Exploring the role of proficiency as a mediator of the relationship between working memory and text comprehension in the L2

Dissertation submitted to the institute of educational sciences

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

English Language Education

By

Bora Demir

Yeditepe University

2015

T.C.
YEDİTEPE ÜNİVERİSİTESİ EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Exploring the role of proficiency as a mediator of the relationship between working memory and text comprehension in the L2

by

Bora DEMİR

Approved by:

Doç. Dr. Gülcan ERÇETİN (Danışman)	(İmza)
Yard. Doç. Dr. Jing Jing KARABACAK (Danışman)	(İmza)
Prof. Dr. Ayşe AKYEL (Üye)	(İmza)
Doç. Dr. Houssein Farhady (Üye)	(İmza)
Yard. Doç. Dr. Zeynep B. Koçoğlu (Üye)	(İmza)

Date of Approval by the Administrative Council of the Institute of Educational Sciences

16.03.2015

VITA

BORA DEMİR

EDUCATION

2008 - 2015	Ph.D., Yeditepe University, Department of English Language Education, Çanakkale, Turkey.
2002 - 2004	M.A., Çanakkale Onsekiz Mart University, Department of English Language Teaching, Çanakkale, Turkey.
1996 – 2000	B.A., Çanakkale Onsekiz Mart University, Department of Foreign Language Education, Çanakkale, Turkey.

WORK EXPERIENCE

2003 - Present	Instructor, Çanakkale University, School of Foreign Languages, Preparatory Program, Çanakkale, Turkey.
2000 - 2003	English Language Teacher, Ministry of Education, Çanakkale, Turkey.

PUBLICATIONS

Demir, B. (2011). Undergraduate FL Learners' Perceptions of Native and Non-Native Language Instructors. The International Journal of Research in Teacher Education 2011, 1(1): 01-29

Demir B. & Dikilitaş, K. (2012). Native and Non-Native Perceptions on a Non-Native Oral Discourse in an Academic Setting. Turkish Online Journal of Qualitative Inquiry. July 3(3), pp. 33-55

CONFERENCE PROCEEDINGS

Demir, B. (2010). The effectiveness of different types of direct written corrective feedback. The Sixth ELT Reseach Conference, İzmir - Turkey.

Demir, B (2010). A proactive evaluation of EAP reading course for ELT and ELL preparatory students at Çanakkale Onsekiz Mart University. 2nd foreign language teaching symposium: a proactive look at english language teaching programs in the preparatory schools of universities, May 28, 2010 Çankaya University -Turkey

Zehir Topkaya, E & Demir,B. (2009). The effects of a course about culture and culture teaching: Do pre-service teachers' conceptions change? X. Uluslararası Dil, Yazın ve Deyişbilim Sempozyumu (The 10th International Language, Literature and Stylistics Symposium), 3-4-5 Kasım/November 2010, Gazi Üniversitesi, Ankara

---

CONFERENCE PRESENTATIONS

Demir,B "Phonological Awareness of a Turkish-English Bilingual and a Turkish Monolingual Child: Is there a bilingual advantage?"The 1st International Educational Research Conference, May, Çanakkale

Demir,B."An investigation into the motivations and learning behaviours of first year university students learning English as a foreign language." 5th International ELT Research Conference. 23-25 May, Çanakkale

Erten, İ. H. & B. Demir. "An Investigation into the Age Factor on the Motivation and Attitudes of Young Learners" IV. İngiliz Dili Eğitimi Araştırma Konferansında sunulmuş Bildiri. 26-28 Mayıs 2005 Çanakkale Onsekiz Mart Üniversitesi, Çanakkale, Türkiye.

Demir, B. (2003). "Motivation and attitudes of primary school students in Canakkale learning English as a foreign language." poster presented at the 3rd International ELT Research Conference, 22-24 May 2003, Çanakkale- Turkey.

Demir, B. (2002). "The expectations of primary school students from English lessons." Poster presented at The Second International ELT Research Conference: Reflections of Twentieth Century ELT Studies on Language Teaching and Teacher Training" Çanakkale: Onsekiz Mart University.

RESEARCH AND TEACHING INTERESTS

Language Teacher Education, Computer Assisted Language Learning
Psychology of learning, Memory and Language

ACKNOWLEDGEMENTS

I would like to express my appreciation to the people who have helped me write this dissertation.

I would like to express my deepest gratitude and thanks to Assoc. Prof. Dr. Gülcan Erçetin who has been a tremendous supervisor for me. I felt her excellent guidance, encouragement, patience, and support in all stages of my research. Without her invaluable guidance, this study would not have been possible.

I also owe a debt of gratitude to Prof. Dr. Ayşe Akyel, whose academic and personal support truly provided me inspirations both for this dissertation and for my future carrier. I am genuinely grateful for her encouragement and trust in me.

My appreciation is also to my second supervisor Asst. Prof. Dr. Jing Jing Karabacak for her valuable feedback and support.

I would like to thank my committee members, Assoc. Prof. Dr. Hossein Farhady for providing invaluable insight on this dissertation, and Assist. Prof. Dr. Zeynep Koçoğlu for her helpful and kind manners.

I would also like to acknowledge Prof. Dr. Dinçay Köksal for his support and help throughout my academic journey.

I must also acknowledge Dr. Faruk Kural for being an excellent friend, colleague, and a great supporter since I knew him.

I extend my special thanks to all my teachers who have made contributions to my life and academic career.

I am also grateful to the students who participated in the study and helped me gather the data.

I would also like to thank my father, my mother and sisters. They have always supported me and encouraged me with their best wishes.

Finally, my deepest gratitude goes to my wife, Ece Demir, and my son, Deniz Demir. They were always there cheering me up and stood by me through this long journey. Thanks for tolerating and believing in me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Exploring the role of proficiency as a mediator of the relationship between working memory and text comprehension in the L2

By
Bora DEMİR

The main purpose of this study is to find out whether there is a relationship between WM capacity and listening/reading, and whether this relationship is mediated by learners' proficiency level in the L2. In addition, with a mixed methods design where the effect of proficiency level is investigated both as a between groups factor but also as a within groups factor, this study aims to make a cross-sectional comparison of different proficiency groups (intermediate and upper-intermediate) at a certain point in time as well as developmental comparison of the same proficiency group before and after instruction. The participants consist of Turkish University students from Çanakkale Onsekiz Mart University. The Working Memory Capacity was measured by, Reading Span Test (English version), Reading Span Test (Turkish version), and Operation Span Test. To measure the participants' reading and listening comprehension, this study made use of the Certificate in Advanced English (CAE). Descriptive statistics, and 2x2x2 mixed ANOVAs were conducted separately on listening and reading scores with time of testing (pre vs. post) as the repeated measures factor, WM capacity (low vs. high), and proficiency level (intermediate vs. upper-intermediate) as between groups factors. The results indicated no significant effects of WM on reading comprehension. On the other hand, its effect on listening comprehension was significant when WM was measured by L1 or L2 RST. Moreover, it was observed that the difference between low- and high-WM capacity learners was greater on the pre-test compared to the post-test for listening comprehension but not for reading comprehension since the interaction between WM and time of testing was significant for the former, not for the latter. Also, when WM capacity is measured by the RST in the L2, the results for listening comprehension indicated a significant interaction between time of testing and WM. However, the difference between the low- and high-WM participants is negligible in the upper-intermediate group in terms of both listening and reading comprehension when WM is measured by the RST in the L1. Results are interpreted in the light of Working Memory Model (Baddeley, 2000), and Declarative/Procedural (DP) Model of language acquisition (Ullman, 2001a, 2001b).

Key Words: Working Memory Capacity, Reading Comprehension, Listening Comprehension, Language Proficiency

KISA ÖZET

İşler Bellek ve İkinci Dilde Metin Anlama Arasındaki İlişkinin Araştırılmasında Dil Yetisinin Rolü

Bora DEMİR

Bu çalışmanın esas amacı, işler bellek kapasitesi ve ikinci dilde okuma/dinleme arasında bir ilişki olup olmadığını ve bu ilişkinin öğrencinin ikinci dildeki dil yetisi seviyesiyle olan ilişkisini bulmaktır. Bunun yanı sıra, bu çalışma, dil yetisi seviyesinin etkisi karışık metot yöntemiyle hem grup içi hem de gruplar arası faktörler olarak araştırılmasıyla, farklı dil yetisi seviyelerinin (orta seviye ve orta ileri seviye) kesitsel karşılaştırmasını amaçlamıştır. Üniversitesi öğrencilerinden olan katılımcıların işler bellek kapasiteleri Okuma Aralığı Testleri (Türkçe ve İngilizce) ve İşlem Aralığı Testi ile ölçülmüştür. Katılımcıların okuduğunu ve dinlediğini anlamalarını ölçmek için CAE testi kullanılmıştır. Okuma ve dinleme skorları üzerinde ayrı ayrı betimleyici istatistikler yapılmıştır. Ayrıca, tekrarlanan ölçümler faktörü olarak ölçme zamanı, gruplar arası faktörler olarak da işler bellek kapasitesi ve dil yetisi olarak 2x2x2 karışık ANOVA testleri yapılmıştır. Sonuçlar, işler belleğin okuduğunu anlama üzerinde önemli bir etkisinin olmadığını göstermiştir. Ancak, Türkçe ve İngilizce Okuma Aralığı Testleri ile ölçüldüğünde işler hafızanın dinlediğini anlama üzerindeki etkisinin önemli olduğu bulunmuştur. Ayrıca, düşük ve yüksek işler bellek kapasiteli öğrenciler arasındaki farkın, okuduğunu anlamaya göre, son teste kıyasla ön testte daha büyük olduğu gözlemlenmiştir. Ancak, bu farkın, okuduğunu anlama için, işler hafıza ve ölçüm zamanı arasındaki ilişkinin son test için değil, ön test için önemli olduğu da gözlemlenmiştir. Ayrıca, öğrencilerin işler bellek kapasiteleri ikinci dildeki işler hafıza Okuma Aralığı Testi ile ölçüldüğünde, dinlediğini anlama açısından, ölçüm zamanı ve işler hafıza arasında önemli bir etkileşim olduğu belirlenmiştir. Ancak, ikinci dildeki işler hafıza Okuma Aralığı Testi ile ölçüldüğünde, düşük ve yüksek işler hafızalı öğrenciler arasındaki farkın orta üstü öğrenciler için hem dinlediğini anlama hem de okuduğunu anlama açısından önemli olmadığı bulunmuştur.

Anahtar Kelimeler: İşler Bellek Kapasitesi, Okuduğunu Anlama, Yazdığını Anlama, Dil Becerisi

CHAPTER 1

INTRODUCTION

1.1. Background of the study

During the years of teaching experience, I have observed that almost all language learners have difficulties in achieving various aspects of language. Although studies on individual differences have managed to explain an important amount of variance, there still exist unanswered issues concerning the question of why some language learners are more successful than others. Hence, understanding the reasons of success and failure in language learning depends partly upon understanding the cognitive processes that take place in the human mind. Understanding the processes underlying the language practice can be used to develop better methods of teaching, assessment, and even a general understanding of the L2 education.

Memory, as an important part of our cognitive capacity, has become one of the most important concepts to emerge from cognitive psychology for 40 years now. In a broader sense, memory is usually defined as one of the main cognitive processes underlying thinking and learning. The idea that the human mind cannot operate without the support of a temporary memory system led the researchers to re-conceptualize the existing memory systems by inserting a new component, working memory. However, as a relatively new concept within the study of memory organization, the nature and functions of the WM system has become the focus of a large number of studies over a few decades.

Today it is well established that holding and processing information to accomplish cognitive tasks is not achievable without the support of a temporary memory system. Hence, the effort to understand how human brain temporarily stores and processes information is essential in understanding the cognitive aspects of language learning.

What makes the concept of Working Memory (WM) critical is its limited capacity in the elucidation of key cognitive tasks, such as problem solving and learning. WM's relations with various aspects of language learning mainly arise from its limited capacity. Due to the essential role of WM in cognitive performance and learning, better learning partly depends on the individual's WM capacity. In this sense, to be able to explain the reasons of success and failure in language comprehension in a second language (L2) largely depends on understanding the role of WM in language related tasks and its interaction with other factors such as proficiency level.

1.2. Purpose of the study

The present study has two purposes. The first one is to find out whether there is a relationship between WM capacity and listening/reading comprehension of the participants. The second purpose of this study is to find out whether this relationship is mediated by learners' proficiency level in the L2.

To this end, the following research questions and hypotheses are addressed:

Research question 1: Are there differences between high- and low-WM participants in terms of their listening and reading comprehension in the L2?

Research question 2: If yes, do proficiency level and instruction in the L2 mediate the performance differences between high- and low-WM participants?

In view of available research findings pointing to a meaningful relationship between WM capacity and reading comprehension (e.g., Alptekin & Erçetin, 2009, 2010, 2011; Harrington & Sawyer, 1992; Miyake & Friedman, 1998; Walter, 2004) as well as WM capacity and oral language comprehension/production (e.g., Juffs & Harrinngton, 2011; Kormos & Safar, 2008; Mackey, Adams, Stafford, & Winke, 2010), learners with larger capacities were expected to have better listening and reading comprehension scores (Hypothesis 1). Since WM operates through conscious controlled processes, it was hypothesized that WM would be more influential at lower levels since WM's influence weakens as skills get more automatic. In other words, since WM operates through conscious controlled processes, it was hypothesized to be more influential at lower levels (Hypothesis 2).

1.3. Significance of the Study

Existing literature points to the central role of WM in text comprehension. Since L2 learners encounter substantial challenge in comprehending written or oral texts, their cognitive processing largely depend upon their WM capacities. However, what remains unclear is the role of language proficiency in explaining the relationship between WM and text comprehension in the L2. For that reason, understanding the link between WM capacity and proficiency level in L2 text comprehension will help us identify the role of each factor in success and failure during reading and listening activities.

Thus far, numerous studies have tried to reveal the connection between WM and text comprehension by indicating that WM accounts for a large amount of variance both in reading and listening comprehension. However, the role of language proficiency in explaining the relationship between WM and text comprehension is still unclear.

Hence, the present study is significant in that it aims to find out whether there is a relationship between WM capacity and listening/reading, and whether this relationship is mediated by learners' proficiency level in the L2.

Another significance of this study is that it provides data from various measures of WMC, rather than relying on one single measure. As such, WM capacity is measured not only through a reading span task (RST) both in the L1 and L2 but also through an operation span task (OST).

In addition, the study has a mixed methods design where the effect of proficiency level is investigated both as a between groups factor but also as a within groups factor. Such a design allows cross-sectional comparison of different proficiency groups at a certain point in time as well as developmental comparison of the same proficiency group at points over time.

This study will also contribute to existing literature, since little research exists on WM's relationship to L2 listening and reading. Besides, these two skills have been rarely investigated in a single study in L2 research in relation to WM capacity.

1. 4. Overview of Methodology

1. 4. 1. Participants

The participants consist of Turkish University students from Çanakkale Onsekiz Mart University. All of the participants were native speakers of Turkish. Their age ranged from 18 to 22. They form two groups at different proficiency levels: a) the prep students, b) the junior students from the departments of English Language Teaching (ELT) and English Language and Literature (ELL). 111 female and 62 male participants were randomly selected from both groups among a target population size of 130 prep and 135 first year students.

1. 4. 2. Setting

The present study was conducted at two different schools of Çanakkale Onsekiz Mart University. The freshmen were from the Faculty of Education, Department of English Language Teaching (ELT), while the preparatory participants were from the School of Foreign Languages.

The ELT program aims to improve the language proficiency of students; to equip students with the contemporary methods and techniques of language teaching; to increase their awareness in intercultural differences to the extent that they become teachers of English who are innovative and knowledgeable professionals and who will be able to be practicing teachers not only in Turkey but also elsewhere in the world; enable them to use the appropriate methods, techniques, materials, resources and technology in the process of teaching and testing English and arrange the classroom environment accordingly, being aware of the universality of social rights, social justice, protection of cultural values, quality culture and environmental protection, occupational health and safety issues.

This is a First Cycle (Bachelor's Degree) program. The students who successfully complete the program are awarded the degree of Bachelor of Arts in English Language Teaching. In order to graduate from this undergraduate program, the students are required to succeed in all of the courses listed in the curriculum of the program by getting the grade of at least DD/S with a minimum of 240 ECTS to have a Cumulative Grade Point Average (CGPA) of 2.00 out of 4.00.

Some of the program key learning outcomes can be listed as, to act in accordance with social, scientific, cultural and ethical values at all phases of performing his/her teaching profession, to have the knowledge and skills of being able to use the appropriate methods, techniques, materials, resources and technology in the process of teaching English, to use the appropriate instruments/tools, approaches and methods, techniques when measuring and testing English language skills and knowledge, and to prepare a teaching plan and assessing accordingly, determining the needs of different students and their language learning strategies and styles, and considering their developmental stages.

Admission of national students to the ELT education program is contingent on success on the nationally centralized Student Selection Examination (YGS and LYS) conducted by the Student Selection and Placement Centre (OSYM). Graduates can work as teachers of English at schools of the Ministry of National Education and as lecturers at higher education institutions, and students who successfully complete this undergraduate program may, on their request, progress to the graduate program.

On the other hand, the preparation school preparatory programs offer a one year preparatory language program which aims at assisting the students who have failed the Language Proficiency Exam in developing the necessary knowledge, skills and confidence to effectively use and further improve the target language throughout their academic studies and in work-related and social environments. The students in the preparatory school attend 8 hours of Basic English grammar, 7 hours of listening and speaking, 7 hours of reading, and 4 hours of writing courses per week. These learners are considered to be intermediate learners of English. As for the reading course, the program aims to improve the language proficiency of students to be able to distinguish between academic and non academic reading texts, analyze different types of discourse, read for main ideas and important details, understand by making inferences, understand the syntactic structure of a sentence and clause, discourse markers, and lexical and grammatical cohesion.

As for the writing, the major aims of the program are to present the important organizational principles of good academic writing, such as writing topic sentences, thesis statements, supporting details, maintaining unity and coherence, using transitions within the ideas and paragraphs, writing clear introductory and concluding paragraphs. Also, help students apply these organizational principles to major rhetorical forms such as classification and clustering, exemplification, cause and effect, comparison and contrast, supporting an argument, problem and solution, personal expression, summarizing, paraphrasing and essay writing, and provide extensive practice in all steps of the writing process, including gathering information from relevant sources, pre-writing, drafting, revising and editing. The major aims of the program in terms of listening the major aims are to ensure students acquire the necessary listening skills so that they will be able and to listen and understand video cassettes, films, debates, conferences seminars, panels, discussions. As for the speaking course, the program aims at understanding and responding to spoken English in social and academic environments, becoming clearly comprehensible through the use of appropriate and rich vocabulary, and also able to exchange ideas.

1.4. 3. Data Collection Instruments

To measure the participants' reading and listening comprehension, this study made use of the Certificate in Advanced English (CAE) by Cambridge University Press, which is an advanced general English examination provided by University of Cambridge ESOL Examinations in England. In the CEFR, CAE is ranked at C1 and C2 levels. Both the reading and listening exams were taken from Exam Essentials: CAE Practice Tests (Osborne, 2006).

This study also made use of 3 tests to measure the working memory capacity of the participants, namely, Reading Span Test (English version), Reading Span Test (Turkish version), and Operation Span Test. Both for Turkish and English versions, this study made use of a modified version of the original RST developed by Daneman and Carpenter (1980). The OST used in the study was a version suggested by Kane et al. (2004). The memory tests were administered on the computer in a computer lab.

1. 4. 4. Data Analysis

The storage scores (i.e., the total number of accurately recalled sentence-final words in the L1 and L2 RSTs and the total number of accurately recalled words associated with the operations in the OST) constituted the WM measure. For each WM measure, the participants were divided into low- and high-WM groups based on a median-split procedure on the storage scores. Thus, the participants below the median were categorized into low-WM group while those above the median were categorized in high-WM group.

In order to answer the research questions, descriptive statistics were first obtained. Then, 2x2x2 mixed ANOVAs were conducted separately on listening and reading scores with time of testing (pre vs. post) as the repeated measures factor, WM capacity (low vs. high), and proficiency level (intermediate vs. upper-intermediate) as between groups factors. The statistical procedures were carried out via IBM SPSS Statistics 20.0.

1.5. The Organization of Chapters

The present study is composed of five chapters. The first chapter comprises, a general introduction of the study, the purpose of the study, the research questions addressed in the study, the overview of methodology (participants, setting, data collection instruments and data analysis), significance of the study are presented. The second chapter presents a detailed literature review including theories of working memory with a specific focus on The Working Memory Model and its components. Following that, the chapter provides background information about working memory and L1/L2 research. Chapter III, provides information on the methodological issues focusing on the participants, the research design, instruments for data collection and the procedures for data analysis. Chapter IV reports the results with reference to the research questions. And finally, discussions and conclusions related to the findings are presented in Chapter V, followed by limitations and suggestions for further research. References and appendices are provided at the end of the study.

CHAPTER 2

REVIEW OF LITERATURE

2.1. Introduction

The review of literature of this study is discussed in four sections. In the first part, the concept of memory is initially discussed. Then, the chapter provides a comprehensive chronological overview of WM models focusing specifically on the construct and components of WM model. Following that, the chapter presents various tests to measure WM capacity. Finally the studies related to L2 reading and listening research are discussed.

2.2. Working Memory

Memory is one of the most important concepts of cognitive science since many aspects of human life depend on it. Rather than a single unitary system, memory is an array of interacting systems, each capable of encoding or registering information, storing it, and making it available by retrieval. Without this capability for information storage, we could not perceive adequately, learn from our past, understand the present, or plan for the future (Baddeley, 1999).

The recognition of memory and its role in life and learning can be traced back to the ancient Greeks. Because of his work on memory, *De Memoria et Reminiscentia*, Aristotle is usually attributed with being the first person to propound the theory of the association of ideas as the basis of memory (Morris & Gruneberg, 1994). Nevertheless, the oldest theoretical distinction between primary and secondary memory was proposed by James (1890). According to James, primary memory is equated with the current contents of consciousness and suffers from a limited capacity, and secondary memory, in contrast, is thought to consist of memories of the distant past and to be unlimited in capacity (Conway, Jarrold, Kane, Miyake & Towse, 2007; Morris & Gruneberg, 1994)

However, it was the mid-twentieth century when psychologists were able to identify and work on distinct dimensions and functions of memory (Dehn, 2008). As a result of debate on the types, dimensions and functioning of the concept of memory, a whole range of memory models were proposed during the 1960s. These models were basically composed of three types of memory namely; *sensory memory*, s*hort-term memory,* and *long-term memory* (Pickering, 2006). The model which received more attention was Atkinson and Shiffrin's 1968 model (Fig.1) and it contained a detailed analysis concerning the structure and functioning of human memory. According to Atkinson and Shiffrin's dual process model, memory is made up of a series of stores. It was assumed that information comes in from the environment through a parallel series of brief sensory memory stores, or registers, and goes into a common short-term store. This system is considered as capable of manipulating information and relating it to long-term storage. Without it, the learning of new material or the recollection of old information would be impossible. (Essentials of human memory, Baddeley, 1999)

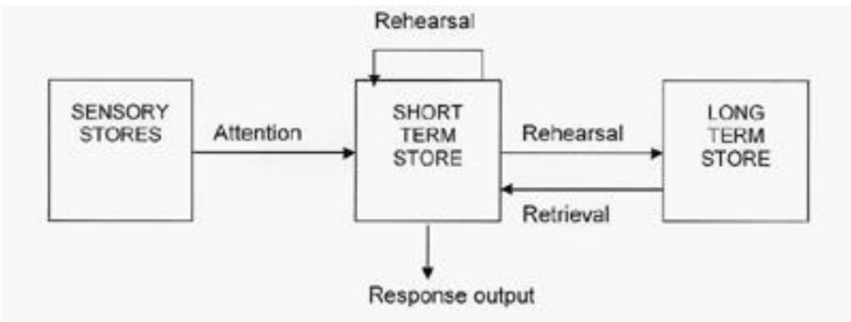

Figure 1. Conceptualization of Multi Store Model of Memory by Atkinson and Shiffrin (1968)

In the model both Short Term Memory (STM) and Long Term Memory (LTM) were conceptualized as passive stores. Although many earlier conceptions of human memory system assumed that STM is simply the activated portion of LTM, they did not present a structure clarifying a locus for manipulation, transformation, and processing of information beyond rehearsal (Feng, 2009).

However, in a later study, Atkinson and Shiffrin (1971) proposed that the flow of information through the short-term store and the subject's control of that flow of information were central to the system underlying human memory. They explained the concept of WM as a system in which decisions are made, problems are solved, and information flow is directed.

Baddeley and Hitch (1974) proposed the concept of WM to overcome the problems in the early models. According to Baddeley and Hitch, short-term storage of information must be considered as part of a more complex system involved in the execution of a specific task. This system has the ability to store and process information simultaneously. According to Baddeley and Hitch (1974) WM refers to a limited capacity system of temporary storage and manipulation of input that is necessary for complex tasks such as comprehension and reasoning. Furthermore, it involves a number of different subsystems, each related to the specific nature of the information to be processed (Baddeley, 1986).

Figure 2. A schematic representation of model of WM (Baddeley and Hitch, 1974)

As mentioned above, the concept of WM was initially proposed by Baddeley and Hitch (1974) and developed by Baddeley (1986). In fact, Baddeley and Hitch proposed the model in order to expand a more accurate model of STM. This model was different from Atkinson and Shiffrin (1968) model, since it offers a far more flexible characterization of STM. One of the most important developments within cognitive psychology over the past 20 years has been the shift away from viewing STM as a receptacle for lists of unrelated items, and towards seeing it as a dynamic system that plays a critical role in a wide range of complex cognitive activities (Gathercole, 1994). The origins of WM are in theories of STM that focused on the temporary storage of information, rather than focusing on the role that temporary storage or transformation played in general cognition (Andrade, 2001). According to Swanson (2000), it is important to note that there has always been an emphasis on WM’s limited capacity to keep information while simultaneously processing the same or other information. This conceptualizes WM distinct from STM.

There have been conflicting views among cognitive psychologists and memory researchers on the definition and conceptualism of STM and WM. According to Baddeley (1996) the concept of WM represents a modification and extension of an earlier concept, short-term memory proposed by Atkinson and Shiffrin (1968). Baddeley differentiates the concept of WM from STM in two respects: the number of subsystems involved during the process, and the considerable emphasis on its functional role in other cognitive tasks such as learning, reasoning, and comprehension.

Some researchers have used the two terms as interchangeable or consider one to be a subtype of the other (McDougall, Hulme, Elllis, & Monk, 1994). Other theorists and researchers argue that WM and STM are distinguishable constructs (Gathercole, 1998; Unsworth & Engle, 2007; Dehn, 2008). Nevertheless it is generally acknowledged that the two concepts are distinct in nature. Unlike WM, STM only refers to the short-term storage of information, and it does not involve the manipulation or organization of material held in memory. STM differs from WM in many ways. While STM passively holds information, WM has stronger relationships with academic learning and with higher-level cognitive functions. STM automatically activates information stored in long-term memory. On the other hand, WM consciously directs retrieval of desired information from LTM. STM can operate independently of LTM, but WM operations rely heavily on LTM structures. STM retains information coming from the environment while WM retains products of various cognitive processes (Dehn, 2008).

There has been numerous efforts to conceptualize and define WM. Baddeley and Hitch (1974) argued that WM is a flexible and limited-resource system with storage and processing capabilities which exist in a traded off fashion. In this system, small memory loads are handled alone by a peripheral phonemic buffer, leaving central processing unaffected, whereas larger loads require additional resources of a central executive. Thus, WM was proposed to be a dynamic system that enabled active maintenance of task-relevant information in support of the simultaneous execution of complex cognitive tasks.

Consistent with the assumption that short-term storage of information must be considered as a part of a more complex system involved in the execution of a specific task. Baddeley (1986) defined WM as comprising a number of different subsystems, each related to the specific nature of the information to be processed.

According to Dehn (2008), WM supports human cognitive processing by providing an interface between perception, STM, LTM, and goal-directed actions. WM is the term that cognitive psychologists use to describe the ability to simultaneously maintain and process goal relevant information. As the name implies, the WM concept reflects fundamentally a form of memory, but it is more than memory, for it is memory at work, in the service of complex cognition (Juffs & Harrington, 2011). Similarly, Baddeley and Logie (1999) defined the term as storing and processing information while performing higher order cognitive tasks such as comprehension, learning and reasoning.

WM has also been defined as a cognitive system that contains a limited computational space in which materials can be temporarily stored, monitored and manipulated (Baddeley, 1986; Just & Carpenter, 1992). According to Juffs and Harrington (2011), WM is not an isolated term within the concept memory. Rather than the capacity to store the products of our experience in the world, it is a system with multiple components, or a collection of interrelated processes, that carries out several important cognitive functions.

A fundamental characteristic of WM is that it has a restricted capacity, which limits cognitive performance. Individuals with larger capacity typically perform better than individuals with smaller capacity on a variety of cognitive tasks including complex learning, reading and listening comprehension. In other words, there exists a variation in WM capacity and this variation is important to everyday cognitive performance. (Juffs & Harrington, 2011). Oberauer et al. (2003) also defines WM as a set of limiting factors for performance in complex cognitive tasks. As said by Miyake and Shah (1999), although there is variety of definitions of WM, there is consensus among WM researchers that it comprises those mechanisms or processes that are involved in the control, regulation, and active maintenance of task-relevant information in the service of complex cognition. Hulme and Mackenzie (1992) WM defined the concept as the use of temporarily stored information in the performance of more complex cognitive tasks. In general, WM is viewed as a comprehensive system that unites various short- and long-term memory subsystems and functions (Baddeley, 1986).

One of the greatest accomplishments of the human mind is perhaps its ability to mentally maintain information in an active and readily accessible state, and selectively process new information simultaneously (Conway, Jarrold, Kane, Miyake & Towse, 2007). In this context, WM refers to the mental processes responsible for the temporary storage and manipulation of information in the course of on-going processing (Juffs & Harrington, 2011).

Although generally accepted as a separate concept of memory, Cowan (2001) defines WM as an activated portion of LTM, and Just and Carpenter (1992) defined WM as a single and limited pool of cognitive resources in which all processes draw and are in competition for sharing that resources. Nevertheless, there is a general acceptance of the need to assume both of a general executive system and specific verbal and visual systems (Baddeley & Hitch, 2001).

2.3. Models of Memory

For the last four decades, researchers have conceptualized various models of memory. Memory models in modern terms, first started with the Atkinson-Shiffrin's *Multi Store Model of Memory* (1968) as an elaboration of the information processing model originally proposed by Broadbent (1958). According to this model, information proceeds from temporary short-term storage to more durable long-term memory. Atkinson and Shiffrin view short-term memory as the workspace for long-term learning. The model simply proposes that information is detected by the sense organs and enters the *sensory memory*. This information enters the STM. Information from the STM is transferred to the LTM only if that information is rehearsed. If rehearsal does not occur, then information is forgotten, lost from short term memory through the processes of displacement or decay.

Atkinson-Shiffrin's Multi Store Model of Memory was criticized for devoting too much importance on structure while ignoring the processes (Baddeley, 1996a). However, Atkinson-Shiffrin's three-part division still provides a useful framework for interpreting memory performance, and it is consistent with the information processing model that persists to this day (Dehn, 2008).

The Levels of Processing Model of Memory (Craig and Lockhart, 1972) was proposed partly as a result of the criticism on the multi-store model of memory. Instead of concentrating on the structures involved, this theory concentrates on the processes involved in memory. Unlike the multi-store model, it is a non-structured approach. The basic idea is that memory is really just what happens as a result of processing information. However, according to the model, memory is just a by-product of the depth of processing of information, and there is no clear distinction between STM and LTM.

2.4. The Working Memory Model

The Working Memory Model was proposed by Baddeley and Hitch in 1974. After studying the former models, they believed that the models were overly simplistic in terms of STM conceptualization. They defined WM as a system for the temporary holding and manipulation of information during the performance of a range of cognitive tasks such as comprehension, learning, and reasoning (Baddeley, 1986). According to Baddeley and Hitch (1974), during the performance of a task that involves two different systems, such as the articulatory control system and the visuo-spatial sketchpad, they can process both together interactively or they can do either on its own. Baddeley and Hitch provide this argument as evidence which supports the idea that WM is composed of different components. The separation of storage from processing is the key feature that distinguishes the WM model from other models.

In this model, WM consists of three basic stores: the phonological loop, the visuo-spatial sketchpad and the central executive, which controls the other two subsystems, referred to as slave systems.

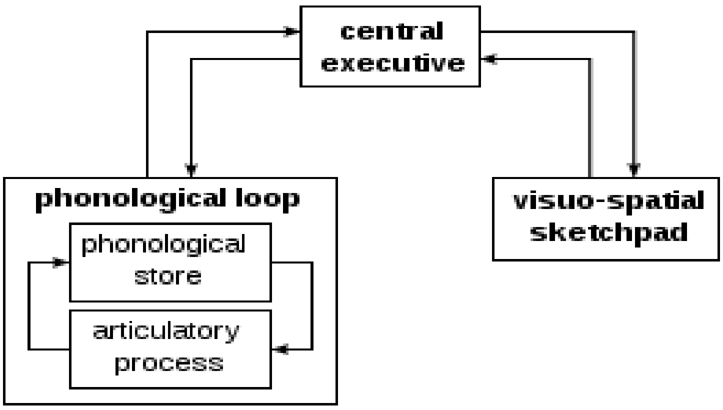

Figure 3. Conceptualization of WM model (Baddeley & Hitch, 1974)

In 2000, Baddeley and Hitch expanded this model with a new component, *multimodal episodic buffer*. According to this model, information is stored in and retrieved from a set of buffers, each specialized for a different kind of information. One of these buffers is for verbal information, one for spatial information, one for visual information, and one for episodic information (Baddeley, 2001). Basically, the model is composed of three sub-components, namely the visuo-spatial sketchpad, the episodic buffer, and the phonological loop. Figure 4 illustrates Baddeley's 2000 model of WM. This study is contextualized within the theoretical framework of Baddeley's WM model.

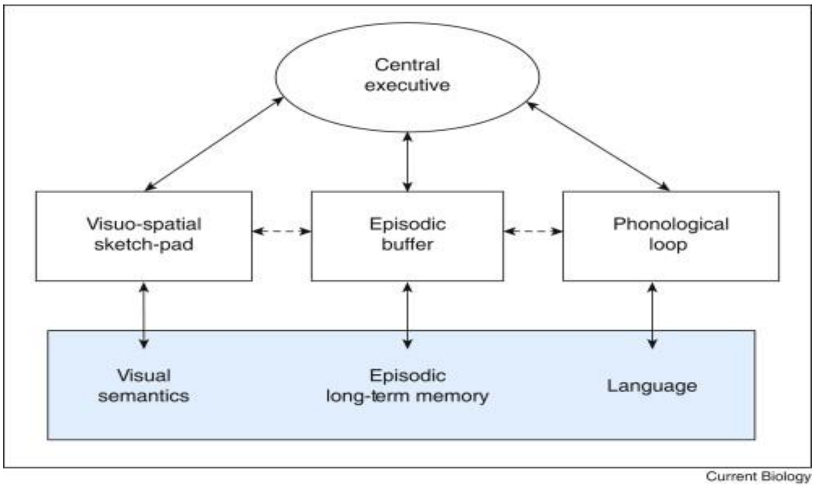

Figure 4. Baddeley's model of WM (2000)

2.4.1. The Phonological Loop

As one of the three basic stores of Baddeley's WM Model, The Phonological Loop (PL) is specialized for the short-term storage and processing of verbal and acoustic information. The terminology of the term has been evolved, since the earlier conception of the component was named by Baddeley and Hitch (1974) as 'phonemic buffer', and as 'articulatory loop' by Baddeley (1986). Basically, as a part of the WM, the PL handles phonological information and rehearses verbal input. According to most models of short-term memory, one characteristic frequently assigned to short-term memory is its reliance on speech coding. Baddeley separated this aspect of memory from the rest and postulated the PL as a slave system (Baddeley, 1999).

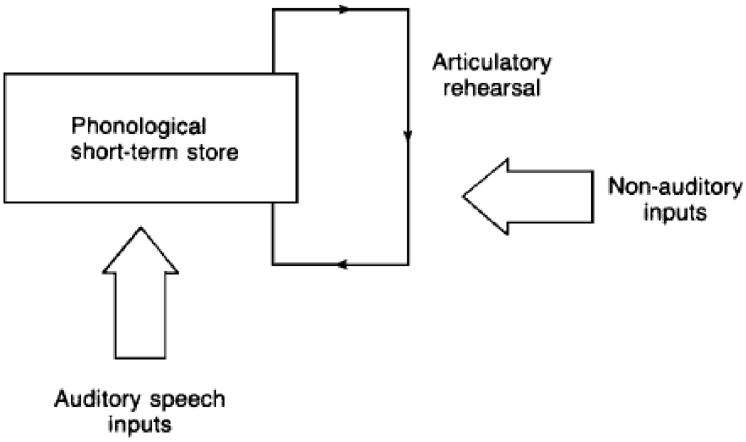

Figure 5. The Phonological Loop Model (Baddeley, 1986)

The PL is composed of two separable components: a *'phonological store'* and '*subvocal rehearsal process'*. The phonological store holds traces of acoustic or speech based material. The input in this store lasts about two seconds unless it is maintained through the use of the second subcomponent, articulatory subvocal rehearsal. Prevention of subvocal rehearsal results in rapid decaying of the information. In other words, the rehearsal system can refresh the information and maintains its storage by subvocal articulation (Morris & Gruneberg, 1994). A phonological loop is thought to have a capacity of 5-9 units (Miller's so-called "magic number" 7+/-2).

Today, it is well established that much more than a slave system the PL may play a key role in the acquisition of vocabulary, and also is vital for learning a second language (Miyake and Shah, 1999). The assumption that such a slave system existed was supported by three clusters of evidence (Baddeley, 2000).

2.4.1.1. The Effect of Phonological Similarity

Lists of words that sound similar are more difficult to remember than words that sound different. Therefore, verbal information is coded largely phonologically in

WM. According to Baddeley (1996), poorer WM for similar sound stimuli supports the existence of a temporary storage system specifically for speech-based or ‘phonological’ input. In other words, the memory traces for similar sound stimuli are assumed to be harder to discriminate at recall (Baddeley, 1990, p. 72).

2.4.1.2. The Effect of Articulatory Suppression

Articulatory suppression has been found to interact with the phonological similarity effect in a highly selective manner which is entirely consistent with the PL model (Morris & Gruneberg, 1994). Memory for verbal material is impaired when subjects are asked to repeat a simple syllable aloud. This process blocks the articulatory rehearsal process, thus leaves memory traces in the phonological loop to decay. According to Morris and Gruneberg (1994) articulatory suppression prevents any visual material from being recoded via subvocal rehearsal into a phonological form suitable for the phonological store. Furthermore, without access to the phonological store, the phonological similarity effect cannot occur.

2.4.1.3. The Word Length Effect

The word length effect is evidence for the finding that short words are easier to remember since the rehearsal of the long words takes longer. According to Baddeley, Thomson & Buchanan (1975), the ability to recall words also depends on a number of characteristics of these words. In other words, lists of short words are better recalled than lists of long words, even though the words of different lengths were matched for variables such as their frequency of occurrence within the language, and their abstractness.

2.4.2. The Visuo-Spatial Sketchpad

Although theoretical development of a detailed model of the sketchpad is less advanced than PL component of the WM model, the Visuo-Spatial Sketchpad (VS) can be defined as a slave system that is specialized for the processing and storage of visual and spatial information. While the PL deals with speech based information, the VS deals with visual information. Baddeley (1986) hypothesized the existence of a temporary visuo-spatial store which is capable of retaining and manipulating images. The VS is a cognitive construct and mental process of temporarily storing visual and spatial information for online use in operations of WM. The sketchpad also displays and manipulates visual and spatial information held in long-term memory. According to Baddeley (2002) The VS is assumed to form an interface between visual and spatial information, accessed either through the senses or from LTM and it allows a range of channels of visual information to be bound together with similar information of a motor, tactile, or haptic nature.

According to Baddeley (2003), the VS is a system that is involved in everyday reading tasks, where it may be involved in maintaining a representation of the page and its layout that will remain stable and facilitate tasks such as moving the eyes accurately from the end of one line to the beginning of the next.

2.4.3. The Episodic Buffer

Baddeley described a number of problems for the WM model which stem from the need to integrate information from the subsidiary systems and from LTM in a way that allows active maintenance and manipulation (Baddeley, 2000). After observing some patients with amnesia, who had no ability to encode new information in long-term memory, Baddeley realized that these patients had good short-term recall of stories, recalling much more information than could be held in the phonological loop. As a result, Baddeley added a fourth component to the model as the third slave system, called the Episodic Buffer (EB).

Baddeley (2000, 2003) needed the EB in order to explain aspects of memory responses that are not explained by other portions of his WM model. He has defined the EB as responsible for integrating the material in the slave systems with information in long term memory (Baddeley, 2000). In other words, the EB deals both with visual and speech-based information. The central executive controls all cognitive activities, and it also shares part of its limited capacity with the slave systems to which it can allocate more capacity.

Baddeley assumed this new component to be episodic in the sense that it holds integrated episodes or scenes and to be a buffer in providing a limited capacity interface between systems using different codes.

Baddeley (2000) assumes the executive to be a purely attentional system with a role extending beyond memory function, whereas he assumes the EB to be purely mnemonic in character. The information retrieved from the buffer is through conscious awareness. This allows multiple sources of information to be processed simultaneously, which in turn creates a model of the environment that may be manipulated to solve problems and plan future behavior. The episodic buffer is also assumed to have links to long-term memory and semantic meaning.

Although postulated by Baddeley to resolve the weaknesses of the WM model, the EB is still insufficient in explaining the problems of what exactly is being communicated between the different components. Besides, the conflation of processing and storage within the slave subsystems, and the relationship between WM and LTM remain unsolved (Andrade, 2001). The relationship between the episodic buffer and the modality specific slave systems, the relationship between the episodic buffer and long-term episodic memory, and the role of executive processes in chunking and binding of information are the issues that need further conceptual and empirical work (Baddeley, 2000).

2.4.4. The Central Executive

As the main component of WM Model, The concept of the Central Executive was first introduced by Baddeley and Hitch in 1974 as a complex system used both for the storage of information, along with the computational processing of that information. They defined the central executive as a work space which is flexible but limited in capacity. Since the CE is used for both storage and processing, greater effort is required to process information, as less capacity remains for the storage of that information. This limited capacity can be used both to regulate and coordinate the flow of information within WM besides performing processing and storage operations (Morris & Gruneberg, 1994).

As the crucial part of the WM model, the CE is the least well understood component of Baddeley's model. According to Baddeley, the reason that makes the CE component of the model complicated to understand is its supervisory functioning between the slave systems with a limited capacity.

> "The central executive component of working memory is assumed to be a limited-capacity attentional system that controls the phonological loop and sketch pad, and relates them to long-term memory. The executive is almost certainly considerably more complex than either of the two slave systems, which makes it considerably harder to investigate." (Baddeley, 1999; 62)

Researchers presume the term CE as the most complex and powerful component of the WM that controls the phonological loop and the visuo-spatial sketchpad, and relates them to long-term memory (Morris & Gruneberg, 1994). As the central mechanism of the model, it plays the role in attending to and switching attention from one cognitive process to another. The executive is almost certainly considerably more complex than either of the two slave systems, which makes it significantly harder to investigate. According to Baddeley (1996),

> "The central executive is assumed to be responsible for the attentional control of WM. The concept was initially used principally as a holding operation, allowing the study of the more tractable problems of the slave systems, while accepting the need for further investigation of the complex processes that are almost certainly involved in the control of memory (Baddeley, 1996; 13470)."

Baddeley further asserts that the CE has the most crucial role in the hierarchy of WM, which functions as the supervisory controlling system. The CE is assumed to be modality free since it acts as a link between three peripheral slave systems. The operation of the flow of information takes place between the subcomponents of either slave system or between two slaves systems and long term memory. Following the completion of the processing, it also moderates and evaluates the accuracy of the final representation of information and makes corrections if necessary (Baddeley, 1998).

Generally, the CE is equated to the Supervisory Attentional System (SAS) within the control of action model described by Norman and Shallice (1980) and by Shallice (1982). SAS model was conceptualized with a broader scope to explain the control of action when compared to the WM model (Jarrold, 1997). According to Shallice (1982), the SAS is a limited capacity system and is used for a variety of purposes, such as tasks involving planning or decision making, novel situations, dangerous or technically difficult situations or situations where strong habitual responses are involved. Baddeley adopted the SAS as the CE in his WM model because of its ability to explain the model as a processing system between its sub-components rather than a separate place of storage. As a result, the CE evolved from being a flexible system which could employ resources for storage or processing at the same time, to being entirely a processing system (Baddeley & Logie, 1999). This system is responsible for functions such as strategy switching, selective attention, retrieval from long-term memory and dual task co-ordination which represent higher level cognition tasks. Hence, the multi-component structure of the model provides a basic theoretical framework for understanding how higher-level cognition is supported by the human WM system (Baddeley, 1996).

More recently, Baddeley (2012) defines the CE as 'homunculus', a little man in the head that is capable of attentional focus, storage, and decision making, with a capacity of doing all the clever things that were outside the competence of its two subsystems. In an attempt to specify the executive functions of the CE, Baddeley postulates four suggestions. First, it would need to be able to focus attention on complex tasks. Second, desirable characteristic would be the capacity to divide attention between two important targets or stimulus streams. The third executive capacity involves switching between tasks, which refers to a specific control system. The fourth executive task that Baddeley assigns to the CE is its capacity to interface with LTM, which is connected to the assumption that the CE was a purely attentional system with no storage capacity (Baddeley & Logie 1999). However, this is a problematic issue in finding an answer to the question of how subsystems using different codes could be integrated without some form of common storage since participants combine the codes rather than simply using either one code or another.

2.4.5. Weaknesses of the WM Model

Although commonly accepted as the predominant memory model, criticisms have been raised about the WM model in terms of the processing of input and the interaction between its components. According to Andrade (2001), a well-specified model of WM should help researchers to explain the memory and executive components of the phenomenon of interest, and to answer the questions about the relationships between those components and the other cognitive processes involved.

However, as one of its limitations, the WM model is not perfectly specified, and it is not always clear which cognitive processes are not a function of WM. The model fails to accomplish its potential as a tool for making predictions and explaining phenomena because the components of the model and their interrelationships are underspecified (Andrade, 2001).

Another weakness of the model is its simplicity. Although simplicity seems to be strength of the WM model, it is also a limitation since the simple model fails to reflect the complexity of real cognition and is hard to apply to phenomena outside the domain of laboratory short-term memory tasks. In other words, the components of the model are too simple, to explain the full range of everyday phenomena and does not go deep enough to find answers to the questions of deep functioning of memory.

The third weakness of the model is related to the CE since it remains as the least understood component of the model. With a framework that describes the organization of CE in detail, developing experimental tasks which clearly tap a particular aspect of executive function, or even tasks which clearly tap the slave systems without imposing on the executive would facilitate researchers to reach more reliable findings. Besides the underspecified nature of the CE, another issue under criticism is the lack of specification of the role of the CE in rehearsal. The CE component of the WM model should be redesigned to provide an answer to the question whether rehearsal is purely a function of the slave systems, a function of the slave systems that is initiated and monitored by the CE, or a function solely of the CE.

Another component that remains unclear is the EB. Although postulated as the most recent component of the WM model, it does not provide the researchers a detailed understanding of how the episodic buffer combines information from the other parts of the model and from LTM as well.

2.4.6. Strengths of the WM Model

Unlike some other models previously discussed, perhaps, the main argument on the strength of the WM model should be its structure which elucidates not only the storage but also the processing of information.

Andrade (2001) postulates three main strengths of the WM model from an applied perspective as, *breadth, specificity*, and its *historical strength*.

Since the model encompasses both auditory and visuo-spatial processing, manipulation as well as temporary storage of the input represents the *breadth* of the model. This structure helps the researchers to analyze complex, real-life situations with various tasks. The second strength is the *specificity* of the model. By specifying separate storage and processing functions and separating verbal and visuo-spatial subsystems, the WM model helps researchers research into the specific cognitive elements of typical and atypical lifetime changes. The specificity of the WM model enables researchers to go beyond correlational studies, allowing them to gather converging evidence from experimental manipulations of specific functions such as verbal rehearsal or visual short-term memory. The third strength of the WM model may be characterized as its *historical strength* since the WM model already holds a central place in cognitive psychology, and there already exists a substantial body of data concerning the role of WM in cognitive function.

According to Morris and Gruneberg (1994), the strengths of the model are its; *flexibility, dual-task methodology, adaptability,* and *generality*. It is *flexible* because, the slave systems can handle memory tasks requiring the maintenance of restricted amounts of either verbal or visuo-spatial information, leaving the valuable limited-capacity resources of the central executive free to support a range of higher-level cognitive activities. As a further strength, *the dual-task methodology* is based on the limited capacity of the WM model which enables researchers to be able to isolate exactly which component of WM, the CE, the PL or the VS contributes to performance on particular tasks. As a third strength of the model is its *adaptability*. Since the model is conceptualized in a modular structure, it has the capacity to preserve its general structure while more local aspects of the model are undergoing revision. Finally, the model has remarkable *generality* since it has been conceptualized by drawing upon evidence from records of brain-damaged patients, children throughout all stages of development, and traditional subject group sampled in short-term memory research (Morris & Gruneberg, 1994).

2.5. Working Memory Capacity

As already been discussed earlier, within human experimental psychology, the term WM is used to refer to a limited capacity system that is capable of storing and manipulating information and that is assumed to be an integral part of the human memory system (Baddeley, 1996). Daneman and Carpenter (1980) proposed that people vary in WM capacity, and these differences are predictors of the amount of information that can be held accessible and processed (Kane & Engle, 2002; Kyllonen & Christal, 1990).

The limited capacity of memory was first suggested by Miller (1956) as the "magical number seven". According to Miller, regardless of the forms of the elements either digits, letters, words, or other units, the memory span of young adults was around seven elements, called chunks. Memory span for digits, letters or words strongly depends on the time it takes to speak the contents aloud, and on the lexical status of the contents. On the other hand, Cowan (2001) proposed that WM has a capacity of about four chunks in young adults.

According to Just and Carpenter (1992) individuals vary in the amount of activation they have available for meeting the computational and storage demands of language processing. This leads to the quantitative differences among individuals in the speed and accuracy with which they comprehend language. They describe capacity as an energy source that some people have more of than other people have. In other words, a person with a larger memory capacity for language may be able to draw on a larger supply of resources.

The WM model is based on the assumption that it is useful to hypothesize a theoretical limited capacity system that provides the temporary storage and manipulation of information that is necessary for performing a wide range of cognitive activities. In addition to that, this system is not unitary but can be divided into an executive component and at least two temporary storage systems, one concerning speech and sound while the other is visuo-spatial (Daneman, 2012).

Cowan (2005) defines WMC simply as the ability to remember things in an immediate-memory task, such as a given list of words to recall. If one recalls six of the words in a given list of words, one's capacity would be six words; if one recalls three words, one's capacity would be three words. According to Cowan, still, that is not a very satisfactory definition of capacity since there is no reason to believe that it could lead to any information beyond what we already know from the data. However, defining WMC in a narrower sense requires investigating the ability of a specific component of the system, rather than asking about the ability of the entire processing system.

According to Cowan (2001) although WM posses the ability to keep representations in an active and accessible state for intelligent behavior and is thought to underlie various diverse cognitive processes, such as language learning and problem solving, it appears to be strikingly limited, as one can only store a small amount of information in WM at any one time (Cowan, 2001). Cowan also asserts that understanding the nature of capacity limit, researchers can find out more detailed answers to the question of how our WM system is organized. Besides, understanding the nature of capacity could enable the researchers to find out whether it is mediated by subsystems that are specialized for maintaining specific types of representations (Baddeley & Logie, 1999) or whether it is limited by a single, domain-general store (Cowan, 1995). Daneman and Carpenter (1980) argued that executive processes are probably one of the principal factors determining individual differences in WMC are dependent both on storage and processing components, which are measured separately or together. Simple WMC is typically measured by the number of unrelated digits or words that can be recalled.

2.5.1. Measuring Working Memory Capacity

WM is often characterized as a system for holding a limited amount of information available for processing. The limited capacity of the WM has been measured across various contents and methods of measurement. In general, processing capacity is measured by using tasks that make simultaneous demands on storage and processing (Kane et al., 2004). WM capacity has been widely measured by WM span tasks, which were based on Baddeley and Hitch's (1974) theory of WM. These measures of WM were designed to require not only information storage and rehearsal, but also the simultaneous processing of additional information.

2.5.1.1. Simple Span Tasks and Complex Span Tasks

Basically, a distinction is often made between simple span tasks and complex span tasks. The former tasks primarily tap storage, while the latter involve both storage and processing. While simple span tasks are used to measure phonological short-term memory, complex span tasks such as counting span, operation span, reading span, speaking span and listening span tasks are designed to minimize the demand on language comprehension. Therefore, complex span tasks include not only the storage of information and rehearsal but also the immediate cognitive processing of incoming information (Daneman & Carpenter, 1980; Turner & Engle, 1989). The most commonly used paradigm for measuring WM capacity is the complex span paradigm.

Variations of complex span tests are, Reading Span (RSPAN) (Daneman & Carpenter, 1980), Listening Span (LSPAN) (Daneman & Carpenter, 1980), Counting Span (CS) (Case, Kurland, & Goldberg, 1982), Operation Span (OSPAN) (Turner &

Engle, 1989), Speaking Span (SS) (Daneman & Green, 1986), and Spatial Span (SS) (Shah & Miyake, 1996).

The general structure of all complex span tasks can be defined as encoding of a list of words or letters for serial recall together with a parallel distracting processing task such as reading a sentence, making syntactic or semantic judgments, or verifying a mathematical equation. A considerable number of studies reported that variants of complex span correlate well with each other and with other indicators of WMC (Daneman & Merikle, 1996; Kane & Engle, 2003; Unsworth & Engle, 2005). Hence, understanding the cognitive processes in the complex span paradigm would be useful in understanding the limited capacity of human cognitive processing.

Complex span tasks are usually constructed in a manner that is intended to replicate real-world memory functioning, which occurs while reading for comprehension. This may be why complex spans generally have higher correlations with academic learning and higher cognitive functions than simple spans do (Dehn, 2008). Complex span tasks are robust predictors of a wide range of complex cognitive skills, such as reading comprehension or learning electronics (Baddeley, 2003).

Although the literature presents conflicting information regarding the reliability of WM span tasks as well as problematic procedures for administration and scoring, they are still considered as current gold standard measure of WMC in the cognitive neuroscience literature (Baddeley, Hitch, & Allen, 2009).

L2 studies of WMC are usually carried out in computer labs, the classroom, or language program setting, and the participants are tested either as individuals or groups. However, group testing is often reported as problematic since the administrator can not ensure that all individuals are involved in the processing task.

The way how the memory tests are administered can affect the validity and reliability of the findings. Hence, how the memory tests are presented, the nature of the secondary processing task in complex measure, and the interpretation of the quantitative results are the significant factors in terms of validity and reliability of the tests. In complex span tests the most important point is that both the processing and storage components are engaged. Besides, the secondary processing task must interfere with rehearsal, to prevent the task from becoming simply a measure of STM.

Daneman & Carpenter (1980) have developed a test paradigm to measure individual executive WMC using WM span tests, which involve a Reading Span Test (RST), and a Listening Span Test (LST). These tests share the common property that the tasks require simultaneous processing and storage (such as listening for comprehension) and short term maintenance (such as remembering the last word in a each sentence). Unlike a single-task test such as digit-span (or word span), WM span test has provided evidence to be reliable predictor of performance on a wide variety of verbal WM measures. Therefore, a dual-task-based WM span test is found to be a sensitive measure of WMC and efficiency (Osaka & Osaka, 1984).

More recently, Redick et al. (2011) investigated the validity and reliability of automated complex span tasks. Over 6000 young adult participants performed at least one of three such tasks (Operation, Symmetry, and Reading Span). According to the authors, the standardized administration and scoring of the automated complex span task is beneficial to researchers interested in measuring WMC without creating their own task. Since automated complex span tasks are just one way to measure WMC, the use of a standardized procedure for administration and scoring would help making comparisons across studies. As a result, the authors concluded that the automated complex span task are valid and reliable tools to understand the nature of WMC, and why individual differences in WMC are related to a variety of behaviors in and out of the lab.

2.5.1.2. Reading Span Test

WM span tasks have been used widely as valid measures of WMC both in L1 and L2 research (Baddeley, Hitch, & Allen, 2009; Daneman & Carpenter, 1980; Daneman & Merikle, 1996; Harrington & Sawyer 1992; Kane et al., 2004; Miyake & Shah, 1999; Turner &Engle, 1989). A commonly used complex span test is the Reading Span Test (RST) (Daneman & Carpenter 1980). The RST assesses an individual's ability to simultaneously read and comprehend a set of sentences and then recall a target word for each, usually the last word in the sentence.

In RST, the participants are asked to read unrelated sentences, and recall the last word of each sentence. These sentences are presented individually ranging from two to six sentences sets. At the end of each set, the participants are asked to recall as many words as possible. In addition to recalling the last words of each set of words, a secondary comprehension task which involves judging the semantic or grammatical judgment of the sentence is additionally performed. This means a heavy cognitive load which requires a high level mental processing. According to Juffs and Harrington (2011), although generally considered as an accurate measure of WM capacity, RST might present various challenges, especially when undertaken outside a controlled laboratory setting.

Van den Noort et al. (2008) criticized the original RST (Daneman & Carpenter, 1980) for a number of reasons by suggesting the need for a standard computerized version of the test. They stated that, the length of the sentences must be better controlled. It would be better to control not only the number of words, but also the number of syllables for each word and letters as well. They also postulated that, unlike the original RST (Daneman & Carpenter, 1980), more attention should be paid to the length of the sentence-final words owing to the difference between remembering a one-syllable word or a four syllable word since shorter words are better recalled than longer words. Another criticism was addressed to the lack of control for the frequency of the sentence-final words since frequent words are better recalled than infrequent words. Moreover, they put forward problems with the presentation of the material. They suggested that presenting the sentences using cards to should be restricted by time. The problem with this method is that participants can read the sentences more slowly to improve their recall.

Besides, there are problems with the use of the RST in other languages. Using only the translations of the original RST in most languages, the researchers cannot control the specific language differences between the original English version and the translation of the RST in terms of word frequency, sentence length, since every language holds its own unique features.

Finally, Daneman and Carpenter (1980) established the reading span by looking at the total number of sentence-final words that participants can recall during the whole RST. Because the former procedure does not adequately reflect the participant's WM capacity (adapted from van den Noort et al. 2008).

Within the same study, they constructed the sentences for a new version of RST in Dutch, English, German, and Norwegian, according to various criteria. They controlled the length of the sentences as ranging from 12 to 17 words, 20 to 22 syllables, and 55 to 73 letters. Unlike the original RST, the number of syllables and the number of letters were controlled for over the five series. A particular lexical database (The CELEX) was used to determine the frequency of the sentence-final words for the Dutch, German, and English versions. Additionally, all the sentence-final words were concrete nouns. All sentences had 6.5 seconds of presentation time and all language versions were evaluated using the same psychometric characteristics. Also within each set, the sentences and sentence-final words were controlled for semantic relations as much as possible.

As another problematic issue, the scoring of RST might be in various ways depending on the design of the study and this variety causes difficulty in referring to the findings of other studies. Friedman and Miyake (2005) examined four scoring methods. The first one is '*total words'* in which the total number of words recalled across all trials is considered as the memory span. Second, '*proportion words'* in which the average proportional recall for each trial is taken into consideration. Third, '*correct sets words'* which means the total number of words in perfectly recalled sets. And the fourth one is '*truncated span',* which means the highest level at which the participant recalled a majority of sets (four out of six).

2.5.1.3. Listening Span Test

Listening Span Test (LST) is a spoken version of the RST, which follows the same format but requires participants to listen to a set of sentences, rather than to read (Daneman & Carpenter 1980). As a dual task test, LST includes complex span tasks for measuring both the storage and process functioning of WM. In the task, the participants hear a set of sentences and recall the last word of each sentence and in the meantime to judge whether the sentences they heard were logical or not by responding either "true" or "false". In the task, similar to RST, these sentences are presented individually ranging from two to six sentences sets. Since it captures many of the processing requirements of sentence comprehension LST is considered as one of the most important predictors of reading comprehension (Daneman & Carpenter, 1980). However, most of the criticism on LST is based on the argument whether comprehension is limited by the capacity of a general WM system or only by one that is specialized for language processes. A reasonable concern about the LST is that it depends much on comprehension. Hence, LST stands as a domain-specific WM measure in the literature.

Several studies on WM have been reported using RST or its listening equivalent LST (Daneman & Carpenter, 1980; Osaka & Osaka, 1994). Cowan (2005) reports several studies that revealed equivalence between RST and LST. Different types of WM span in which a processing task and a storage task are combined yield similar results (Daneman & Merikle,1996; Engle, Cantor, & Carullo, 1992), and correlate with one another, and with intellectual aptitude tasks, (Conway, Cowan, Bunting, Therriault & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999; Kyllonen & Christal, 1990). So, if it is a process that is critical for the correlations with intellectual abilities, it is a rather general process, though there of course are likely to be domain-specific skills involved, such as the distinction between verbal and spatial abilities (Shah & Miyake, 1996). The current study assumes that, WM capacity is a modality-free concept which has a central form of storage and capacity.

2.5.1.4. Operation Span Test

Turner and Engle (1989) developed OST for the purpose of predicting reading ability with a WM span task that did not involve the reading of sentences. As another variation of complex span tasks, OST requires the participants to solve simple mathematical equations instead of making logical or grammatical judgments about the sentences while trying to recall the final words of each set of sentence (Turner & Engle, 1989). The idea of using letters, instead of words is in order to lessen the dependence on prior language knowledge (Conway et al. 2005). In other words, while comprehending sentences in RST and LST requires knowledge of the language, OST has been developed as a measure more of the domain-general capacity rather than the content-specific capacity.

According to Turner and Engle (1989), the OST is a measure of WMC that strongly implicates the operations of the CE. The OST, like other complex span tasks, involve participants to continuously shift back and forth between the processing and storage requirements of the task. While the participants try to verify equations they also try to remember the target words of each task which represent the activation of CE during the operations.

The next section will discuss the concept of WM in terms of L1 and L2 reading research and L1 and L2 listening research. Besides, examples of empirical studies related to the relationship between WM and reading-listening both in L1 and L2 will be presented within this section.

2.6. The Relationship of Working Memory with Reading and Listening Comprehension and Language Proficiency

Over the past decades, identifying comprehension skills have become one of the major research interests. In the field of cognitive psychology, researchers have conducted studies to understand the basic processes both in reading and listening comprehension. A large body of the research has focused on the analyses of how readers and listeners represent text in memory. However, the research has paid less attention to the lower level processes that take place during language comprehension, either written or spoken.

Since reading involves complex and difficult processes, it becomes a critical issue as a basic life skill and an important means of learning (Harrington & Sawyer, 1992). In order to understand the complexities involved, it helps to consider reading as a form of discourse processing that involves multiple levels of knowledge representation (Kintsch, 1998). During reading, reader must first decode the letters and words in a given text and then try to get a meaning from the string of words along with an attempt to understand the grammatical relationships between all these components.

The close interaction between WM and reading is significant since insufficient WMC can cause problems in reading or reading comprehension. WM enables the coding, processing and recording of information during the reading processes. Research has revealed that reading comprehension relies on the capacity of WM to retain text information that facilitates the comprehension of succeeding sentences (Just & Carpenter, 1992). Hence, individuals with greater WM capacities are more successful not only at integrating information during reading, but also carrying the information from one sentence to the next. The fundamental nature of comprehension is the integration of information that is derived from any text. Therefore, in this sense, integration can be accomplished by an effective functioning WM (Cain, 2006).

In general means, although the reading process has been described as the process of receiving and interpreting information encoded in language form via the medium of print (Urquhart and Weir, 1998), this definition does not indicate other components of the linguistic processing such as text structure, length, lexical and linguistic complexity. Therefore, a definition of reading should include information of how readers engage in phonological, morphological, syntactic, semantic, and discourse processing (Leeser, 2007).

Once considered as a simple decoding process, today it is well established that reading is an interactive, constructive and contextualized process through which individuals make meaning due to the interaction between the reader and the text that the reader creates meaning from by activating a number of cognitive linguistic processes (Grabe, 1988).

According to Grabe (2009) for reading, the role of WM in lower-level processing is well-established. According to him;

> "Working memory supports phonological, orthographic, and morphological processing for word recognition. It stores and combines words that have been activated, it carries out syntactic and semantic processing at the clause level, and it stores the relevant information for building text comprehension." (Grabe, 2009; 35)

Even though WM capacity have been correlated with reading comprehension (Harrington and Sawyer, 1992; Cain, 2006), syntactic processing, (Miyake and Friedman, 1998), and vocabulary learning (Atkins and Baddeley, 1998), there is still need for research on the role of lower-level processes during the process of reading. Starting from phonemic level, *phonological*, *orthographic*, *morphological*, *syntactic*, and *semantic processes* build up text comprehension. However, researchers typically define reading comprehension itself as a global construct and pay little attention to its multilevel representational structure and the role played by each level in comprehension (Alptekin & Erçetin, 2009).

During reading, the reader has to convert written words into phonemes. This phonological processing depends on visual processing of any given text. By accessing learned phonetic codes, the readers match the graphemes with the phonemes they represent to recode visual stimuli. This corresponding process occupies a large part of the WM capacity of the reader since it requires segmenting, blending, and holding of phonemes. Hence, there remains less space for semantic and syntactic procedures in WM capacity of the reader. However, as decoding becomes automated, more WM capacity becomes available for reading comprehension. As a result, automaticity appears as a factor that facilitates reading comprehension by reducing the WM resources necessary for decoding words (Dehn, 2008). On the other hand, poor readers who focus on reading decoding have less WM resource for comprehension because of their inefficient word processing.

As a result, the link from word-level reading to comprehension was through the assumption that comprehension included both lower- and higher-level processes that require cognitive resources that rely on the WM capacity Perfetti (2007:358).

On the other hand, cognitive processes that take part in listening comprehension is more complicated when compared to reading. In this respect, listening is a distinct skill from reading since it involves real-time processing, without the option of going back to earlier sections of the text that have been missed by the listener (Flowerdew, 1994). Due to a number of phonological and lexical features such as, pronunciation differences and reduced forms of language, listening comprehension becomes a more challenging task for language learners.

Listening can be defined as the ability to process spoken language automatically, to understand the linguistic information that is attached to the oral text. Listening comprehension on the other hand is more than just hearing the spoken language. The listener often has to construct the meaning using linguistic knowledge, the context of the situation, and his/her background knowledge on the topic of the spoken text (Brown & Yule, 2001; Flowerdew & Miller, 2005). Hence to decode and comprehend what is spoken requires complex cognitive processes that take place in a limited time.

During listening, WM plays the vital role of constructing and integrating ideas from a flow of following words (Just & Carpenter, 1992). To be able to understand the meaning of an utterance, the listener has to remember the earlier parts of the utterance in order to connect them to later occurring words. During this multifaceted process, WM stores the incomplete results of comprehension, as well as encoding some items for later retrieval.

Anderson (1995) proposed a cognitive framework for listening comprehension including three phases as, *perception*, *parsing* and *utilisation*. These three phases represent different levels of processing, and are interrelated and recursive and can take place at the same time during a single listening event.

Within the framework of the model, perceptual processing is the encoding of the acoustic or written message which involves segmenting phonemes from the continuous speech stream. In the parsing phase, words are transformed into a mental representation of the combined meaning of these words according to syntactic structures or cues to meaning. During utilisation, the listener may draw different types of inferences to complete the interpretation and make it more personally and meaningful to respond to the speaker of the utterance (Anderson, 1995, p. 37). According to this model, the listeners have to rely on their WM capacities in order to store and process the input simultaneously. As a result of this cognitive load, L2 listeners experience greater demands on their WM processing when being exposed to a spoken text in the L2. Thus, an oral utterance containing more information poses a greater challenge for the WM capacity of the listener (Baddeley & Hitch, 1974).

All in all, the discussions on the role of proficiency in reading and listening comprehension within the framework of WM, with respect to a large body of research, reveal that learners with greater WM capacities are more successful at interpreting information both in oral and written texts. Hence, less proficient language learners spend much of their WM capacities on holding information in their memories while processing listening comprehension tasks. Conversely, more proficient learners can process the spoken information more easily and share more cognitive resources to operationalize the information representations. As a result, it becomes evident that the role of WM capacity in L2 listening comprehension is mediated by L2 proficiency.

2.7. Working Memory and L1 Reading Research

As one of the four language skills, reading is often defined as a complex process involving word recognition, comprehension, fluency, and motivation, along with a variety of cognitive processes. However, researchers typically define reading comprehension itself as a global construct and pay little attention to its multilevel representational structure and the role played by each level in comprehension (Alptekin & Erçetin, 2009).

Since reading involves complex and difficult processes, it becomes a critical issue as a basic life skill and an important means of learning (Harrington & Sawyer, 1992). In order to understand the complexities involved, it helps to consider reading as a form of discourse processing that involves multiple levels of knowledge representation (Kintsch, 1998). During reading, reader must first decode the letters and words in a given text and then try to get a meaning from the string of words along with an attempt to understand the grammatical relationships between all these components.

There is much evidence for WMC as a strong predictor of reading comprehension both in L1 (Daneman & Carpenter, 1980; Miyake & Friedman, 1998) and in L2 (Harrington & Sawyer, 1992). Reading comprehension involves drawing inferences, which in turn means to comprehend a text (Calvo, 2001; Alptekin & Erçetin, 2010). However, whether inferential or literal, working-memory capacity is differentially involved in the comprehension process depending on the reader's interaction with a given reading task (Alptekin & Erçetin, 2009).

There have been attempts in discovering the role of various span tests and their relationship with reading comprehension in L1 during early 1990s. Turner & Engle (1989) tested the relationship between simple and complex memory span tests and reading comprehension. They presented the participants lists of words or digits to remember along with the RST. The analysis of the data showed that, complex memory span tests are significantly correlated with reading comprehension. They also reported complex span as a more accurate measure of WMC since it prevents the use of rehearsal whereas simple memory does not.

La Pointe and Engle (1990) carried out a further study on the relationship between simple and complex spans and reading comprehension. The aim of the study was to follow the finding of Daneman and Carpenter (1980) and Turner and Engle (1989) that the simple word span did not predict comprehension. They designed five experiments to investigate the effects of word length in simple word span tasks and complex operation and reading span tasks and the relationship between these tasks and reading comprehension. They utilized a computerized version of RST, a word span test, OST, and a reading comprehension test. Experiments 1 and 2 showed that word length effects both simple and complex memory span tasks correlated with reading comprehension. The third experiment revealed that, articulatory suppression did not eliminate word length effects. And the final experiments showed that articulatory suppression eliminated the effect of word length when words were sampled with replacement from small fixed pools but not when sampled without replacement from a large pool. The authors concluded that the reading span does not measure WM specific to reading.

Engle, Nations and Cantor (1990) tried to find out whether the correlation between complex WM spans and reading comprehension occurs because the complex spans reflect the capacity of a structural WM that plays a causal role in comprehension or because another factor, word knowledge, plays a causal role in both the span tasks and comprehension. Ninety college students were given both simple and complex versions of the word span task with high- and low-frequency words. They measured the reading comprehension of their participants with the Verbal Scholastic Aptitude Test (VSAT). Their findings revealed that the correlation between span and VSAT was somewhat higher when span was tested with low-frequency words, but was significant with both low- and high-frequency words. According to the authors, this suggests that both word knowledge and a content-free WM play a fundamental role in the relationship between word span and higher level cognitive tasks.

Engle, Cantor and Carullo (1992) carried out a study related to individual differences in WM and reading comprehension. They tested four hypotheses in three experiments. In the first two experiments, a moving window procedure was used to present the operation–word and reading span tasks. The correlation between span and comprehension was reduced when the viewing times were partialed out. Exp 3 compared a traditional experimenter-paced simple word-span and an S-paced span in their relationship with comprehension. The experimenter-paced word-span correlated with comprehension, but the S-paced span did not. The results of the analysis supported a general capacity explanation for the relationship between WM and comprehension.

Similarly, a PhD dissertation by Convirs (1994) investigated the role of WM in reading comprehension by using simple and complex span tests. The aim of the dissertation was to have a better understanding of how each aspect of WM was related to reading comprehension of children. Several aspects of WM system was analyzed by a non-verbal test of intelligence in relationship to reading ability of children of different ages. Convirs hypothesized that if intelligence accounts for most of the variance in scores of reading comprehension, then WM might not have been a specific factor independent from general intelligence. Conversely, if general ability does not account for most variance in reading scores, then WM might have played a separate role in enabling the reading process. Finally, the analysis of the data revealed that participants with different levels of reading ability also vary in WM ability. The study concluded that good readers have more efficient WM abilities when compared to poor readers which in turn show the linear relationship between WMC and reading ability. .

Ikeda and Kitagami (2012) investigated the effect of WMC and mental effort on the ability to monitor the accuracy of text comprehension. WMC of the participants were measured by the operation span test and were given five scientific expository reading texts. After reading, participants assigned comprehension rating to each text and then completed a comprehension test. The analysis of the data revealed a significant interaction between WMC and mental effort. They concluded that both WMC and mental effort affected the ability of participants to monitor their accuracy.

2.8. Working Memory and L2 Reading Research

The role of WM in L2 reading has also been investigated in various empirical studies by using Daneman and Carpenter's (1980) original measure of WM, with span tasks such as RST, Counting span, and OST (Chun & Payne, 2004; Harrington & Sawyer, 1992; Leeser, 2007; Walter, 2004). In their series of experiments, Daneman and Carpenter (1980) devised an RST that predicted reading comprehension better than a simple digit span task did. As a result, they postulated that reading ability was positively correlated with WMC relating to both processing and storage capacity.

Harrington & Sawyer (1992) examined the role of L2 WM capacity on reading skills of advanced L2 learners. WM capacity was measured by the RST (Daneman & Carpenter, 1980). They found that, participants with larger WMC scored higher on measures of reading skill. They also found that FL readers with higher WM spans performed better on both the reading and grammar/vocabulary sections of the Test of English as a Foreign Language (TOEFL) proficiency exam. They concluded that their findings are consistent with an interpretation of the RST as an index of WMC, in which capacity is defined functionally in terms of a trade-off between active processing and storage.

Osaka and Osaka (1992) investigated the relationship between WMC in the L1 and L2 to determine whether WMC was language dependent. L1 Japanese/L2 English participants were administered Daneman and Carpenter's (1980) RST along with two other versions of the RST: a Japanese version and an L2 English (ESL) version. Significant positive correlations were found between the scores on the Japanese and ESL versions of the RST, as well as between Daneman and Carpenter's original RST and the Japanese version.

In a follow-up study, Osaka, Osaka, and Groner (1993) measured the efficiency of WM with German and French versions of the RST. Fifteen Swiss German/French bilinguals participated in the study. They reported a significant correlation between German and French versions of the RST for L1 German/L2 French participants; they also reported significant differences between the two tests (performance on L1 German and L2 French). The results of the correlational analyses indicated that the efficiency of WM for reading is independent of language, reconfirming the results obtained by M. Osaka and N. Osaka (1992)

Walter (2004) investigated the transfer of reading comprehension skills to L2. The study hypothesized that transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 WM. She based the theoretical framework of the research on two notions from cognitive psychology in relation to the transfer of reading comprehension skills from L1 to L2 (i.e. the notion that reading comprehension proceeds by the reader's building of a mental structure representing the text and the notion of WM). The participants were French learners of English from two groups of proficiency level. The members of both groups were proficient readers in L1 French, but they differed in their ability to comprehend texts in L2 English, even when the lower-intermediate learners had no problem in processing the individual sentences of those texts. The findings provided strong support for the hypothesis that the lower-intermediate group had failed to transfer to L2 the ability to build well-structured mental representations of texts, while the upper-intermediate group had succeeded in transferring this ability. This structure-building ability was in turn linked to the development of WM in L2.

Another study by Chun and Payne (2004) investigated the role that individual differences in 13 students in a second-year German language course who read a German short story in L2 on a multimedia CD-ROM with access to multimedia annotations of difficult vocabulary words, and their actions while completing a vocabulary test and comprehension exercises. A non-word repetition test and a version of Daneman and Carpenter's (1980) RST were used to measure WM. They found a relationship between phonological WM (non-word repetition test) and the number of times the learners consulted vocabulary glosses (i.e., look-ups).

Van den Noort, Bosch, and Hugdahl (2006) researched the hypothesis that WMC interacts with FL proficiency. They tried to find an answer to the question of whether the interaction between WMC and language proficiency is language-specific. The multilingual participants were native (L1) Dutch speakers who were fluent in their second (L2) language, German, and had recently started learning Norwegian as their third (L3) language. Hence, they aimed to control the interaction between WMC and language proficiency with a multilingual design. As for instrumentation they used a forward digit span as simple WM task and the RST and letter-number ordering as complex WM tasks. The results revealed that differences in performance between L1, L2, and L3 can be found on both simple and complex WM tasks. The findings suggested that WM interacts with FL proficiency when moving from the participants' L1 to their L2 and L3. The reason is that the shift between the languages reduces WM resources. Besides, the authors also reported significant correlations between the RST scores of their participants in their three languages. This finding supports the WMC interaction hypothesis which suggests that WMC is constant. This means that the relationship between the WMC and FL proficiency is language-independent rather than language specific.

Leeser (2007) examined how topic familiarity and WM capacity affect beginning L2 Spanish learners' reading comprehension. Participants consisted of 146 high-beginner L2 Spanish learners enrolled in an accelerated, elementary Spanish course. A computerized version of a RST was used as a measure of WM capacity, and all participants read texts about familiar or unfamiliar topics and completed recall protocols to measure passage comprehension. Results revealed that WM's effect depended on the participants' previous knowledge about text topics. In other words, learners with high prior knowledge and high WM capacity showed the best performance.

Van den Noort, Bosch, Haverkort, & Hugdahl (2008) carried out a study by making use of a standard computerized version of the RST in four different languages (i.e. Dutch, English, German, and Norwegian). Their new version of RST exposed the same methodological criteria for all four languages. They conducted a plausibility test, an abstract-concrete rating scale, and a pilot-study on native speakers to test the new RST. Besides, they tested the internal and external reliability and the ecological validity of the new the new RST. The results indicated the new RST as a suitable test to investigate verbal WM. Finally, an important advantage of the new RST is that the different language versions make cross-linguistic comparisons of RST results possible.

Sunderman and Kroll (2009) investigated the WMC of university students who studied abroad and those who did not. They aimed at demonstrating the important role of WM in comprehension and reading abilities, as well as language skills more generally. They examined the role of WM resources in lexical comprehension and production for learners who had or had not studied abroad. Participants included native English learners of Spanish. Of the 48 participants, 34 had not studied abroad, but 14 had. The span task was adapted from a reading span measure designed by Waters and Caplan (1996). The comprehension measure consisted of a translation recognition task, in which the participants were asked to decide if two words were translation equivalents of each other. The authors found that individuals who lack a certain threshold of resources are unable to benefit from that experience. They concluded that WM is an important internal resource, necessary for the individual to benefit from studying abroad. Those individuals with higher WM may be able to attend to more linguistic factors at once, thereby increasing their ability to parse grammatical structures.

Another study on WMC and L2 university students' comprehension by Fontanini and Tomitch (2009) specifically focused on comprehension of linear texts and hypertexts. The study aimed to investigate the relationship between WM capacity and L2 reading comprehension of both linear texts and hypertexts and utilized three instruments. To measure comprehension they used recall and comprehension questions. They measured the WMC by the Reading Span Test (Daneman & Carpenter, 1980). The participants were forty-two speakers of English as an L2 from two different L1 backgrounds (i.e. Brazilians and Chinese). The results revealed that hypertexts might compromise comprehension, especially for participants with lower WMC. The conclusion of the study was that different variables including readers' WMC, their L1, and the mode of text presentation may interfere in L2 reading, and each one of these aspects might impede, in different ways, the construction of a coherent mental representation of the text being read.

Alptekin and Erçetin (2011) investigated the effects of WM capacity and content familiarity on literal and inferential comprehension in second language reading. 62 students from advanced level English courses participated in the study. They measured WMC via a computerized RST. Experimental and control groups were given either the modified text or the original text and then the participants were given a multiple-choice test that included both literal and inferential comprehension. Findings revealed significant effects of WMC and content familiarity on inferential comprehension, but not with literal comprehension.

Oh (2011) carried out a study on the construct of WM and its role in LI and L2 reading comprehension. The study tried to explain a construct of WM as a multi-component model by reviewing empirical studies on the impact of WM in LI and L2 reading comprehension. The author introduced a newly adopted construct to the model of WM, EB and discussed the role of background knowledge in relation to WM. The review indicated the CE as a significant predictor both for LI and L2 reading comprehension. Conversely, although the PL is significantly related to CE, it was not found as a direct significant predictor for LI and L2 reading comprehension but, it explains significant variances of vocabulary acquisition in early language acquisition, which in turn is a direct significant predictor for reading comprehension.

Juffs and Harrington (2011) investigated the aspects of WM in L2 learning by an article reviews research. They reviewed recent developments in the WM model and issues surrounding the operationalization of the construct itself. They further discussed various methods of measuring WMC some of which are word and digit span tasks, reading, listening and speaking span tasks. They also outlined the role proposed for WM in explaining individual differences in L2 learning processes and outcomes, including sentence processing, reading, speaking, lexical development and general proficiency. They concluded the review study by indicating WM as a non-unitary construct which depends on the age of the L2 learners, the task and the linguistic domain.

Rai, Loschky, Harris, Peck and Cook (2011) investigated the effects of stress, WMC, and inferential complexity on Spanish FL readers' inferential processing during comprehension. They predicted that stress would disrupt FL reading comprehension, with stronger effects on higher level inferences involving the situation model level of representation than on lower level inferences such as bridging inferences, or non-inferences, such as memory for facts. Another prediction was that such stress effects would be greater for those with higher WM resources but primarily affecting their reaction times. Fifty five intermediate-level Spanish FL learners' participated in the study. Reading comprehension of the participants was measured by using questions with three levels of inferential complexity: non-inference (factual), bridging inference (pronoun referent), and pragmatic inference. They measured participants' WM capacity with a computerized version of the Operation Span Task developed by Unsworth et al. (2005) and the participants' stress level using a video camera. The analysis of the data revealed that learners with higher WMC were more accurate overall. Inference construction during comprehension was negatively related to inferential complexity. The results also showed that participants with higher WMC strategically traded reading speed for greater comprehension accuracy, whereas lower WM learners only did so under stress and did so less successfully. They concluded that stress impedes FL reading comprehension through interactions between WMC and inferential complexity.

Pae and Sevcik (2011) investigated the role of verbal WM in L2 reading fluency and comprehension. The study provided a comparison of fifty English-

Korean and Korean-English bilinguals' reading fluency and comprehension both in their L1 and L2. Each participant was individually given measures of forward and backward digit recall, sentence recall, and reading skills. The results showed that participants' L1 and L2 performance was consistent with the findings of previous research, indicating that the students demonstrated a greater strength in L1 than L2. Besides, L1 forward and backward digit spans accounted for the significant variances in L2 reading fluency and comprehension for the English-speaking participants in the U.S., whereas L1 forward digit span was more predictive of L2 reading fluency and comprehension than backward digit span and sentence recall for the Korean-speaking counterparts in Korea.

Kaushanskaya & Yoo (2012) examined bilinguals' phonological short-term and WM performance in their L1 and L2. Twenty Korean–English adult bilinguals participated in the study, including undergraduate and graduate students at a university. Short-term memory was measured by a non-word repetition task, and WM was measured by a complex task. The results of the study showed a better performance of bilinguals on the STM task than on the WM task, and with shorter non-words than with longer non-words. The authors concluded that L1 STM performance was superior to L2 STM performance, but only for the longest non-words, whereas L1 WM performance was superior to L2 WM performance across all length levels. In addition, correlation analyses between bilinguals' L1 and L2 performance revealed stronger cross-linguistic associations for the WM task than for the STM task. In the light of these findings, they suggest that WM tasks may engage domain-general CE processes in bilinguals, whereas STM skills may depend on language-specific knowledge in the L1 and the L2.

Roberts (2012) investigated individual differences in L2 sentence processing. The paper focuses on the effect of individual factors such as WMC and proficiency on L2 sentence processing. The results showed that learners of higher WM capacity or proficiency appear to process the input differently from those of lower memory capacity and/or proficiency. In other words, L2 learners with a higher WMC or greater proficiency are more likely to process the input like native speakers. Otherwise, learners appear to shallow process the input, irrespective of individual variability.

2.9. Working Memory and L1 Listening Research

As already been discussed, the term WM refers to a cognitive system that contains a limited computational space in which materials can be temporarily stored, monitored and manipulated (Baddeley, 1996). Caplan and Waters (1999) argued that this limited capacity WM system restrains the time and linguistic operations one operationalize during comprehending a complex sentence. Hence, the close interaction between WM and listening is significant since insufficient WMC can cause problems in listening or listening comprehension. Although there are clear academic opinions suggesting that WM is critical in listening comprehension (Engle, 2002), a review of the literature, nevertheless, reveals the fact that few studies exist on the relationship between WM and listening comprehension in the L1.

An early unpublished doctoral dissertation presented by Futransky (1992) aimed to explore the role of verbal WM in listening comprehension for fifth-grade students in L1. Listening comprehension of the participants was tested by Profiles in Listening and Reading test (PILAR). To address both storage and processing components of verbal WM, a variation of sentence span memory test was used

(Daneman and Carpenter, 1980). Results showed that memory demands affected listening comprehension accuracy which means that having limited memory capacity is associated with reduced listening comprehension.

Kosaka and Yamazaki (2000) examined the relationship in pre-schoolers' WM capacity and text comprehension. They grouped forty one children in a low or high WM group according to the results of a Listening Span Test. Participants' text comprehension was measured by Sentence Verification Technique tests. The results revealed that although children with low WM constructed inferential representations during listening, they failed to grasp the gist of the text. The concluded that global integration may be followed by grasping the gist of text and that WMC may strongly affect the comprehension of higher levels of text.

Florit, Roch, Altoe and Levorato (2009) carried out a study to find out the role of memory in listening comprehension in preschoolers. The participants were administered measures to evaluate listening comprehension ability (story comprehension), short-term and WM skills (forward and backward word span), verbal intelligence and receptive vocabulary. Results showed that both short-term and WM predicted unique and independent variance in listening comprehension after controlling for verbal abilities, with WM explaining additional variance over and above short-term memory. Their results supported the conclusion that a direct relation exists between memory and listening comprehension. They indicated that WM accounts for a large amount of variance in listening comprehension which means that not only the maintenance but also the processing of information in memory provides a unique and independent contribution to listening comprehension.

2.10. Working Memory and L2 Listening Research

An important part of communication for individuals, listening has an essential role in language learning, and research has revealed that listening ability is the major factor between more and less successful learners in their ability to use listening as a means of language learning. Although it is commonly accepted that L2 learners need a large amount of comprehensible input to help them become better language learners, in an FL environment, L2 learners are usually confronted with a number of difficulties, such as fast speech or unfamiliar vocabulary (Brown, 2001).

However, the role of WM in the listening process has not been well explored yet. One of the reasons might be the little importance that teachers attach to listening courses. As a result of this, in contrast to other language skills, teachers tend to provide less time and effort for listening activities and tasks. Yet, there are a number of studies that aimed at exploring the relationship between WM and listening process in L2.

Tsuchihira (2007) assessed the listening span of junior college students in their freshman year both in L1 (Japanese) and in the L2 (English), and evaluated the participants' listening ability. The participants were at the age of 18-20, and their proficiency was at beginner level. The study used an L2 listening memory span test designed in both languages. As for the listening measure, listening questions were taken from the listening sections of The Society for Testing English Proficiency STEP (Eiken) for 2nd and pre-2nd grade. The results revealed a significant relationship between L1 and L2 WMC. The results also showed that both L1 and L2 WMC were related to L2 listening ability.

Shanshan and Tongshun (2007) investigated the effect of WM on listening process and its relationship with listening comprehension. Fifty-nine freshmen students from a Chinese University participated in the study. The participants' WM capacity was measured by a modified listening span test developed by Daneman and Carpenter (1980). The participants' L1 (Chinese) and L2 (English) were both measured by a WMC test in order to examine which one is more efficient in studying the relationship between WM and the foreign language listening comprehension. The study also used the listening part of CET-4 (College English Test Band four) as a measure of listening comprehension. The results demonstrated that WM was an effective predictor of EFL listening comprehension. Learners with larger WMC were more likely to have better abilities in listening comprehension.

An unpublished doctoral dissertation by Londe (2008) aimed to investigate the role of WM in L2 listening comprehension tests. Participants consisted of Hungarian college students between ages 19-25 who speak English as a second language at various levels of proficiency. Participants' WM capacities were tested by Hungarian version of LST both in L1 and L2. L2 listening comprehension performance was assessed by a standard English as a second language test which included an authentic English language lecture recorded at university. The results of the study indicated WM as essential in the listening comprehension construct.

Kormos and Safar (2008) conducted a study to find out whether WMC was related to performance of 121 secondary school students aged 15-16 in the first intensive language training year of a bilingual education program in Hungary. The participants' general proficiency including listening was tested by Cambridge First Certificate Exam. Fifty students were also tested with a backward digit span test, measuring their WMC. They indicated that the backward digit span test correlated very highly with the overall English language competence, including listening comprehension test scores.

Mackey, Adams, Stafford & Winke (2010) examined the relationship between learners' production of modified output and their WMC. Forty two English-speaking learners of Spanish participated in the study. Participants were given a listening span task as a measure of WMC which was modeled on a standard Reading Span measure (Waters & Caplan 1996b; Waters & Caplan 2004). They found a relationship between the learners' WMC and their tendency to modify output. They reported that the larger processing capacity cause greater production of modified output during interaction. The authors concluded that WM accounted for 17–18% of the variance in the scores, indicating that other factors were also playing a role in the amount of modified output.

Juffs and Harrington's (2011) recent review of research on the role of WM in L2 learning presents the recent developments in the issues surrounding the operationalization of the construct of WM and discusses various methods of measuring WM including word and digit span tasks, reading, listening, and speaking span tasks. They report that the findings from the Listening Span Tests do indicate a relationship between WM and listening comprehension. But they also point to the fact that a great deal of research is necessary to confirm the findings of existing research.

Andringa, Olsthoorn, van Beuningen, Schoonen and Hulstijn (2012) investigated individual differences in both native and non-native listening comprehension. The authors aimed to find determinants of success in native and non-native listening comprehension from an individual differences approach. They tested 121 native and 113 non-native speakers of Dutch on various linguistic and nonlinguistic cognitive skills that underlie listening comprehension. They used structural equation modeling to discover the predictors of individual differences in listening comprehension. They found that listening comprehension for native speakers was a function of knowledge of the language and the efficiency with which one can process linguistic information. On the other hand listening comprehension for non-native speakers was found as a function of knowledge and reasoning ability. Nevertheless, WM did not explain unique variance in listening comprehension in either group.

Marx and Roick (2012) investigated determinants of listening comprehension in L1 and L2 students. Stating the need for studies comparing listening comprehension and its determinants in L1 and L2 students, they based the study on the relationship between students' reading comprehension and listening skills. 424 ninth-grade students participated in the study. They measured listening comprehension, phonological awareness, WM, morphosyntactic skills, and vocabulary of their participants. The results indicated that L2 students' vocabulary was less developed than L1 students' vocabulary. However, only Turkish-speaking L2 students showed weaker listening comprehension skills than L1 students. They concluded that listening comprehension was affected by vocabulary and morphosyntactic knowledge.

Moussa-Inaty, Ayres and Sweller (2012) investigated the result of simultaneously reading and listening to the same materials when learning English as a foreign language from a cognitive load theory framework. They carried out three experiments in which the participants were exposed to varieties of written and oral presentations of materials. Following that, the students were given reading, writing, and listening tests. The findings indicated that participants exposed to reading alone performed better on listening tests than participants exposed to both reading and listening simultaneously. Within a cognitive load theory framework, the findings suggest that at least some categories of learners will enhance their listening skills more by reading the materials only rather than simultaneously reading and listening. In addition to that, simultaneous presentation of spoken and written text should be avoided, when learning to listen is the aim, considerably more emphasis should be placed on reading.

Martin and Ellis (2012) investigated the roles of phonological short-term memory and WM in L2 grammar and vocabulary learning in an artificial foreign language. Fifty monolingual native English speakers participated in the study. As for instrumentation, they used non-word repetition, non-word recognition, and listening span tests as measures of memory. Participants were tested on their ability to induce the grammatical forms and to generalize the forms to novel utterances. Results indicated that depending on the measure, individual differences in final abilities in vocabulary and grammar correlated between 0.44 and 0.76. However, the results demonstrated significant independent effects of phonological STM and WM on L2 vocabulary and L2 grammar learning.

In sum, a review of the studies on the relationship between WM and L2 listening comprehension revealed that WM is likely to affect L2 listening comprehension, and that these effects are predominantly strong in situations that impose additional WM capacity load.

2.10. Chapter Summary

This chapter presented an overview of the theoretical background related to WM theories and text comprehension as well as a review of major findings from studies conducted on the relationship between WM and text comprehension both in L1 and L2.

After discussing the concept of memory and its importance in language learning, models of WM were overviewed. Atkinson-Shiffrin's Multi Store Model of Memory (1968), Craig and Lockhart's The Levels of Processing Model of Memory (1972), and Baddeley and Hitch's The Working Memory Model (1974) were presented in a chronological fashion.

Since the theoretical framework for the present study is based upon Baddeley's model of WM (2000), each component of the WM was discussed in detail in relation with the strengths and weaknesses of the model.

Then, the chapter focused on the concept of WMC along with the discussion of various instruments used in measuring WMC.

Finally, this chapter overviewed the research on the relationship between WM and L1-L2 Reading Research and WM and L1-L2 Listening Research.

CHAPTER 3

METHODOLOGY

3.1. Introduction

This chapter covers the methodological procedures implemented in this study. First, the chapter will present the details of the participants and a brief report on the procedures followed during the pilot study. Next, the instruments used in the data collection procedure will be discussed in detail. Finally, procedures for data analysis will be provided.

3.2. Participants

This study has a pre-experimental design with the pretest-posttest nonequivalent groups, composed of intact classes. The participants consist of Turkish University students from Çanakkale Onsekiz Mart University who received course credit for taking part in the study. All of the participants were native speakers of Turkish. Their age ranged from 18 to 22. They form two groups at different proficiency levels: a) the prep students, b) the junior students from the departments of English Language Teaching (ELT) and English Language and Literature (ELL).

According to the regulations of the university, at the beginning of the term, all incoming students from ELT and ELL departments are given a placement test including grammar, reading, essay writing, listening, and speaking parts. Participants who pass the placement exam are enrolled as freshmen. Students who get scores below 60 points have to attend English preparatory classes for one academic year. The students in the preparatory school attend 8 hours of Basic English grammar, 7 hours of listening and speaking, 7 hours of reading, and 4 hours of writing courses per week. These learners are considered to be intermediate learners of English. According to the Common European Framework of Reference for Languages

(CEFR) the participants are considered at B1 level which means that they can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, can deal with most situations likely to arise whilst travelling in an area where the language is spoken, can produce simple connected text on topics which are familiar or of personal interest, can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans (CEFR, 2001).

The second group of participants is composed of students who pass the placement exam and enroll as freshmen. They are considered upper-intermediate level English learners and start taking courses from their programs. According to the CEFR, the participants of this group can be considered as B2 level which means that they can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in their field of specialization, can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party, and can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options (CEFR, 2001).

The participants were randomly selected from both groups among a target population size of 130 prep and 135 first year students. The students who had participated in the pilot study were excluded for the main study. Table 1 shows the distribution of the participants according to gender and class year.

Table 1.

Distribution of the participants according to gender and year of education.

	Female	Male	Total
Prep students	55	31	86
1st year students	56	31	87
Total	111	62	173

3.3. Instrumentation

3.3.1. Measures of Reading and Listening Comprehension

To measure the participants' reading and listening comprehension, this study made use of the Certificate in Advanced English (CAE) by Cambridge University Press. CAE is the advanced general English examination provided by University of Cambridge ESOL Examinations in England. In the CEFR, CAE is ranked at C1 and C2 levels. Candidates who have obtained an A grade are awarded a C2 certificate, those obtaining grade B or C, are awarded a certificate at C1. CAE is a required qualification for international students applying to many British universities.

CAE comprises reading, writing, use of English, listening, and speaking parts. However, since the aim of the present study is to find out the relationship between WM capacity and reading and listening comprehension in the L2, the participants were only given reading and listening sections of the CAE. For the reading section, the testees are supposed to understand texts from publications such as fiction and non-fiction books, journals, newspapers and magazines. On the other hand for the listening section, the participants need to show that they can understand the meaning of a range of spoken material, including lectures, radio broadcasts, speeches and talks.

Both the reading and listening exams were taken from Exam Essentials: CAE Practice Tests by Charles Osborne (2006). With reference to the standard application of CAE, the participants were given 75 minutes for reading section, and 40 minutes for listening section both for pre-test and post-test.

The reading section of the CAE consisted 4 sections. In section A there were 2 multiple choice reading comprehension questions for each of three short passages. Section B comprised 7 multiple choice reading comprehension questions for a longer passage. Section C consisted of a magazine article in which six paragraphs have been removed. The participants had to choose the right paragraph which fits each gap. In Section D, there were 15 questions and the participants had to choose which question was related to any of the given category. The total number of questions was 34. Each correct answer was awarded 1 point.

The listening test of the CAE included 4 sections. In Section A, the participants listened to three different extracts, and they had to choose the answer (A, B, or C) which fitted best according to what they heard. There were two multiple-choice questions for each extract. In Section B the participants were asked to listen to part of a talk and complete the sentences with the appropriate word. There were 8 fill in the blanks questions. In Section C, the participants listened to an interview and chose the best answer for 6 multiple-choice listening comprehension questions. In Section D, the participants had to accomplish two tasks. For each part, they listened to five short extracts and were asked to choose from the list A-H what each speaker talks on a certain topic. There were 5 questions for each task. There were 30 questions for the listening part. Each correct answer was awarded 1 point.

Cronbach's alpha as an indicator of the internal consistency reliability of the pre-test was .64 for the listening test and .44 for the reading test. The alpha coefficients for the post-test were found to be .67 and .49, respectively. These low values can be attributed to the difficulty of the test for the participants.

3.3.2. Measures of Working Memory Capacity

This study made use of 3 tests to measure WM capacity, namely, Reading Span Test (English version), Reading Span Test (Turkish version), and Operation Span Test.

3.3.2.1. Reading Span Test

Both for Turkish and English versions, this study made use of a modified version of the original RST developed by Daneman and Carpenter (1980). The RSTs were matched for the syntactic and vocabulary complexity, frequency and the length of target-word, and sentence length. Two language versions of the test consisted of 42 unrelated simple sentences in the active voice. Each sentence ended with a different word. The test was administered on the computer in a computer lab. The researcher presented instructions and examples in detail. Next, the participants were given two practice trials. Finally, the participants were asked to read the instructions and start the experiment.

In the RST-Eng, each sentence ended with a different concrete noun, and occurred only once during the test. The test consisted of five sets of sentences. Starting from two sentences, the set size increased to five sentences with three practice trials for each set size. Each sentence was 11–13 words in length and presented on-line by displaying one sentence after another at 7-second intervals. Additionally, as a processing task, a syntactic judgment task was integrated into the RST-Eng to ensure that participants process every sentence for syntax rather than focusing only on the final words. The test included 21 grammatical (e.g. His younger brother played guitar in a rock and roll band.) and 21 ungrammatical sentences (e.g. Secretaries usually have an older computer and on telephone their a desk.).

During the test, each sentence appeared only once. The sentences and the order of set sizes were randomized and delivered by the computer program automatically. During the test the participants pressed “T” on the keyboard to indicate whether a given sentence was grammatical or “F” for ungrammatical. Once all the sentences for a set are displayed, a question mark appeared on the empty screen and the participants were instructed to recall the words by typing their responses into a text field onscreen, and they had unlimited time to recall the words.

The Turkish version of RST was adapted from Alptekin & Erçetin (2010). Parameters of sentence length, syntactic and vocabulary complexity, frequency and the length of target-words for the Turkish version of the RST were matched with the English version. Unlike English, word order in simple Turkish sentences is generally as '*Subject Object Verb'*. Owing to that reason, the sentences always ended with a verb. All the verbs were motion verbs and third person singular. Parallel with the English version, the Turkish version of the RST alsoincluded 21 grammatical (e.g. Bahçede kuşların göçünü seyrederken bir kırlangıç gelip Ceren'in omuzuna kondu.) and 21 ungrammatical sentences (e.g. Kayaların ve makilerin arasından sekerek, vadiye çevikliğiyle ceylan bir indi.). During the test, each sentence appeared only once and was delivered randomly. Each sentence was 10–11 words in length. The same procedures of English version were followed for the Turkish version of RST.

Cronbach's α for the internal consistency reliability coefficients for the processing and storage tasks on the Turkish version were found to be .972 (M = 26.82, SD = 13.62) and .973 (M = 25.63, SD = 13.93) while those on the English version were .629 (M = 24.57, SD = 4.89) and .809 (M = 29.35, SD = 6.33) respectively.

3.3.2.2. Operation Span Test

The OST used in the study was a version suggested by Kane et al. (2004). The test was administered on the computer in a computer lab. The researcher presented instructions and examples in detail. Next, the participants were given two practice trials. Finally, the participants were asked to read the instructions and complete the experiment. Each time the program randomly displayed a mathematical operation and a to-be-remembered word on the computer screen. The test included 21 correct (e.g. (2x2) +2 = 6 KALEM), and 21 incorrect (e.g. (4x2)-1 = 6 GEMİ) equations.

The test consisted of five computer-paced trials of operation-words pairs, ranging from 2 to 5 sets. Each operation appeared only once and the order of set sizes were randomized and delivered by the computer program automatically. During the test the participants pressed “T” on the keyboard to indicate whether a given equation was correct or “F” for incorrect. Once all the operations and words for a set were displayed, a question mark appeared on the empty screen and the participants were instructed to recall the words by typing their responses into a text field onscreen. The OSPAN task consisted of 42 trials that included three to seven operation–word pairs.

Cronbach’s α for the internal consistency reliability coefficients for the processing and storage tasks on the OST found to be .663 (M = 38.00, SD = 3.15) and .784 (M = 37.63, SD = 4.00).

3.3.3. The software

Traditionally, memory span tests are administered under the control of experimenters. Reliability and validity of experimenter-paced task administration have been repeatedly reported by various studies. However, with the development of computer software, computer-paced span tasks have become more popular during recent years (Conway et al., 2005; Unsworth et al., 2005; Unsworth and Engle, 2006). Research revealed that experimenter-paced and computer-paced span tasks share some overlap, but also measure additional and distinct processes. Besides, computer-paced span tasks were demonstrated to be highly reliable measures memory span (Bailey, 2012). When compared to experimenter-paced task administration, computer-paced span tasks possess the advantage of being administered in a group setting. These tasks can be used in web-based studies since the presence of the researcher is not required. Besides, they can be shared easily between the researchers at different research sites (Bailey, 2012). Since the present study involved a correlational design, which requires large sample sizes, computer-paced software was used. The software was developed by the author with the help of a computer expert. The memory test program made it possible to administer the tests in a group setting. The program was based on the same principles of other standard programs (e.g., Superlab). The responses of each participant during a test were automatically stored in a folder, which made the scoring procedure easier for the researcher.

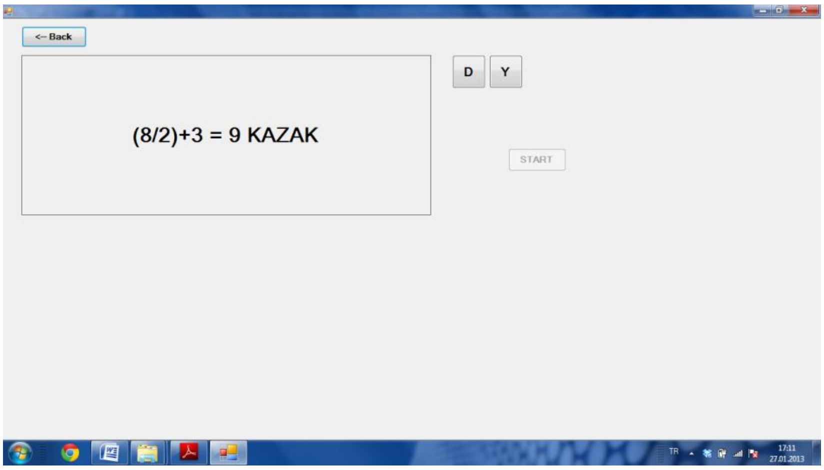

Figure 6. A sample screen shot of the interface of the computer program

3.4. Pilot Study

In order to monitor possible problems both for administration and implementation of the computer based WM tests, a pilot study was carried out with randomly selected 20 students (12 female, 8 male). In the first session the participants were informed briefly about the aim and the scope of the study in a computer lab. Then the researcher made a demonstration of a trial set of two sentences by the help of a projector. The researcher presented instructions and examples in detail. Following that the participants were given two practice trials. In addition, the participants were asked to read instructions and start the tests.

The participants performed the OST as the initial test. The OST took about 5 minutes. After all the participants had completed the first test, they were informed about the RST-Tr and followed the same procedures. At the end of the tests, the researcher asked each participant if they had difficulty in using the program. The participants' feedback was positive in that they did not have any problem in using the software. The second session was carried out three days later. The participants followed the same procedures for the RST-Eng. The test lasted about 6 minutes. The participants who participated in the pilot study were excluded from the main study. The pilot study enabled the researcher to monitor possible problems both for administration and implementation of the computer based WM tests.

3.5. Procedures for data collection

The participants completed 5 sessions during the 14-week semester for each session the following steps were followed for data collection.

The first session was carried out during the first week of the semester and consisted of the CAE including one reading comprehension and one listening comprehension test. For each group of participants, the CAE tests were administered in groups and invigilated by the researcher. After two weeks, in session 2, the participants completed both the RST-Tr and the OST one after another in the same session. The tests were held during the semester on different days for each group of participants. After two weeks, in session 3, the participants completed the RST-Eng in groups. Again the tests were held during the semester on different days for each group of participants. Finally, session 5 was carried out during the last week of the semester and consisted of an equivalent version of CAE given as the pre-test. For each group of participants, the CAE tests were administered in groups and invigilated by the researcher.

Table 2.
Language distribution of the tests used in the study

	Turkish (L1)	**English (L2)**
Working Memory	Reading Span Test Operation Span Test	Reading Span Test
Text Comprehension		Cambridge Advanced Exam Reading Section Listening Section

3.6. Data Analysis

The storage scores (i.e., the total number of accurately recalled sentence-final words in the L1 and L2 RSTs and the total number of accurately recalled words associated with the operations in the OST) constituted the WM measure. For each WM measure, the participants were divided into low- and high-WM groups based on a median-split procedure on the storage scores. Thus, the participants below the median were categorized into low-WM group while those above the median were categorized in high-WM group.

In order to answer the research questions, descriptive statistics were first obtained. Then, 2x2x2 mixed ANOVAs were conducted separately on listening and reading scores with time of testing (pre vs. post) as the repeated measures factor, WM capacity (low vs. high), and proficiency level (intermediate vs. upper-intermediate) as between groups factors. The statistical procedures were carried out via IBM SPSS Statistics 20.0.

3.7. Chapter Summary

This chapter focused on the methodological procedures implemented in the present study. The research is based on a quasi-experimental design with the pretest-posttest nonequivalent groups with intact classes. The first part of the chapter comprises the procedures followed during the pilot study.

Second, the details of the participants were the presented in detail. Next part of the methodology section focused on the instrumentation. The measures of WMC that are used in the data collection procedure were discussed in detail. Besides, the working principles of computer software developed by the author were explained in detail. Finally, procedures for data analysis were discussed at the end of the chapter.

CHAPTER 4

RESULTS

The main purpose of the current study was to investigate the role of WM capacity on listening and reading comprehension in the L2 with a specific focus on the mediating role of L2 proficiency. To this end, it explored whether the relationship between WM and L2 listening/reading differed depending on the level of proficiency before and after instruction. Thus, the study examined the interactions among WM capacity (low vs. high), proficiency level (intermediate vs. upper-intermediate) and time of testing (before instruction vs. after instruction) in relation to their effects on L2 listening and reading comprehension. It was hypothesized that WM would have a more influential role in listening/reading comprehension at the intermediate level than the upper-intermediate since its influence weakens, as skills get more automatic. Thus, it was expected that high- and low- WM participants would perform differently on tests of listening and reading comprehension at the beginning of semester.

Since WM capacity was assessed through multiple measures, namely OST, L1 RST and L2 RST, separate analyses were conducted with each WM measure in order the main effect of WM as well as its interaction with proficiency level and time of testing. Therefore, the results are presented separately for each WM measure.

4.1. WM as measured by OST, time of testing and proficiency level in L2 listening and reading comprehension

Table 3 displays the descriptive statistics for WM as measured by OST, time of testing and proficiency level in terms of listening and reading comprehension in the L2. Table 3 indicates that both the listening and reading scores of the intermediate learners increased from the pre-test to the post-test. On the other hand, the upper-intermediate learners seem to have improved in terms of reading rather than listening. In fact, the post-test listening scores of the latter group is slightly lower than their pre-test scores, which can be explained by lack of listening instruction during the freshman year. Comparing the means of the low- and high-WM participants in each proficiency group, it can be seen that the latter group has a slightly higher mean than the former. However, the difference seems to be negligible.

Table 3.

Descriptive statistics for WM as measured by OST, time of testing and proficiency level

Test	Proficiency	WM	Pre-test M	Pre-test SD	Post-test M	Post-test SD	N
Listening	Intermediate	Low	8.35	2.67	9.25	2.39	28
		High	8.78	2.82	9.40	2.38	42
		Total	8.61	2.75	9.34	2.38	70
	Upper-intermediate	Low	10.52	2.79	10.42	1.77	19
		High	11.86	2.69	10.73	1.38	15
		Total	11.11	2.79	10.55	1.59	34
Reading	Intermediate	Low	11.44	2.81	12.34	2.61	29
		High	11.14	2.89	13.19	2.89	42
		Total	11.26	2.84	12.84	2.79	71
	Upper-intermediate	Low	10.00	2.24	14.20	1.97	15
		High	10.33	2.35	14.11	2.54	18
		Total	10.18	2.27	14.15	2.27	33

In order to determine the interaction of WM, proficiency level, and time of testing, separate 2x2x2 mixed ANOVAs were conducted on listening and reading scores with time (pre vs. post) as the repeated measures factor, WM (low vs. high), and proficiency (intermediate vs. upper-intermediate) as between groups factors.

The ANOVA results for listening comprehension (Table 4) revealed a significant main effect of proficiency level as well as a significant interaction between time of testing and proficiency level. The other effects were not significant. The interaction between time of testing and proficiency level is illustrated in Figure 7.

Table 4.

ANOVA summary table for WM as measured by OST, proficiency level, time of testing and listening comprehension

Source	SS	df	MS	F	Sig.	Partial η^2
Proficiency	167.928	1	167.928	22.10	.000	.181
WM	13.979	1	13.979	1.84	.178	
Proficiency * WM	3.197	1	3.197	.42	.518	
Error	759.851	100	7.599			
Time	.209	1	.209	.04	.834	
Time * Proficiency	21.153	1	21.153	4.45	.037	.043
Time * WM	4.739	1	4.739	.99	.320	
Time * Proficiency * WM	1.591	1	1.591	.33	.564	
Error(time)	475.053	100	4.751			

Figure 7 shows that the mean difference between the intermediate and upper-intermediate group is much larger on the pretest compared to the posttest. The mean of the intermediate group increased from the pre-test to the post-test whereas that of the upper-intermediate group decreased.

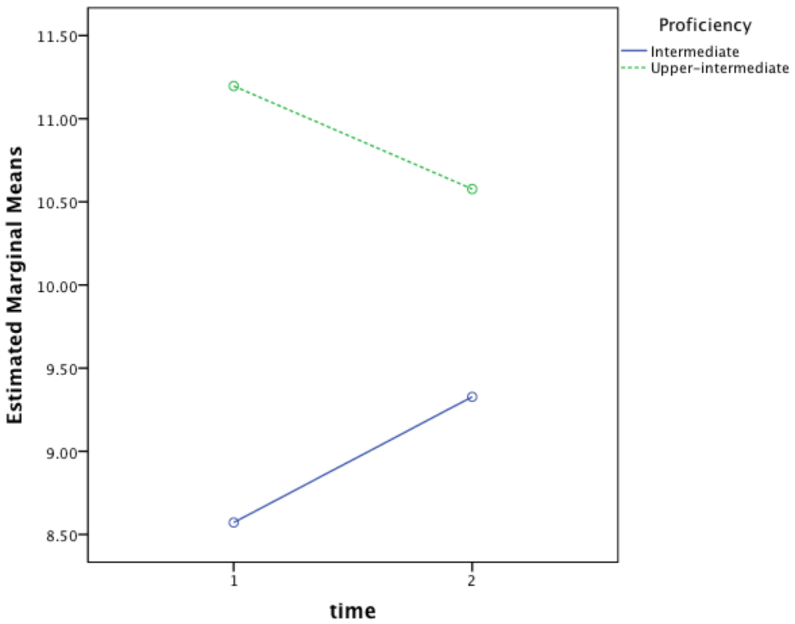

Figure 7. Interaction between time of testing and proficiency level for listening

Tests of simple main effects with Bonferroni adjustment revealed that the difference between the two groups was statistically significant both the beginning of semester ($p < .001$) and at the end of semester ($p < .01$). In other words, the upper-intermediate group outperformed the intermediate group at both times of testing.

The ANOVA results for reading comprehension (Table 5) indicate a significant main effect of time of testing as well as a significant interaction between time of testing and proficiency level. The other effects were not significant.

Table 5. ANOVA summary table for WM as measured by OST, proficiency level, time of testing and L2 reading comprehension

Source	SS	Df	MS	F	Sig.	Partial η^2
WM	1.705	1	1.705	.20	.653	
Proficiency	.743	1	.743	.08	.766	
WM * Prof.	.242	1	.242	.02	.865	
Error	837.468	100	8.375			
Time	330.414	1	330.414	56.18	.000	.360
Time * Proficiency	70.181	1	70.181	11.93	.001	.107
Time * WM	1.471	1	1.471	.25	.618	
Time * Prof. * WM	6.856	1	6.856	1.16	.283	
Error(time)	588.053	100	5.881			

The interaction between time of testing and proficiency level is illustrated in Figure 8, which shows that the intermediate group has a slightly higher pre-test mean than the upper-intermediate group. On the other hand, the reverse is the case on the post-test. In other words, the latter outperforms the former group. Tests of simple main effects with Bonferroni adjustment revealed that the difference between the two groups was not significant ($p > .05$) at Time 1 while it was significant at Time 2 ($p < .05$). In other words, the groups did not differ in reading performance at the beginning of semester. The upper-intermediate group surpassed the intermediate group at the end of the semester.

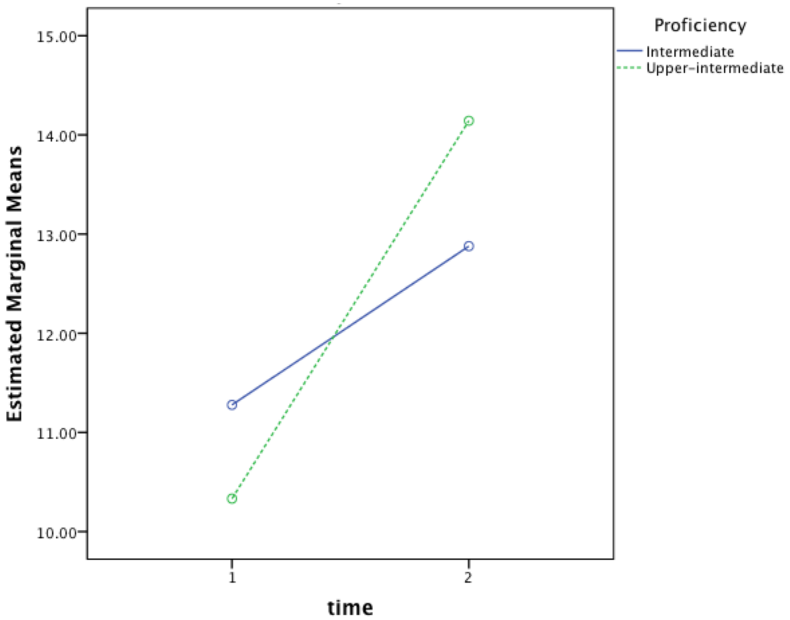

Figure 8. Interaction between time of testing and proficiency level for reading comprehension

4.2. WM as measured by L1 RST, time of testing and proficiency level in listening and reading comprehension in the L2

Table 6 presents the descriptive statistics for WM as measured by the RST in the L1, time of testing and proficiency level in terms of listening and reading comprehension in the L2. Since the marginal means for time of testing and proficiency level are the same as those in Table 1, they are presented here but not discussed any further (see p. 66). On the other hand, the cell means for WM capacity are different from those of Table 1. Specifically, Table 4 indicates that when WM is measured by the RST in the L1, the mean difference between low- and high WM participants is worth noting especially in the intermediate group and in terms of listening comprehension. In other words, the difference between the low- and high-WM participants is negligible in the upper-intermediate group in terms of both listening comprehension and reading comprehension.

Table 6.

Descriptive statistics for WM as measured by L1 RST, time of testing and proficiency level

Test	Proficiency	WM	Pre-test M	Pre-test SD	Post-test M	Post-test SD	N
Listening	Intermediate	Low	7.55	2.69	9.14	2.05	27
		High	9.27	2.60	9.46	2.57	43
		Total	8.61	2.75	9.34	2.37	70
	Upper-intermediate	Low	10.65	2.65	10.80	1.60	26
		High	12.62	2.87	9.75	1.38	8
		Total	11.11	2.79	10.55	1.59	34
Reading	Intermediate	Low	11.32	2.55	13.03	3.01	28
		High	11.23	3.04	12.72	2.66	43
		Total	11.26	2.84	12.84	2.79	71
	Upper-intermediate	Low	10.04	2.40	14.16	2.40	25
		High	10.62	1.84	14.12	1.88	8
		Total	10.18	2.27	14.15	2.26	33

The ANOVA results for listening comprehension (Table 7) point to a significant main effect of proficiency level as well as a significant interaction between time of testing and proficiency level. Also, there is a significant interaction between time of testing and WM capacity. The other effects were not significant.

Table 7.

ANOVA summary table for WM measured by L1 RST, proficiency level, time of testing and L2 listening comprehension

Source	SS	df	MS	F	Sig.	Partial η^2
Proficiency	157.247	1	157.247	21.33	.000	.176
WM	19.499	1	19.499	2.64	.107	
Proficiency * WM	2.838	1	2.838	.38	.536	
Error	737.076	100	7.371			
Time	1.985	1	1.985	.45	.501	
Time * Proficiency	45.246	1	45.246	10.38	.002	.094
Time * WM	43.960	1	43.960	10.09	.002	.092
Time * Proficiency * WM	5.881	1	5.881	1.35	.248	
Error(time)	435.645	100	4.356			

The interaction between time of testing and WM capacity is illustrated in Figure 9, which indicates that the difference between low- and high-WM capacity learners is greater on the pre-test compared to the post-test. Tests of simple main effects with Bonferroni adjustment revealed significant differences between the low- and high WM groups at Time 1 ($p > .01$) but not at Time 2 ($p > .05$). In other words, the influence of WM capacity on L2 listening comprehension decreases as the proficiency level increases. Since the interaction between time and proficiency level is already presented in the previous section (see p. 69), it is not provided here again.

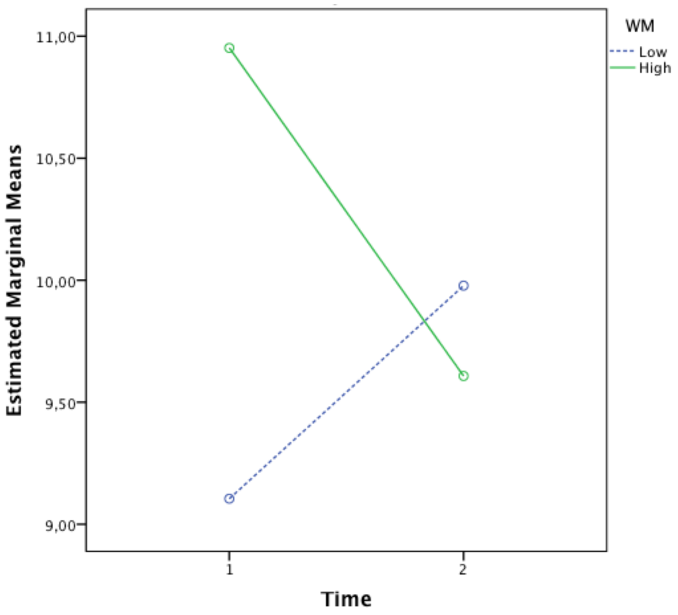

Figure 9. The interaction between time of testing and WM in L2 listening comprehension

As for reading comprehension, Table 8 indicates a significant main effect of time of testing and a significant interaction between time of testing and proficiency level (see p. 69). Unlike listening comprehension, the main effect of WM on reading comprehension was not significant nor was the interaction between time of testing and WM.

Table 8.

ANOVA summary table for WM as measured by L1 RST, proficiency level, time of testing and L2 reading

Source	SS	df	MS	F	Sig.	Partial η^2
Proficiency	.913	1	.913	.10	.742	
WM	.048	1	.048	.00	.940	
Proficiency * WM	2.030	1	2.030	.24	.624	
Error	837.918	100	8.379			
Time	261.486	1	261.486	43.68	.000	.304
Time * Proficiency	43.561	1	43.561	7.27	.008	.068
Time * WM	1.597	1	1.597	.26	.607	
Time * Proficiency * WM	.347	1	.347	.05	.810	
Error(time)	598.549	100	5.985			

4.3. WM as measured by L2 RST, time of testing and proficiency level in listening and reading comprehension in the L2

Table 9 presents the descriptive statistics for WM as measured by the RST in the L2, time of testing and proficiency level in terms of listening and reading comprehension in the L2.

Table 9.

Descriptive statistics for WM as measured by L2 RST, time of testing and proficiency level

			Pre-test		Post-test		
Test	Proficiency	WM	M	SD	M	SD	N
Listening	Intermediate	Low	8.19	2.78	9.51	2.32	31
		High	8.94	2.71	9.20	2.44	39
		Total	8.61	2.75	9.34	2.37	70
	Upper-intermediate	Low	10.58	2.79	10.58	1.58	24
		High	12.40	2.45	10.50	1.71	10
		Total	11.11	2.79	10.55	1.59	34
Reading	Intermediate	Low	10.97	3.07	12.55	2.74	34
		High	11.54	2.62	13.10	2.84	37
		Total	11.26	2.84	12.84	2.79	71
	Upper-intermediate	Low	10.10	1.97	14.20	2.37	20
		High	10.30	2.75	14.07	2.17	13
		Total	10.18	2.27	14.15	2.26	33

The ANOVA results for listening comprehension (Table 10) indicate a significant main effect of proficiency level, and a significant interaction between time of testing and proficiency level. Also, there is a significant interaction between time of testing and WM.

Table 10.

ANOVA summary table for WM as measured by L2 RST, proficiency level, time of testing and L2 listening comprehension

Source	SS	df	MS	F	Sig.	Partial ŋ²
Proficiency	168.595	1	168.595	22.130	.000	.181
WM	11.880	1	11.880	1.559	.215	
Proficiency * WM	4.164	1	4.164	.547	.461	
Error	761.849	100	7.618			
Time	.258	1	.258	.056	.813	
Time * Proficiency	30.324	1	30.324	6.627	.012	.062
Time * WM	22.043	1	22.043	4.818	.030	.046
Time * Proficiency * WM	1.742	1	1.742	.381	.539	
Error(time)	457.555	100	4.576			

The interaction between time of testing and WM capacity is illustrated in Figure 10, which indicates that the difference between low- and high-WM capacity learners is greater on the pre-test compared to the post-test. Tests of simple main effects with Bonferroni adjustments revealed a significant difference between the two groups at Time 1 ($p < .05$) but not at Time 2 ($p > .05$). In other words, the influence of WM capacity in L2 listening comprehension decreases as the proficiency level increases. Since the interaction between time and proficiency level is already presented in the previous section (see p. 69), it is not provided here again.

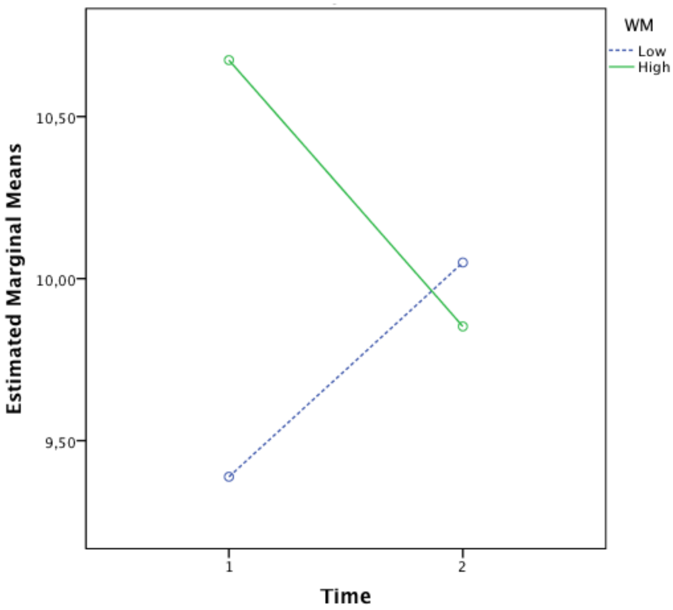

Figure 10. The interaction between time of testing and WM in L2 listening comprehension

As for reading comprehension, Table 11 indicates a significant main effect of time of testing and a significant interaction between time of testing and proficiency level (see p. 69). However, the results do not show any significant effects of WM on reading comprehension. These findings are similar to those obtained through L1 RST.

Table 11.

ANOVA summary table for WM as measured by L2 RST, proficiency level, time of testing and L2 reading comprehension

Source	SS	df	MS	F	Sig.	Partial η^2
Proficiency	.700	1	.700	.08	.772	
WM	3.952	1	3.952	.47	.492	
Proficiency * WM	2.919	1	2.919	.35	.554	
Error	829.090	100	8.291			
Time	331.452	1	331.452	55.27	.000	.356
Time * Proficiency	60.581	1	60.581	10.10	.002	.092
Time * WM	.337	1	.337	.056	.813	
Time * Proficiency * WM	.262	1	.262	.04	.835	
Error(time)	599.712	100	5.997			

To summarize the results presented in this section, findings do not point to a significant role of WM capacity in terms of reading comprehension. On the other hand, its influential role could be detected in terms of listening comprehension when WM is measured by L1 or L2 RST rather than OST. Specifically, the difference between low- and high-WM capacity learners is greater on the pre-test compared to the post-test, suggesting that the influence of WM capacity in L2 listening comprehension decreases as the proficiency level increases. Since L1 and L2 RSTs lead to similar conclusions, the role of WM capacity in L2 listening and reading comprehension can more reliably be demonstrated when WM is measured through the RST either in the L1 or the L2 rather than the OST. Thus, it can be concluded that the findings obtained through the RST are robust. These findings will be discussed in the next section, excluding those obtained through the OST.

CHAPTER 5

DISCUSSION and CONCLUSION

This chapter will present the results of quantitative analyses in detail and discuss the findings of the study in accordance with the research questions and hypotheses stated in the previous chapter. Then the chapter will present implications. Finally, the limitations and delimitations of the study will be discussed and suggestions for further research will be provided.

Findings of the present study in relation to the first research question investigated the role of WM capacity in listening and reading comprehension in the L2. The study hypothesized that learners with larger capacities would have better listening and reading comprehension scores. However, the results indicated that WM did not have a significant role in reading comprehension. On the other hand, its role in listening comprehension was significant when WM was measured by L1 or L2 RST. Nevertheless, the OST scores did not show any significant relationship to either reading or listening comprehension. This seems to be surprising given the general view of WM capacity (Turner & Engle, 1989), which assumes that the same resources are used to support working memory storage and processing activities, regardless of the nature of the task (Dehn, 2008). Thus, one would expect the OST scores to yield similar results as the RST scores. The lack of a significant relationship between WM measured through OST and reading/listening comprehension can be explained by two main reasons. First, the descriptive statistics for OST scores revealed higher means and lower SDs both for processing (M = 38.00, SD = 3.15) and storage tasks (M = 37.63, SD = 4.00) on both tasks compared to L1 RST (M = 26.82, SD = 13.62) and L2 RST (M = 24.57, SD = 4.89). Thus, the internal consistency reliability values for the OST were higher than those for the RSTs (see pp. 59-60). Conway et al. (2005) emphasize that the secondary task of complex span tests such as reading span or operation span should be sufficiently demanding so that WM is actively involved in "keeping task-relevant information active and accessible in memory during the execution of complex cognitive tasks" (p. 771). Thus, it seems the processing task of the OST was not demanding enough to stimulate WM operations of active maintenance of information in the face of ongoing processing. Second, as revealed by Daneman and Merikle's (1996) meta analysis on the relationship between WM and L1 language comprehension, measures of WM that include a verbal processing component and a verbal storage component as in RSTs have the best predictive power. Based on these concerns, the discussion of the results in this chapter will exclude the results based on the OST.

The findings of the present study regarding the lack of a relationship between reading comprehension and WM capacity do not support the findings of studies demonstrating the existence of such a relationship either in L1 (e.g., Daneman & Carpenter, 1980; Daneman & Merikle, 1996) or L2 (e.g, Alptekin & Erçetin, 2009, 2010, 2011, Harrington & Sawyer, 1992; Leeser, 2007; Miyake & Friedman, 1998; Walter, 2004). On the other hand, the findings are in line with a number of studies that point to the lack of a significant relationship between WM and L2 learning. For instance, Juffs (2005) found that overall scores based on the Michigan Test's grammar and vocabulary sections did not correlate with either L1 or L2 RST scores. This finding was consistent for learners from different L1 groups, namely Chinese, Japanese, and Spanish. Similarly, Felser and Roberts (2007) did not find any significant effects of WM on Greek-speaking English learners' processing of wh-dependencies. Adams and Shahnazari-Dorcheh (2014), investigating the role of WM in L2 reading comprehension with regard to different proficiency levels, demonstrated a significant relationship for the beginner group but non-significant relationships for the intermediate and advanced learners. Thus, WM's role seems to depend on how challenging the reading task is for the learner group. In this regard, treating the concept of reading comprehension as a two-dimensional construct, Alptekin and Erçetin (2010) demonstrated that WM was significantly associated with inferential comprehension for advanced learners of English while its relationship to literal comprehension was not significant. They concluded that unlike literal understanding, inferential comprehension requires controlled and effortful processing even for proficient L2 learners. Thus, the way reading comprehension is conceptualized is crucial in exploring how it is related to WM capacity.

As for listening, the results revealed a significant relationship between listening comprehension and WM when it was measured by L1 or L2 RST. These findings support those of other studies that point to a meaningful relationship between WM capacity and listening comprehension (e.g., Juffs and Harrington, 2011; Kormos & Safar, 2008; Londe, 2008; Mackey, Adams, Stafford & Winke, 2010; Marx & Roick, 2012; Shanshan & Tongshun , 2007; Tsuchihira, 2007). For instance, Kormos and Safar (2008) showed that WM capacity measured by a complex span measure, correlated very highly with overall English language competence including listening test scores of students. Similarly, Tsuchihira (2007), Shanshan and Tongshun (2007), and Londe, 2008 concluded that learners with larger WM capacities were more likely to have better abilities in listening comprehension based on the significant relationship of listening comprehension with both L1 and L2 WM capacities.

Findings of the present study in relation to the second research question investigated whether the role of WM in L2 listening and reading comprehension was mediated by L2 proficiency. It was hypothesized that WM would be more influential at lower levels since WM's effect weakens as skills get more automatic. In other words, since WM operates through conscious controlled processes, it was hypothesized to be more influential at lower levels. This question was investigated not only through a cross-sectional comparison of intermediate and upper-intermediate learners at the beginning and end of semester but also through a developmental follow-up of the same proficiency group from the start of instruction to the end. Cross-sectionally, it was observed that the interaction between WM and proficiency groups was not statistically significant either in terms of listening comprehension or reading comprehension. Developmentally, it was observed that the difference between low- and high-WM capacity learners was greater on the pre-test compared to the post-test for listening comprehension but not for reading comprehension since the interaction between WM and time of testing was significant for the former, not for the latter. When WM was measured by the RST in the L1, the results indicated that the difference between the low- and high-WM participants was negligible in the upper-intermediate group in terms of both listening and reading comprehension before and after instruction. However, when WM capacity was measured by the RST in the L2, the results for listening comprehension indicated a significant interaction between time of testing and WM.

The cross-sectional and developmental comparisons seem to contradict each other. This could be explained by low degrees of reliability in the placement of learners into proficiency groups. Although the results of the current study shows that the proficiency groups were somewhat different in their listening and reading performance, error associated with placement decisions based an institutional placement test as in the current study, might contaminate the results. Additionally, proficiency level differences between the two groups may not be large enough; a comparison of the intermediate group with an advanced group might have revealed different results. Therefore, the developmental results may be more dependable. As such, it can be concluded that, while no significant role of WM is observed in L2 reading comprehension, the level of proficiency mediates its role in listening comprehension.

The theoretical rationale behind these findings could be that listening requires different cognitive processes when compared with reading (Engle, Kane, & Tuholski, 1999). While readers can regulate their pace during the comprehension process, listeners cannot do the same since the procedure requires simultaneous processing of the received input (Just & Carpenter, 1987). Listeners create understanding by starting with the smallest units of such as individual sounds, or phonemes that are combined to form words in order to compose phrases, clauses, and sentences. Finally, individual sentences are combined to create ideas and concepts to make the aural input comprehensible. Hence, during all these stages of L2 listening, the learners' performances depend on their L2 proficiency. As they master their L2, they use less cognitive sources to comprehend the spoken input. In other words, less proficient language learners spend much of their WM capacities on holding information in their memories while processing listening comprehension tasks. Conversely, more proficient learners can process the spoken information more easily and share more cognitive resources to operationalize the information representations. As a result, it becomes evident that the role of WM capacity in L2 listening comprehension is mediated by L2 proficiency.

It is possible to state that limited capacity of WM affects L2 listening comprehension when the listener is not competent enough both in terms of linguistic knowledge and processing skills in the L2. For the present study, lower-level listeners needed to spend more cognitive resources to comprehend what they hear during the listening process. Hence, the high-WM participants were more successful on listening comprehension before instruction. As the participants were exposed to instruction and their skills got more automatic, the influence of WM weakened, suggesting that WM plays a greater role in controlled processing of L2 listening.

L2 comprehension skills at lower levels of proficiency are performed through controlled processes that depend on explicit linguistic knowledge of declarative memory sources. According to the declarative/procedural (DP) model of language acquisition (Ullman, 2001a, 2001b), L2 learners tend to depend more on their declarative memory (DM) for the morphosyntactic computations. With sufficient exposure to language and with practice, L2 learners may depend more on their procedural memory (PM) (Ullman, 2005). Therefore, the DP model assumes that with increasing exposure and proficiency, L2 learning may come to rely on the procedural memory system. This switch from declarative to procedural memory may be mediated by practice and experience with the L2. As learners proceduralize their L2 skills, they rely less on their WM capacities (Carpenter, Morgan-Short, and Ullman, 2009).

To conclude, the role of WM in L2 listening comprehension before instruction changed after instruction only for the participants from low-proficiency group. This can be explained by the fact that, between pre-test and post-test, the intermediate learners received instruction since they were preparation year students. So, as a result of extensive exposure to L2, their PM became more active in terms of language processing and enabled the participants to spend most of their WM capacity on comprehending rather than decoding what they heard. On the other hand, no significant effect of testing time was found for the upper-intermediate group, since their exposure to language, as freshmen, was not as much as the intermediate.

5.1. Pedagogical implications

There is substantial evidence for the hypothesis that WM processes underlie individual differences in learning ability. WM is required in any learning task that requires manipulation and simultaneous storage and processing of information. Since most of the incoming information must pass through WM, the capacity and effective functioning of it determines the rate and extent of learning.

Classroom performance and the development of academic skills, such as reading comprehension and listening comprehension rely heavily on the sufficient functioning of WM sources. A large body of evidence has established strong relationships between WM components and certain areas of academic achievement such as reading comprehension, listening comprehension, spelling, vocabulary development, writing, and oral language production. In typical language classroom learning environments, common classroom activities involve the potential to impose heavy demands on storage and processing. Hence, learning becomes less successful when WM capacity is overloaded by activities in which new information needs to be integrated with previously stored knowledge simultaneously. It is evident that some learners process input more effectively than others because they are more likely to process cognitive based linguistic operations easily during language comprehension process. This is due to the fact that some learners have greater WM capacity, and are able to carry out both lower- and higher-level processes at greater speed and accuracy when compared to the others. Yet, learners with lower WM capacity fail to operationalize higher-level processes when focusing on lower-level decoding processes. Within the classroom setting, WM capacity differences between the learners should be taken into consideration and teachers should move away from what learners do not know about the text and put more emphasis on what they do know about it.

Since comprehension tasks play a crucial role in learners' reading and listening practice, appropriate tasks are essential in improving language comprehension. Therefore, it is critical for material developers to design effective language tasks for learners from lowers language levels. Hence, both oral and written texts should be designed considering the limited capacity of WM. Language teachers can achieve this by simplifying the vocabulary, using common vocabulary rather than unusual, using simplified forms of sentences rather than using sentences with complex structures, and reducing the sentences length.

As for listening comprehension, more importance should be given to pre-listening activities to enable learners to apply their prior knowledge during listening. Besides, rather than focusing on the outcome of listening activities teachers should focus more on the listening process. With regard to higher task demands, unfamiliar information might overload the limited processing capacity and hinder a better performance. Hence, building activities that will enable learners to use their prior knowledge such as topic, culture, and genre, will lower WM load and help them to compensate for failure to understand speech sounds. Also, other factors such as speaker accent, speech rate, and cultural content should be taken into consideration when designing listening courses.

As for reading, differences in WM capacity can greatly influence the success of reading comprehension, when the text is difficult or complex. Especially at early stages of language learning, learners lose track in complex tasks while trying to focus on decoding process. However, as learners become more proficient in the target language, they decode the letter automatically and facilitate reading comprehension by reducing the WM resources necessary for decoding words. Also, other textual factors such as, overall length of texts, information density, directness and concreteness, and culturally specific vocabulary occupy a large part of the WM capacity. Hence, instructors should be careful when selecting coursebooks or extra materials to be presented in the classroom.

An important implication for L2 reading comprehension is that automatic word-recognition ability is a critical component of fluent reading. So, language instructors should enhance automaticity skills by inserting fluency practices in any reading syllabus.

We should also remember that, reading comprehension also depends on the development of various language skills such as vocabulary development, verbal abilities, prior knowledge, and reading decoding skills. While trying to comprehend the text, WM simultaneously manipulates recently read information and recently retrieved information. Hence, any deficiency in any component of learners' language might increase both processing and storage load on the WM capacity. So, while designing language courses, instructors should consider language as a whole, and start teaching all the skills right from the beginning.

Also, the amount of cognitive load that can be caused by any learning material in the L2 should be taken into account during the stages of material development and presentation of it. Theoretically, this argument is based on

Sweller's (1988) Cognitive Load Theory of learning in relation to WM capacity of L2 learners, which postulates that because of its limited capacity architecture, if the WM resources are not used efficiently, the cognitive load on WM capacity would deteriorate learning. Sweller (1988) defines mental load as the load that is created by the characteristics of a particular task in question while mental effort is the amount of cognitive capacity or resources allocated by the learner to do a given task. If the internal complexity (intrinsic cognitive load) of an instructional material is high, the learners use their WM capacities at a maximum level. Hence, the cognitive load that is created by the instructional design of any learning material (extraneous cognitive load) should be decreased. Sweller suggests that, although intrinsic cognitive load is hard to manipulate, extraneous cognitive load can be changed through carefully designed instructions, and can promote an efficient and effective learning environment.

Another implication for improving the success of learners with low WM capacity is to help learners develop effective strategies for dealing with situations in which they experience WM failures. These strategies support the effective operation of WM processes, such as semantic encoding. By implying effective strategy training in their courses, language teachers can help their learners to have the required skills to cope with complex tasks and activities.

A concluding implication to compensate for low WM capacity is WM capacity training programs. The effect of WM training to other skills such as nonverbal and verbal ability, and word decoding might enhance cognitive functioning in language processing. By practicing computer based training tasks, the learners carry out numerous processes such as encoding, inhibition, maintenance, manipulation, shifting and controlling attention. As a result of this extensive practice, the learners can improve their ability to manage complex tasks by enhancing their WM capacity.

5.2. Limitations and Suggestions

This study has several limitations. The first limitation of the study is related to sample size. 86 Intermediate and 87 upper-intermediate level Turkish learners of English comprised the participants of the study. To be able to generalize the findings of the study, it should be replicated with learners at different levels of proficiency. Besides, the present study was carried out at a state university. Yet, different findings could be obtained with participants from different contexts.

Second, in the present study, the groups were placed according to an institutional placement test. However, the analysis of the data revealed that the cross-sectional and developmental comparisons seem to contradict each other. In other words, error associated with placement decisions based on the institutional placement test might have contaminated the results. Different results could be obtained by a more objective and reliable placement test.

Third, the tests were given in a controlled laboratory setting. This might have influenced the results, because some participants might have spent a longer period of time for typing and this could have affected their scores negatively. Hence, individual testing of the participants with their own hand writings could have eliminate this factor and help the researchers to find out the actual time that is spent on the retaining of the information.

APENDIX A

Turkish RST

1. Mezeleri afiyetle midesine indirirken bakarak denize keyifle rakısını masmavi içti.
2. Bisikleti görünce mutluluktan teyzesinin atladı ve şapur boynuna şupur öptü.
3. Sokağın köşesini döndü, karşıya geçti, biraz sonra bakkalın önünde durdu.
4. Tam dört yıl sonra, kılıç balığı avlamak için Küba'ya yelken açtı.
5. Gece yaklaşınca yırtıcı hayvanların korkusuyla bir ağaca çıktı ve orada uyudu.
6. Kayaların ve makilerin arasından sekerek, vadiye çevikliğiyle ceylan bir indi.
7. Üzerinde hala şortlar olduğu halde, kalkan Meksika'dan uçağa Amsterdam'a bindi.
8. Ali, evin içinde koştururken sehpanın üzerinde duran değerli vazoyu düşürüp kırdı.
9. Köylü kadın sabahın erken saatlerinde pazara gelip, yetiştirdiği ürünleri sattı.
10. Bay Tavşan, uyuklamakta olan tilkiye korkuyla sokulup, yavaşça ayağına dokundu.
11. Bahçede kuşların göçünü seyrederken bir kırlangıç gelip Ceren'in omzuna kondu.
12. Çekmecede bir makas, altı çatal bıçak, altın biraz gümüş ve buldu.
13. Sezon hazırlıklarını sürdüren takım, maçında 17 ilk tam gol attı.
14. Elini ceketinin cebine soktuğunda buldu kağıt ve bir hemen okudu.
15. Babam, her zamanki gibi gazetesini ve alıp koltuğa kahvesini oturdu.
16. Yaramaz Ali, adını büyük harflerle düzensiz bir şekilde defterine yazdı.
17. Annem yüzümü sildi, taburenin üstüne oturdu, sonra benim kolumdan tuttu.
18. Yeniden arkasını döndü ve hızla eve girerek kapıyı arkasından kapadı.
19. Bütün gün sıcaktan pişen ayaklarını bileğine serin kadar suyuna deniz soktu.
20. Puroyu dudaklarının arasında çevirdikten sonra ucunu ısırdı ve halıya tükürdü.
21. Dansın sonunda deri ceketini şöyle bir döndürüp hızla ayaklarımızın önüne serdi.
22. Günlerce balkonda baktığımız minik nihayet çırparak serçe kanatlarını bugün uçtu.
23. Paraşütü açılmadığı için 160 km hızla üstüne ağacının böğürtlen düştü.
24. Yüzlerce askere yemek yapmak için, patates kadar akşama sabahtan soydu.
25. Eşsiz doğa görüntüsünü rahatlıkla seyretmek camı için buğulanan mendiliyle sildi.
26. Köpek, arka bacakları üzerinde sıçrayarak kar çalışırken yakalamaya tanelerini havladı.
27. Evin dış duvarlarını pembenin, mavinin, sarının ve yeşilin yumuşak tonlarıyla boyadı.
28. Manastıra karşı oturdu ve ayaklarının dibine uzanan bir köpek yavrusunu okşadı.
29. Ukrayna'nın güneyinde bir Nil atlatarak, timsahı Azak Denizi'ne bakıcısını kaçtı.
30. Gök gürültüsünden çok korkan Deniz, yatağına yatıp yorganını başına çekti.
31. Polonya'da 19 yıl önce komaya şahıs bir geçen giren hafta uyandı.
32. Öğretmen, notları okumadan önce sınavdaki yanlışlarımız hakkında uzun uzun konuştu.
33. Gümüş rengi Bursa ipeğinden gömleği renkli üzerine kaşmir yelek sarı giydi.
34. Gözleri dolan yaşlı adam, cebindeki çocuğa bozuk mendil satan paraları verdi.
35. Annesi istediği oyuncağı almayınca kendini yere attı ve saatlerce ağladı.
36. Yıkama programı biten makineden çıkardığı çamaşırları sobanın yanına açtığı çamaşırlığa astı.
37. Babam, bugün işten erken çıkıp istasyonuna almaya tren anneannemi gitti.
38. 80 yaşındaki Çinli ihtiyar, saçlarını 26 yıl sonra ilk kez yıkadı.
39. İzmir'de birkaç aydır boş olan iki katlı binanın balkonu çöktü.
40. Bisikletle bayır aşağı giderken düştüm, iki bacağım balon günde gibi şişti.
41. Yağmurda ıslanmış köpek, içtikten sobanın sütünü yanında keyifle sonra yattı.
42. Kahraman itfaiyeci, suyu yangının üstüne püskürttü ve yangını büyümeden söndürdü.

APPENDIX B

English RST

1. All that remained in the box one lunch was salted nut.
2. His younger brother played guitar in a rock and roll band.
3. We had a tasty lunch, consisting of chicken, salad, and an apple.
4. The girl picked up her bag and down to went the gym.
5. Jenny helped her father put a blanket the old on bed.
6. He quickly drank some of the milk and then washed the glass.
7. Secretaries usually have an older computer and on telephone their a desk.
8. I saw a child and her father river the playing near ball.
9. He looked across the room and saw a person holding a gun.
10. The only thing left in the kitchen cupboard was a broken cup.
11. Several guests joined us for our first by camp the dinner lake.
12. We faced trouble because the stuck got car in the sand.
13. I dreamed that I was in with my field a sheep.
14. You have a copy of the list of words on the board.
15. The last thing he did was take to nice a hot bath.
16. The young woman and her boyfriend they thought a saw dog.
17. The woman screamed and old slapped the man in the face.
18. In order to attend the dinner she needed to buy a dress.
19. We need at least a week to visit such a historic site.
20. She took a deep breath and into reached rusty the box.
21. A beautiful brown and green grasshopper jumped onto my new shoe.
22. A young female employee accidentally a coffee of cup spilled into his lap.
23. He wore his red pajamas and set his small alarm clock.
24. The man was too afraid to drive he the caught so bus.
25. The people in northern Europe always like to travel by train.
26. The drums beat and the flutes sang and held spectators their the breath.
27. He wanted to leave his bags and jacket in the hotel room.
28. They knew that it was impolite to eat spaghetti with a spoon.
29. She thought it would be nice to drink at a fancy bar.
30. She woke up very early and then chocolate made a cake.
31. You will be amazed by the number cars on the of street.
32. Laura had little education and she never paying had a good job.
33. In Spain and Latin America go workers lunch for home and a nap.
34. At night the prisoners escaped a through in hole the wall.
35. The hunting knife was so sharp that it cut his right hand.
36. The clerk in the department store put the presents in a bag.
37. She soon realized that the man forgot to leave the room key.
38. The first driver out in the morning up picks the always mail.
39. All morning the two children sat under and talked a tree.
40. I am the friend of the man who brought you the duck.
41. He had long hair and a beard that reached down to his chest.
42. The Egyptians were the civilization recorded first to use wedding the ring.

APPENDIX C

Operation Span Test

1. (4x2)-1 = 6 GEMİ
2. (2x6)-4 = 8 KARPUZ
3. (2x2)+1 = 5 BİNA
4. (8/2)+3 = 9 KAZAK
5. (2x2)-3 = 1 ÇANTA
6. (4/4)+5 = 7 ÇİÇEK
7. (4x2)-3 = 3 RADYO
8. (1x6)-4 = 2 DEFTER
9. (3/3)+8 = 9 HOROZ
10. (1/1)+7 = 6 SAAT
11. (3x2)+2 = 8 ÇİLEK
12. (1x2)+2 = 5 KOLTUK
13. (4x2)-4 = 4 KAPLAN
14. (8/4)+4 = 7 AĞAÇ
15. (6/3)+1 = 3 MUTFAK
16. (2x3)+3 = 9 AYAK
17. (2x2)+2 = 6 KALEM
18. (2x3)+3 = 7 UÇAK
19. (4/1)+1 = 5 DOLAP
20. (8/1)+1 = 9 GÖMLEK
21. (3x2)+1 = 6 KUZU
22. (7/1)+1 = 8 BURUN
23. (6/1)+2 = 7 GÖZLÜK
24. (5/1)-1 = 3 LASTİK
25. (1x6)-2 = 5 ODA
26. (1x7)-2 = 4 KEDİ
27. (4/2)+5 = 8 KAMYON
28. (4/2)+1 = 4 TABAK
29. (1x1)+1 = 2 KASA
30. (8/2)-2 = 4 ORMAN
31. (4/4)+4 = 6 KAVUN
32. (1x1)+2 = 3 KÖPEK
33. (2/2)+1 = 2 BİLET
34. (2x3)-1 = 7 HAVLU
35. (8/2)+1 = 5 KÜPE
36. (3x1)+3 = 9 AYNA
37. (3x1)+1 = 4 BANYO
38. (2x2)+4 = 8 ÜTÜ
39. (6/1)+1 = 7 ŞİŞE
40. (6/1)-2 = 7 ÖRTÜ
41. (3/1)+2 = 5 ETEK
42. (8/2)+3 = 8 KİTAP

APPENDIX D

A Sample Page of the Reading Test

You are going to read three extracts which are all concerned in some way with environmental issues. For questions **1–6**, choose the answer (**A, B, C** or **D**) which you think fits best according to the text.

Mark your answers on a separate answer sheet.

How much does the environment matter?

I've always been keen on environmental matters, but in the seventies, when I was growing up, the subject wasn't discussed as much as it is today. Perhaps the problem was just as bad then, but fewer people understood the implications of what was going on or realised something ought to be done about it. I remember being pleased when laws restricting the use of certain pesticides came in. They appeared to signal the start of a vast movement to improve the environment, an acknowledgement that all human beings had a right to clean air and water.

With hindsight, I think we were naive. There is much greater awareness of environmental issues these days, but I still don't think enough is being done about the problem. And since it's a global issue, individual countries can't tackle it by themselves. Reducing damage to the environment really must be an international effort. In a paradoxical way, the more we discover about the extent of the problem, the less we do about it. This is because problems like global warming are so huge that ordinary individuals don't feel they could possibly make any difference. I think that's the real danger facing us today – that we'll succumb to a feeling of helplessness instead of making a concerted effort to make our planet a safer and cleaner place for future generations.

1 Why does the writer mention the law restricting pesticides?
- **A** to show that her initial optimism about the environment was misplaced
- **B** to illustrate the steps that have been taken to protect the environment
- **C** to make people aware of the dangers of using pesticides indiscriminately
- **D** to demonstrate her early involvement in environmental movements

2 What fear does she express in the second paragraph?
- **A** People are still not aware of the extent of the problem.
- **B** Individuals' actions won't make any difference to the problem.
- **C** Nothing can be done to stop the effects of global warming.
- **D** The scale of the problem may prevent action being taken.

APPENDIX E

A Sample Page of the Listening Test

PART 1. ***You will hear three different extract. For questions 1-6, choose the answer (A, B, or C) which fits best according to what you hear. There are two questions for each extract.***

Extract One

You hear part of an interview with a Tai Chi instructor.

1 According to Ruth, Tai Chi
A needs to be performed indoors.
B is a series of exercises.
C represents a way of life. | 1 |

2 One of the most difficult things about learning Tai Chi is
A that you have to do it out of doors.
B achieving harmony between your movements.
C remembering the complex dance sequences. | 2 |

Extract Two

You hear two people discussing taking a year off before going to university.

3 Before going to work in China, the man had not expected the job to be
A rewarding.
B easy.
C boring. | 3 |

4 The woman's experience made her realise that
A marketing was a difficult career for women.
B working in a hotel was very challenging.
C she had made the wrong choice of career. | 4 |

Extract Three

You hear part of an interview with an illusionist.

5 Daniel's interest in magic arose from
A his ambition to become an entertainer.
B his desire to impress someone.
C his trips to the cinema as a boy. | 5 |

6 According to Daniel, one similarity between magic and film making is
A the debt they owe to technology.
B the disbelief they arouse in the audience.
C the power they have over the audience. | 6 |

PART 2. ***You will hear part of a talk by a writer who has written a biography. For questions 7-13, complete the sentences.***

7. The soaker has written a book about _________called Robert Tewbridge.
8. Tewbridge's father was a __________ in Scotland.
9. Tewbridge's parents wanted him to become ___________.
10. Tewbridge earned his living by writing __________ for various publications.
11. The speaker learnt a great deal about Tewbridge's characte from studying his ________.
12. It appears that Tewbridge and his _________ were close friends.
13. Tewbridge spent many years studying _________.

REFERENCES

Adams R, Shahnazari-Dorcheh M. (2014). The relationship between working memory and L2 reading comprehension. *Applied Research on English Language.* 3 (6):19-34

Alptekin, C. & Erçetin, G. (2009). Assessing the relationship of working memory to L2 reading: Does the nature of comprehension process and reading span task make a difference? *System*, 37(4):627-639.

Alptekin, C. & Ercetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading*, 33, 206–219.

Alptekin, C., & Erçetin, G. (2011). The effects of working memory capacity and content familiarity on literal and inferential comprehension in L2 reading. *TESOL Quarterly*, 45, 235–266.

Alptekin, C., & Erçetin, G. (2012). The explicit/implicit knowledge distinction and working memory: Implications for second-language reading comprehension. *Applied Psycholinguistics,* 34, 727-753.

Andrade, J. (Ed.). (2001). *Working memory in perspective*. Hove, UK: Psychology Press Ltd.

S. Andringa, N. Olsthoorn, C. van Beuningen, R. Schoonen & J. Hulstijn (2012). Determinants of success in native and non-native listening comprehension: an individual differences approach. *Language Learning*, 62, 49-78.

Atkinson, R.C. & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence and J. T. Spence (Eds) *The psychology of learning and motivation (Vol 2)*. London: Academic Press

Atkinson, R.C. & Shiffrin, R.M. (1971). The control of short term memory. *Scientific American,* August 1971, 225(2):82-90.

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

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Science,* 4, 417–423.

Baddeley, A. D. & G. Hitch (1974). Working memory. In G. H. Bower (ed.), *Recent advances in learning and motivation*. New York: Academic Press, 47–89

Baddeley, A., Logie, R., Nimmo-Smith, I., & Brereton, N. (1985). Components of fluent reading. *Journal of Memory and Language*, 24, 119-131.

Baddeley, A.D., Hitch, G.J., & Allen, R.J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, 61, 438-456.

Broadbent, D (1958). *Perception and Communication.* London: Pergamon Press.

Brown, G. & Yule, G. (1983). *Teaching the Spoken Language.* Cambridge: CUP.

Calvo, M. G. (2001). Working memory and inferences: Evidence from eye fixations during reading. *Memory*. 9,365–381.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z.,Wilhelm, O., & Engle, R.W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*, 769–786.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62–101). New York: Cambridge University Press.

Cowan, N. (2005). *Working memory capacity*. Hove, UK: Psychology Press.

Cowan, N. (2011).Working memory and attention in language use. In J. Guendouzi, F. Loncke, & M. J. Williams (Eds.), *The handbook of psycholinguistic and cognitive processes: perspectives in communication disorders* (pp. 75-97). New York, NY: Psychology Press.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466. 134.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(4), 561–584.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. *Journal of Memory and Language, 25,* 1–18.

Daneman, M., & Hannon, B. (2001). Using working memory theory to investigate the construct validity of multiple-choice reading comprehension tests such as the SAT. *Journal of Experimental Psychology: General, 130,* 208-223.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: meta-analysis. *Psychonomic Bulletin & Review, 3,* 422–433.

Dehn, M. J. (2008). *Working memory and academic learning : assessment and intervention*. Hoboken, NJ: John Wiley & Sons.

Convirs, S.M. (1994).*The role of working memory in reading comprehension.* Unpublished doctoral dissertation. University of California

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal behavior,* 11, 671-684.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 972–992.

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two factor theory of cognitive control. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol.44, pp.145-199). New York, NY: Elsevier.

Engle, R. W., Kane M. J., & Tuholski S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: mechanisms of active maintenance and executive control* (pp. 102–134). Cambridge: Cambridge University Press.

Felser, C. & L. Roberts (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research* 23, 9-36.

Feng, G. (2009). Mixed Responses: Why Readers Spend Less Time at Unfavorable Landing Positions. *Journal of Eye Movement Research*, 3(2):3, 1-26.

Florit, E., Roch, M, Altoè, G, & Levorato, M.C. (2009). Listening comprehension in preschoolers: The role of memory. *British Journal of Developmental Psychology*, 27 (4), 935-951.

Flowerdew, J., & Miller, L. (2005). *Second language listening: Theory and practice*. Cambridge: Cambridge University Press.

Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language, 51,* 136–158.

Futransky J. S. (1992). *Relations among verbal working memory, listening comprehension, and reading skills.* Unpublished doctoral dissertation. University of Rhode Island.

Garrod, S., Freudenthal, S., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language*, 33, 39–68.

Gathercole, S.E., Lamont, E., & Alloway, T.P. (2006). Working memory in the classroom. In S. Pickering (Ed.). *Working memory and education,* pp. 219-240. Elsevier Press.

Harrington, M. W. & Sawyer, M. (1992). L2 working memory capacity and L2 reading skills. *Studies in Second Language Acquisition*. 14.1, 25–38.

Hulme, C. & Mackenzie, S. (1992). *Working memory and severe learning difficulties.* Hove: Lawrence Erlbaum Associates Publishers, Hove.

James, W. (1890), The Principles of Psychology, 2 vols, New York: Henry. Holt.

Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society* 102.2, 199–225.

Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research,* 21.2, 121–151.

Juffs, A. (2006a). Processing reduced relative vs.main verb ambiguity in English as a Second Language: A replication study with working memory. In R. Slabakova, S. Montrul & P. Prevost (eds.), *Inquiries in linguistic development in honor of Lydia White*. Amsterdam: John Benjamins,213–232.

Juffs, A. & Harrington, M. W. (2011) Aspects of working memory in L2 learning. *Language Teaching*, 44 2: 137-166.

Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working-memory capacity: A latent variable approach to verbal and visuo-spatial memory span and reasoning. *Journal of Experimental Psychology*, 133, 189_217.

Kaushanskaya, M. & Yoo, J. (2012). Phonological memory in bilinguals and monolinguals. *Memory and Cognition.* 40(8):1314-30.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, England: Cambridge University Press.

Kormos, J. & A. Sáfár (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. Bilingualism: Language and Cognition 11(2), 261–271.

Kosaka, K & Yamazaki, A. (2000). Working memory capacity and text comprehension in young children", *JPN J EDU P*, 48(3), 2000, pp. 343-351

La Pointe, L. B., & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1118–1133.

Leeser, M.J. (2007). Learner-based factors in L2 reading comprehension and processing grammatical form: topic familiarity and working memory. *Language Learning, 57,* 229–270.

Londe Z. C. (2008). *Working memory and English as a second language listening comprehension tests: A latent variable approach*. Unpublished doctoral dissertation. University of California.

Mackey, A., Adams, R., Stafford, C. & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning,* 60.3, 501–533.

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

McDougall, S.,Hulme, C., Ellis,A., & Monk, A.(1994). Learning to read: The role of short-term memory and phonological skills. *Journal of Experimental Child* Psychology. 58, 112-133.

Miyake, A. & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy & L. E. Bourne (eds.), *Foreign language learning: Psycholinguistic studies on retention and training*. Mahwah, NJ: Lawrence Erlbaum, 339–364.

Miyake, A. & Shah, P. (1999). Toward unified theories of working memory. In A. Miyake & P. Shah (eds.), *Models of memory: Mechanisms of active maintenance and recall*. Cambridge: Cambridge University Press. 442–481.

M. M. Gruneberg, P. E. Morris, & R. N. Sykes (1994). *Practical aspects of memory.* London: Academic Press.

Morgan-Short, K., Finger, I., Grey, S., & Ullman, M. T. (2012). Second language processing shows increased native-like neural responses after months of no exposure. *PLoS ONE*, 7(3): e32974.

Moussa-Inaty J; Ayres P; Sweller J, (2012) 'Improving Listening Skills in English as a Foreign Language by Reading Rather than Listening: A Cognitive Load Perspective. *Applied Cognitive Psychology,* 26, 391 - 402

Osaka, M. & N. Osaka (1992). Language independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, 30, 287–289.

Osaka, M., Osaka, N. & Groner, R. (1993). Language independent working memory as measured by Japanese and English reading span tests. *Bulletin of the Psychonomic Society*, 31, 117–118.

Rai, M.K., Loschky, L.C., Harris, R.J., Peck, N.R., & *Cook, L. (2011). The effects of stress and working memory capacity on foreign language readers' inference comprehension. *Language Learning,* 61(1), 187-218.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., et al. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*, 28, 164–171.

Roberts, L. (2012). Individual Differences in Second Language Sentence Processing. *Language Learning.* 62, 72–188.

Sunderman, G. & Kroll, J.F. (2009).When study-abroad experience fails to deliver: The internal resources threshold effect. *Applied Psycholinguistics.* 30, 79–99.

Tsuchihira, T. (2007). L2 Working memory capacity and L2 listening test scores of Japanese junior college students. *Journal of Bunkyo Gakuin University*, 7, 159-175.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language*, *28*, 127–154.

Ullman, M. T. (2001a). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, 30(1), 37–69.

Ullman, M. T. (2001b). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, 4(1), 105–122.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92(1–2), 231–270.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory and practice* (pp. 141–178). Washington, DC: Georgetown University Press.

Unsworth, N., & Engle, R.W. (2007). On the division of short-term and working memory: An examination of simple and complex spans and their relation to higher-order abilities. *Psychological Bulletin, 133,* 1038-1066.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37,* 498-505.

Van den Noort, M., Bosch, P., & Hugdahl, K. (2006). Foreign language proficiency and working memory capacity. *European Psychologist, 11*, 289–296.

Van den Noort, M.W.M.L., Bosch, M.P.C., Haverkort, M., & Hugdahl, K. (2008). A standard computerised version of the reading span test in different languages. *European Journal of Psychological Assessment*, 24(1), 35-42.

Walter, C. (2004). Transfer of reading comprehension skills to L2 is linked to mental representations of text and to L2 working memory. *Applied Linguistics, 25(3),* 315–339.

Waters, G. S., & Caplan, D. (1996). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology, 49A*(1), 51- 79.