors for the impacts of various tasks as well as task complexity. Therefore, while categorizing the L2 pedagogic tasks as simple or complex tasks, proficiency should be taken into consideration.

6.2 Limitations and suggestions for future research

This study has several limitations. Firstly, only general measures of accuracy and syntactic complexity were employed in this study because these measures can capture the variation in a larger-scale (Norris & Ortega, 2009) and they allow for a comparison between the results of the current study and the ones reported by the previous studies. However, specific measures of L2 performance which are more task-relevant (e.g. structural variety for complexity and target-like use of articles for

accuracy) can be more sensitive in capturing the differences between two tasks (Norris & Ortega, 2009; Robinson et al., 2009).

Secondly, although accuracy levels were assessed with three measures, the types of errors that learners had were not examined. However, especially for highly proficient L2 speakers, investigating the error types could provide a more detailed analysis for the effects of task complexity on accuracy. In addition, this study did not examine the fluency levels of participants. Since participants were highly proficient in English, fluency scores may provide a better evidence for the automatization levels of these participants.

Thirdly, as indicated by the results of perceived difficulty scores, the tasks employed in the study may be too simple for the participants. Since they were all highly proficient learners, even There-and-Then task which was assumed to be a complex task may not have challenged them enough. Therefore, more challenging tasks can indicate different results. Besides, although picture-based narrative tasks are widely employed in task complexity studies, their transferability to language classrooms is questionable. Although picture narration in Here-and-Now condition is utilized in TBLT classrooms, picture narration in There-and-Then condition is not a common pedagogical task and it may not reflect a real-world task. Instead of pictures, employing silent videos could increase the authenticity of the tasks used in the study.

Additionally, the two picture sets, the Picnic and the Chase, were assumed to be comparable based on some criteria explained in the Methodology section. Therefore, they were expected to result in similar oral performances when the complexity of the tasks was the same. However, further research investigating

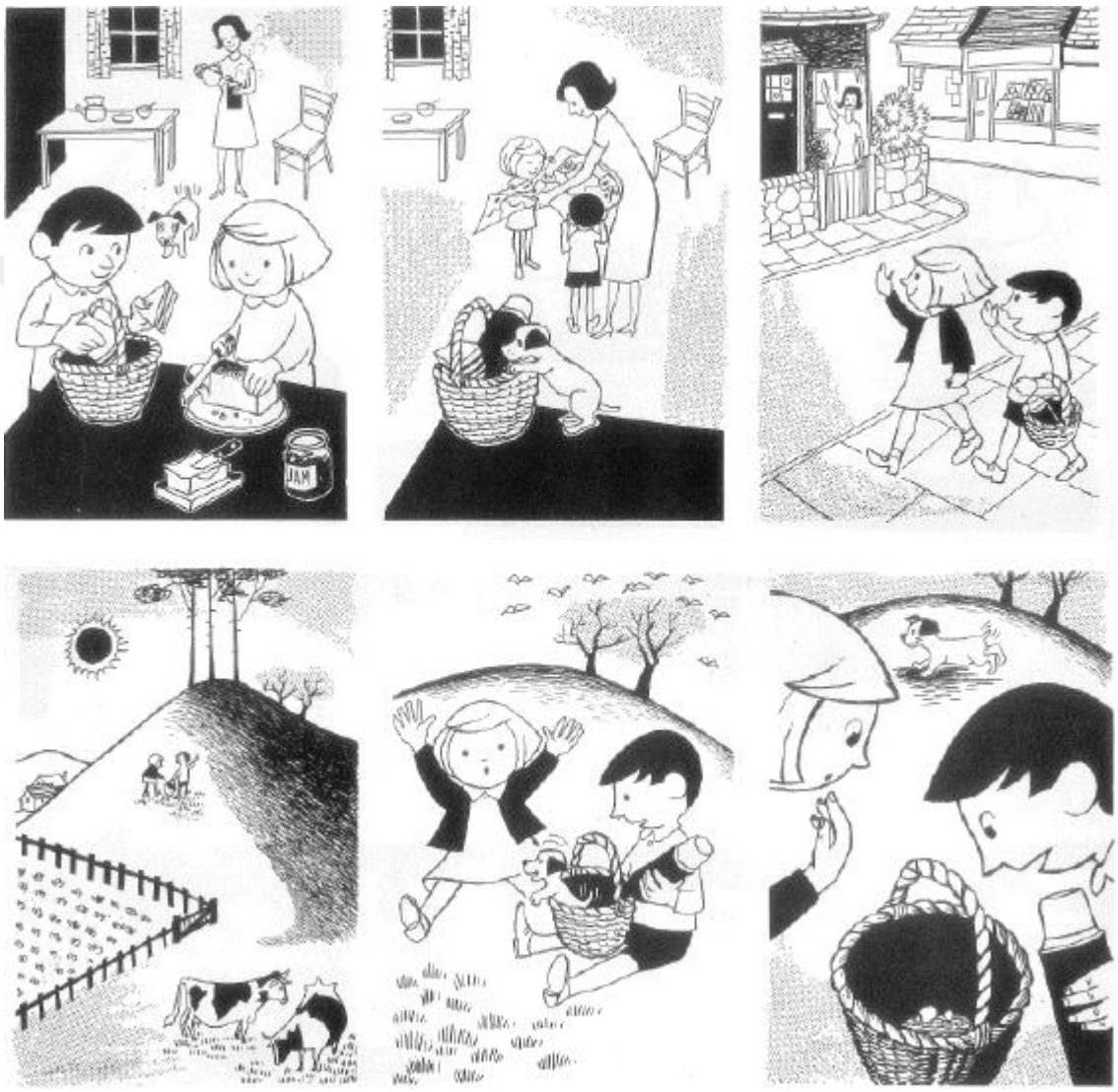
learners' perception about the comparability of the tasks with a retrospective questionnaire can provide more evidence for the comparability of these picture sets.

Finally, during the data collection and data analysis I observed some variation between learners in terms of the use of the elements that were critical for a thorough understanding of the story described in the pictures. In other words, some important elements in the stories (such as the dropped package in first frame of the Chase story or the small dog getting into the basket in second frame of the Picnic story) were not reported at all by a few participants, whereas others incorporated those elements in their stories. Therefore, it would be interesting to examine to what extent the elements in six picture frames are included in participants' stories. However, when the content of participants' stories are to be investigated, as pointed out by my thesis co-advisor, visual saliency of the pictures could be a key factor in the stories that learners produce. A further eye tracking research examining the eye movements of the participants when engaged with the pictures can provide a further understanding of how the visual saliency can affect oral performance in picture-based narrative tasks.

APPENDIX A

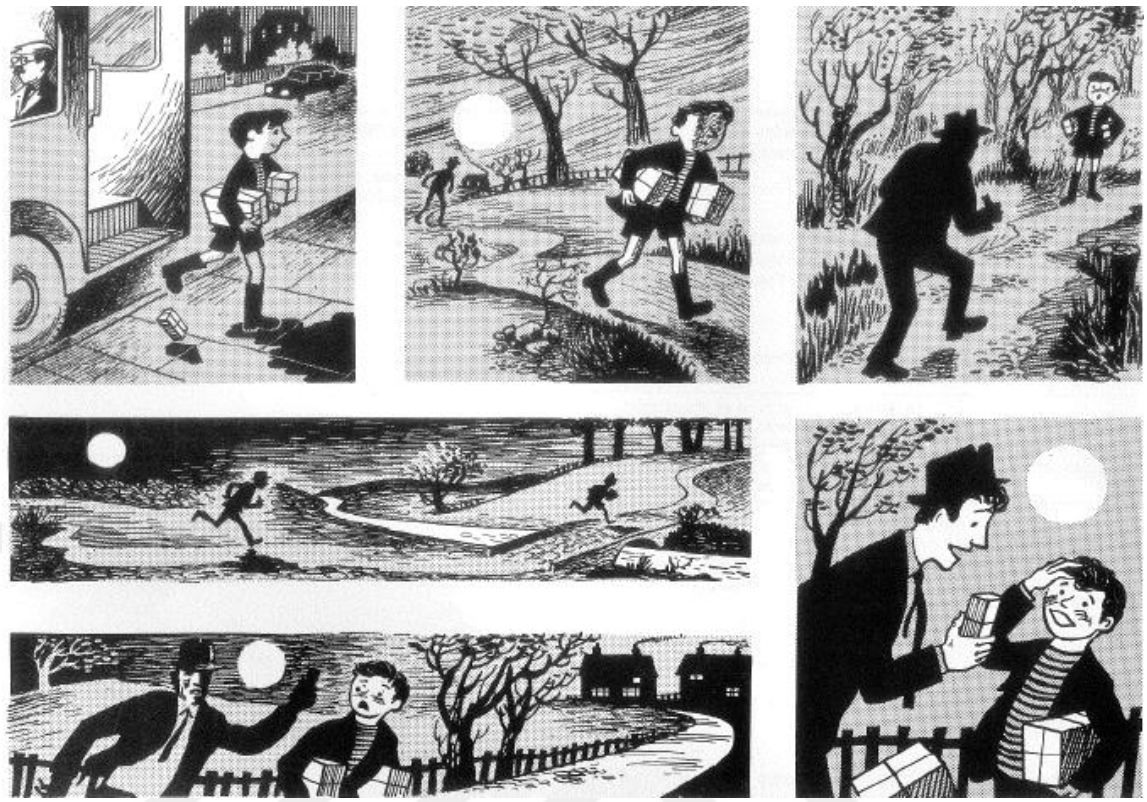
PICTURES USED IN TASKS

Picture 1: The Picnic



Source: [Heaton, 1966]

Picture 2: The Chase



Source: [Heaton, 1975]

APPENDIX B

BACKGROUND QUESTIONNAIRE

1. Name/ Surname: _____
2. Age: _____
3. Nationality: _____
4. Mother Tongue: _____

5. Foreign Language(s) and Proficiency: Oral Written
   1. _____ High/ Medium/ Low High/ Medium/ Low
   2. _____ High/ Medium/ Low High/ Medium/ Low
   3. _____ High/ Medium/ Low High/ Medium/ Low

6. BUEPT (Boğaziçi University English Proficiency Exam) score: Overall: _____
   Listening: _____
   Reading: _____
   Writing: _____

7. How long have you been studying English?

_____

8. Have you ever lived abroad? Where? How long?

_____

9. Do you use English regularly? In which contexts (classrooms, social network, etc.)?

_____

10. Do you like speaking English?

_____

11. Do you feel comfortable speaking English?

_____

12. How would you rate your own ability to speak English? (very poor, poor, fair, good, excellent)

_____

APPENDIX C

LIKERT SCALE FOR PERCEIVED TASK DIFFICULTY

Affective Questionnaire

Name: _____ Task: _____

Please evaluate the speaking task that you have just narrated choosing the appropriate answer.

1. How was the task?

1 Very Easy	2 Easy	3 Moderately Easy	4 Moderately Difficult	5 Difficult	6 Very Difficult

APPENDIX D

TASK INSTRUCTIONS IN ENGLISH

1 Task instructions for Unpressured Online Planning Group

1.1 Here-and-Now task

In this study we will show you pictures ordered from 1 to 6. These pictures narrate a story. You need to tell the story shown on the pictures in English.

- Before starting to tell your story, you have 1 minute to examine and understand the pictures.
- After 1 minute, you should tell the story.
- You will be able to see the pictures while telling your story.
- There is no time limit to tell the story.
- You should tell the story as detailed as possible as if you were telling it to a friend who has never seen the pictures.
- You should use present tense while telling the story.

If you have any questions, you should ask them before starting. You should not ask any questions after seeing the pictures, and you should tell the story without any breaks.

1.2 There-and-Then

In this study we will show you pictures ordered from 1 to 6. These pictures narrate a story. You need to tell the story shown on the pictures in English.

- Before starting to tell your story, you have 1 minute to examine and understand the pictures.
- After 1 minute, you should tell the story.
- You will not be able to see the pictures while telling your story.
- There is no time limit to tell the story.
- You should tell the story as detailed as possible as if you were telling it to a friend who has never seen the pictures.
- You should use past tense while telling the story.

If you have any questions, you should ask them before starting. You should not ask any questions after seeing the pictures, and you should tell the story without any breaks.

2 Task instructions for Pressured Online Planning Group

2.1 Here-and-Now task

In this study we will show you pictures ordered from 1 to 6. These pictures narrate a story. You need to tell the story shown on the pictures in English.

- Before starting to tell your story, you have 1 minute to examine and understand the pictures.
- After 1 minute, you should tell the story.
- You will be able to see the pictures while telling your story.
- You have only 90 seconds to tell the story.
- You can check the remaining time on the timer.
- You should tell the story as detailed as possible as if you were telling it to a friend who has never seen the pictures.
- You should use present tense while telling the story.

If you have any questions, you should ask them before starting. You should not ask any questions after seeing the pictures and you should tell the story without any breaks.

2.2 There-and-Then

In this study we will show you pictures ordered from 1 to 6. These pictures narrate a story. You need to tell the story shown on the pictures in English.

- Before starting to tell your story, you have 1 minute to examine and understand the pictures.
- After 1 minute, you should tell the story.
- You will not be able to see the pictures while telling your story.
- You have only 90 seconds to tell the story.
- You should tell the story as detailed as possible as if you were telling to a friend who has never seen the pictures
- You should use past tense while telling the story.

If you have any questions, you should ask them before starting. You should not ask any questions after seeing the pictures and you should tell the story without any breaks.

APPENDIX E

TASK INSTRUCTIONS IN TURKISH

1 Task instructions for Unpressured Online Planning Group

1.1 Here-and-Now task

Bu çalışmada size 1'den 6'ya kadar sıralanmış resimler göstereceğiz. Bu resimler bir hikaye anlatmaktadır. Bu resimlerde betimlenen hikayeyi İngilizce olarak anlatmalısınız.

- Hikayeyi anlatmaya başlamadan önce resimleri incelemeniz ve anlamanız için 1 dakika süreniz var.
- 1 dakika sonunda sizden hikayeyi anlatmanızı isteyeceğiz.
- Hikayeyi anlatırken resimlere bakmaya devam edeceksiniz.
- Hikayeyi anlatırken herhangi bir süre kısıtlaması bulunmamaktadır.
- Hikayeyi bu resimleri hiç görmemiş olan bir arkadaşınıza anlatırmış gibi mümkün olduğu kadar ayrıntılı bir şekilde anlatın.
- Hikayeyi anlatırken geniş zaman/ şimdiki zaman kullanın.

Eğer bir sorunuz varsa çalışmaya başlamadan önce sormalısınız. Resimleri gördüğünüz andan itibaren herhangi bir soru sormamalı, istendiğinde hikayeyi ara vermeden anlatmalısınız.

1.2 There-and-Then task

Bu çalışmada size 1'den 6'ya kadar sıralanmış resimler göstereceğiz. Bu resimler bir hikaye anlatmaktadır. Bu resimlerde betimlenen hikayeyi İngilizce olarak anlatmalısınız.

- Hikayeyi anlatmaya başlamadan önce resimleri incelemeniz ve anlamanız için 1 dakika süreniz var.
- 1 dakika sonunda sizden hikayeyi anlatmanızı isteyeceğiz.
- Hikayeyi resimlere bakmadan anlatacaksınız.
- Hikayeyi anlatırken herhangi bir süre kısıtlaması bulunmamaktadır.
- Hikayeyi bu resimleri hiç görmemiş olan bir arkadaşınıza anlatırmış gibi mümkün olduğu kadar ayrıntılı bir şekilde anlatın.
- Hikayeyi anlatırken geçmiş zaman kullanın.

Eğer bir sorunuz varsa çalışmaya başlamadan önce sormalısınız. Resimleri gördüğünüz andan itibaren herhangi bir soru sormamalı, istendiğinde hikayeyi ara vermeden anlatmalısınız.

2 Task instructions for Pressured Online Planning Group

2.1 Here-and-Now task

Bu çalışmada size 1'den 6'ya kadar sıralanmış resimler göstereceğiz. Bu resimler bir hikaye anlatmaktadır. Bu resimlerde betimlenen hikayeyi İngilizce olarak anlatmalısınız.

- Hikayeyi anlatmaya başlamadan önce resimleri incelemeniz ve anlamanız için 1 dakika süreniz var.
- 1 dakika sonunda sizden hikayeyi anlatmanızı isteyeceğiz.
- Hikayeyi anlatırken resimlere bakmaya devam edeceksiniz.
- Hikayeyi bu resimleri hiç görmemiş olan bir arkadaşınıza anlatırmış gibi mümkün olduğu kadar ayrıntılı bir şekilde anlatın.
- Hikayeyi anlatmak için 90 saniye süreniz var.
- Hikayeyi anlatırken kalan vakti kronometreden görebilirsiniz.
- Hikayeyi anlatırken geniş zaman/ şimdiki zaman kullanın.

Eğer bir sorunuz varsa çalışmaya başlamadan önce sormalısınız. Resimleri gördüğünüz andan itibaren herhangi bir soru sormamalı, istendiğinde hikayeyi ara vermeden anlatmalısınız.

2.2 There-and-Then task

Bu çalışmada size 1'den 6'ya kadar sıralanmış resimler göstereceğiz. Bu resimler bir hikaye anlatmaktadır. Bu resimlerde betimlenen hikayeyi İngilizce olarak anlatmalısınız.

- Hikayeyi anlatmaya başlamadan önce resimleri incelemeniz ve anlamanız için 1 dakika süreniz var.
- 1 dakika sonunda sizden hikayeyi anlatmanızı isteyeceğiz.
- Hikayeyi resimlere bakmadan anlatacaksınız.
- Hikayeyi bu resimleri hiç görmemiş olan bir arkadaşınıza anlatırmış gibi mümkün olduğu kadar ayrıntılı bir şekilde anlatın.
- Hikayeyi anlatmak için 90 saniye süreniz var.
- Hikayeyi anlatırken kalan vakti kronometreden görebilirsiniz.
- Hikayeyi anlatırken geçmiş zaman kullanın.

Eğer bir sorunuz varsa çalışmaya başlamadan önce sormalısınız. Resimleri gördüğünüz andan itibaren herhangi bir soru sormamalı, istendiğinde hikayeyi ara vermeden anlatmalısınız.

APPENDIX F

EXAMPLE AS-UNIT CODING

1 Transcription Sample before AS-unit Coding

Matt and Matt and his sister uh planned to go somewhere. So they prepared something for their travelling. And their mother helped them. Also uh by the way they also had a small nice dog. And uhh after they uh prepared for the journey their mother showed the map so that they could find the place easily. And so after they uh did everything for journey they leaved the they left the home. And also finally they reached the point the location they wanted to sit and watched watch the nature. And uhh so uhh birds uh around co- there were a lot of cows and birds. And uh [long pause] then later on they discovered that the things they got for dinner for eating uhh they realized:: that they lost them. probably the dog ate the all things.

2 Transcription Sample after AS-unit Coding

|{Matt and} Matt and his sister {uh} planned:: to go somewhere|

|{Uhh and} so they prepared something for their travelling|

|And their mother helped them|

|{Also uh} by the way they also had a small nice dog|

|{And uhh} after they {uh} prepared for the journey:: their mother showed the map|

|so that they could find the place easily|

|{And so} after they {uh} did everything for journey:: {they leaved the} they left the home|

|{And also} finally they reached {the point} the location:: they wanted to sit:: and {watched} watch the nature|

{And uhh so uhh}

|{Birds uh around co-} there were a lot of cows and birds|

{And uh [long pause]}

|{Then} later on they discovered:: that the things they got {for dinner} for eating {uhh}|

|they realized:: that they lost them|

|probably the dog ate the all things|

Notes.

1. |…| shows the AS unit boundaries
2. :: shows the clause boundaries
3. {…} shows the dysfluency markers (i.e., reformulations, repetitions, false starts, and replacement) that were omitted in the pruned narrative.

3 Pruned Narrative Sample

|Matt and his sister planned:: to go somewhere|

|so they prepared something for their travelling|

|And their mother helped them|

|by the way they also had a small nice dog|

|after they GR prepared for the journey:: their mother showed GR the map:: so that they could find the place easily|

|after they did everything for the journey:: they left the home|

|finally they reached the location:: they wanted to sit:: and watch the nature|

|there were a lot of cows and birds|

|later on they discovered:: that the things:: they got for eating|

|they realized:: that they lost them|

|probably the dog ate the all things|

APPENDIX G

CODING GUIDELINES FOR SECOND CODER

Please code the data according to the instructions given below.

Note: In the examples below, disfluency markers (i.e., reformulations, repetitions, false starts, and replacement) are put inside brackets {…}.

1 AS-unit

An AS-unit is an independent clause together with any subordinate clause(s) associated with it.

Write each AS unit on a different line and put an upright slash …|… at the beginning and at the end

Example:

Segment of text before AS-unit parsing:

- {Uhh} Matt and his sister is preparing for a picnic {uhh} their mother helps them and then {uhh} Matt likes chocolate pie very much.

Segment of text after AS-unit parsing:

- |{Uhh} Matt and his sister is preparing for a picnic|
  |{Uhh} their mother helps them|
  |And then {uhh} Matt likes chocolate pie very much|

2 Clause

A clause consists minimally of a finite Verb or a non-finite Verb element plus at least one other clause element (Subject, Object, Complement, or Adverbial)

Mark each clause boundary with a double colon (::)

2.1 A finite Verb

2.1.1 Coordination: Coordinated main clauses should be treated as separate AS-units

Examples:

- Then he was followed by a man and the man was so scary and John was scared too.

|then he was followed by a man|

|And the man was so scary|

|and John was scared too| (3 AS-units)

- suddenly he drops one of his packets but he doesn't notice that

  |suddenly he drops one of his packets|

  |but he doesn't notice that| (2 AS-units)

- the dark man he wears a hat so in the dark John cannot recognize his face

  |the dark man he wears a hat|

  |So in the dark John cannot recognize his face| (2 AS-units)

When there is a coordination of verb phrases, the coordinated phrases will be considered to belong to the same AS-unit, unless the first phrase is marked by falling or rising intonation and is followed by a pause of at least 0.5 seconds (this is a clearly noticeable pause).

Example:

- |they are putting some jam and butter on their breads:: and putting them in their {buc- bucket} bucket| (1 AS-unit, 2 clauses)

2.1.2 Subordination:

Subordinated clauses belong to the same AS-unit.

Examples:

- |And then when he {got on back} got back on:: {umm uhh} he had to walk like {uh} a distance very far| (1 AS-unit, 2 clauses)
- |they just realized:: that the dog ate everything (1 AS-unit, 2 clauses)

The subordinated clauses starting with "because" normally belongs to the same AS-unit

Example:

- |John was very afraid:: because he thought:: that he was going to do something to him| (1 AS-unit, 3 clauses)

HOWEVER

If it functions as an ellipted version of an independent clause (such as, I say this because…) write it as a separate AS-unit.

Example:

- |probably John forgot something at the bus|

  |so he brought back to him|

  |because he was giving him another box| (3 AS-units)

2.1.3 Embedding

Embedded clauses belong to the same AS unit but counted as separate clauses

Example:

- |he noticed that:: it was just a nice person:: who was bringing him his wallet back| (1 AS unit, 3 clauses)

2.2 A non-finite Verb element plus at least one other clause element (Subject, Object, Complement, or Adverbial)

Examples:

- |{Uhh} then after that John started to run| (1 AS-unit)
- |He was following him:: to give him the box| (1 AS-unit, 2 clauses) (In this example "to give" is followed by an object, so it is counted as a separate clause)

3 False Starts, repetitions, and self-corrections:

Where these occur, count the final version and exclude the previous versions.

4 Topicalization

Do not exclude topicalized noun phrases unless they are separated by falling intonation and a noticeable pause.

Example:

|The children, Matt and his sister, are preparing a picnic basket| (1 AS-unit)

Syntactic Complexity Measures

Measure	Explanation	Example
Overall complexity	AS-unit length or mean length of AS-unit. Divide the number of words by the number of AS-units in each pruned narrative.	|Matt and his sister planned:: to go somewhere| |so they prepared something for their travelling| |And their mother helped them| 20 words/ 3 AS-units = 6.7
Phrasal complexity	Clause length or mean length of clause. Divide the number of words by the number of clauses in each pruned narrative.	|Matt and his sister planned:: to go somewhere| |so they prepared something for their travelling| |And their mother helped them| 20 words/ 4 clauses = 5
Subordination	Clause/AS-unit ratio. Divide the number of clauses by the number of AS-units in each pruned narrative.	|Matt and his sister planned:: to go somewhere| |so they prepared something for their travelling| |And their mother helped them| 4 clauses/ 3 AS-units = 1.3

Accuracy Measures

Measure	Explanation	Example
Errors per 100 words	Divide the total number of errors by the total number of words in each pruned narrative and then multiply the result by 100%.	|Matt and his sister decides:: to go <u>to</u> (ON) a picnic <u>in</u> (AT) (<u>missing article 'the'</u>) weekend| |And their mother prepares some foods and beverages (<u>missing 'for'</u>) them| (4 errors/ 21 words) × 100% = 19.05
The ratio of error-free clauses	Divide the total number of error-free clauses by the total number of clauses in each pruned narrative.	|Matt and his sister prepares a picnic box| |and their brown dog looks at them| |and the dog <u>saw</u> the picnic box:: and <u>jumped</u> in it| |and Matt and his sister after <u>finish</u> the <u>preparings</u>:: leave <u>from</u> home| 2 clauses without errors/ 6 clauses in total = 0.33

The ratio of error-free AS units	Divide the total number of error-free AS-units by the total number of AS units in each pruned narrative.	|Matt and his sister prepares a picnic box| |and their brown dog looks at them| |and the dog saw the picnic box:: and jumped in it| |and Matt and his sister after finish the preparings:: leave from home| 2 AS-units without errors/ 4 AS-units in total = 0.5

APPENDIX H

INTERCODER RELIABILITY SCORES

	Measures	N	Pearson Correlation Coefficient r
Complexity	Overall Complexity	44	.85
	Complexity by Subordination	44	.87
	Phrasal Complexity	44	.93
Accuracy	Errors per 100 words	44	.94
	The ratio of error-free clauses	44	.88
	The ratio of error-free AS units	44	.89

Note. All measures were calculated per pruned narrative.

REFERENCES

Ahmadian, M. J., & Tavakoli, M. (2010). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research, 15*(1), 35–59.

Ahmadian, M. J., Tavakoli, M., & Dastjerdi, H. V. (2015). The combined effects of online planning and task structure on complexity, accuracy and fluency of L2 speech. *The Language Learning Journal, 43*(1), 41-56.

Albert, A. (2011).When individual differences come into play: The effect of learner creativity on simple and complex task performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 239-266). Amsterdam: John Benjamins Publishing.

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

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

Cumming, G. (2012). *Understanding the new statistics. Effect sizes, confidence intervals, and Meta-Analysis*. New York, NY: Routledge.

de Bot, K. (1992). A bilingual production model: Levelt's 'speaking' model adapted. *Applied Linguistics, 13*, 1-24.

Ellis, R. (1987). Interlanguage variability in narrative discourse: Style in the use of past tense. *Studies in Second Language Acquisition, 9*, 12-20.

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

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

Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167–92). Amsterdam: John Benjamins.

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

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 11*(2), 354-375.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299-323.

Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research, 3*, 215–247.

Foster, P., & Skehan, P. (2013). Anticipating a post-task activity: the effects on accuracy, complexity, and fluency of second language performance. *The Canadian Modern Language Review, 69*(3), 249-273.

Gilabert, R. (2007a). The simultaneous manipulation of task complexity along planning time and [+/- here-and-now]: Effects on L2 oral production. In M. d. P. Garcia Mayo (Ed*.), Investigating Tasks in Formal Language Learning* (pp. 44-68). Clevedon: Multilingual Matters.

Gilabert, R. (2007b). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics in Language Teaching, 45*(3), pp. 215–240.

Gilabert, R., Barón, J., & Llanes, À. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *International Review of Applied Linguistics in Language Teaching, 47*, 367–395.

Gilabert, R., Barón, 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* (pp. 105-138). Amsterdam: John Benjamins Publishing.

Heaton, J. B. (1966). *Composition through pictures*. Essex: Longman.

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

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.

Huck, S. W. (2004). *Reading statistics and research*. Boston: Pearson.

Hulstijn, J., & Hulstijn, W. (1984). Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning, 34*, 23–43.

Ishikawa, T. (2007). The effect of manipulating task complexity along the [+/- here-and-now] dimension on L2 written narrative discourse. In M. d. P. Garcia Mayo (Ed.), *Investigating Tasks in Formal Language Learning* (pp. 136–156). Clevedon: Multilingual Matters.

Iwashita, N., McNamara. T., & Elder, C. (2001). Can we predict task difficulty in an oral proficiency test? Exploring the potential of an information-processing approach to task design. *Language Learning, 51*, 401–436.

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

Kim, Y. (2009). The effects of task complexity on learner–learner interaction. *System, 37* (2), 254–268.

Kim, Y., & Tracy-Venture, N. (2011).Task complexity, language anxiety, and the development of the simple tasks. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 287-306). Amsterdam: John Benjamins Publishing.

Kormos, J. (2000). The role of attention in monitoring second language speech production. *Language Learning, 50*(2), 343-384.

Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kormos, J. (2011). Speech production and the Cognition Hypothesis. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of Language Learning and Performance* (pp. 39-60). Amsterdam: John Benjamins.

Kormos, J., & Trebits, A. (2011). Working memory capacity and narrative task performance. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 267–285). Amsterdam: John Benjamins Publishing.

Kormos, J., & Trebits, A. (2012). The role of task complexity, modality and aptitude in narrative task performance. *Language Learning*, 1-34.

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

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

Kuiken, F., & Vedder, I. (2011). Task complexity and linguistic performance in L2 writing and speaking: The effect of mode. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 91-104). Amsterdam: John Benjamins Publishing.

Larson-Hall, J., & Plonsky, L. (2015). Reporting and interpreting quantitative research findings: What gets reported and recommendations for the field. *Language Learning, 65*(S1), 127-159.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 83-122). Oxford: Oxford University Press.

Long, M. H. (1985). A role for instruction in second language acquisition. In K. Hyltenstam and M. Pienemann (Eds), *Modelling and assessing second language acquisition* (pp. 77–99). Clevedon, Avon: Multilingual Matters.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego, CA: Academic Press.

Long, M. H. (2015). *Second language acquisition and Task-Based Language Teaching*. Oxford: Wiley-Blackwell.

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

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* (pp. 141-173). Amsterdam: John Benjamins Publishing.

Michel, M. C., Kuiken, F., & 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.

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), pp. 11–37.

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.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge: Cambridge University Press.

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

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

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

Plonsky, L., & Oswald, F.L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878-912.

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

Revesz, A. (2011). Task complexity, focus on L2 constructions, and individual differences: A classroom-based study. *The Modern Language Journal, 95*, 162-181.

Robinson, P. (1995). Task complexity and second language narrative discourse. *Language Learning, 45*(1), 99-140.

Robinson, P. (2001a). 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. 287-318). Cambridge: Cambridge University Press.

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

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

Robinson, P. (2007a). Criteria for grading and sequencing pedagogic tasks. In M.P. Garcia Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 7–27). Clevedon: Multilingual Matters.

Robinson, P. (2007b). 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, 45*, 193-213.

Robinson, P. (Ed.). (2011a). *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance*. Amsterdam: John Benjamins.

Robinson, P. (2011b). Task-based language learning: A review of issues. *Language Learning, 61*, 1–36.

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

Robinson, P., Cadierno, T., & Shirai, Y. (2009). Time and motion: measuring the effects of the conceptual demands of tasks on second language speech production. *Applied Linguistics, 30*(4), 533–554.

Sasayama, S. (2015). *Validating the assumed relationship between task design, cognitive complexity, and second language task performance* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.

Sasayama, S., & Izumi, S. (2012). Effects of task complexity and pre-task planning on Japanese EFL learners' oral production. In A. Shehadeh and C. A. Coombe (Eds.), *Task-Based Language Teaching in foreign language contexts: Research and implementation* (pp. 23-43). Amsterdam: John Benjamins Publishing.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Schmidt, R. W. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 287-318). Cambridge: Cambridge University Press.

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

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

Skehan, P. (2016). Tasks versus conditions: Two perspectives on task research and their implications for pedagogy. *Annual Review of Applied Linguistics, 36*, 34-49.

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

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

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review, 50*, 158–164.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H.G. Widdowson* (pp. 125–144). Oxford: Oxford University Press.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*, 371–391.

Swain, M., & Lapkin, S. (2002). Talking it through: two French immersion learners' response to reformulation. *International Journal of Educational Research, 37*, 285–304.

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. 239-273). Amsterdam: John Benjamins.

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

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

Wang, Z. (2014). On-line time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27 - 62). Amsterdam: John Benjamins.

Wang, Z., & Skehan, P. (2014). Structure, lexis, and time perspective: Influences on task performance. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 155 – 186). Amsterdam: John Benjamins.

Wickens, C. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3*, 159–177.

Wickens, C. (2007). Attention to the second language. *International Review of Applied Linguistics, 45*, 177–191.

Willis, D. (1996). *A Framework for Task-Based Learning*. London: Longman.

Willis, D., & Willis, J. (2001). Task-based language learning. In R. Carter and D. Nunan (Eds.), *The Cambridge guide to teaching English to speakers of other languages*. Cambridge: Cambridge University Press.

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