T.C.

KAFKAS UNIVERSITY

THE INSTITUTE OF SOCIAL SCIENCES

THE DEPARTMENT OF ENGLISH LANGUAGE AND LITERATURE

THE EFFECTS OF AGE AND PRODUCTION MODE ON THE NATIVE AND FOREIGN LANGUAGE SERIAL RECALL PERFORMANCE OF TURKISH STUDENTS OF ENGLISH

A THESIS FOR THE DEGREE OF MASTER

By Catherine AKÇA

Supervisor: Dr. Gencer ELKILIÇ

KARS, 2009

T.C.

KAFKAS UNIVERSITY

THE INSTITUTE OF SOCIAL SCIENCES

THE DEPARTMENT OF ENGLISH LANGUAGE AND LITERATURE

THE EFFECTS OF AGE AND PRODUCTION MODE ON THE NATIVE AND FOREIGN LANGUAGE SERIAL RECALL PERFORMANCE OF TURKISH STUDENTS OF ENGLISH

A THESIS FOR THE DEGREE OF MASTER

By Catherine AKÇA

Supervisor: Dr. Gencer ELKILIÇ

KARS, 2009

T.C.
KAFKAS ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜ'NE

Catherine AKÇA' ya ait "The Effects of Age and Production Mode on the Native and Foreign Language Serial Recall Performance of Turkish Students of English" konulu çalışma, jürimiz tarafından İngiliz Dili ve Edebiyatı Anabilim Dalı Yüksek Lisans tezi olarak oybirliğiyle kabul edilmiştir.

Öğretim Üyesinin Ünvanı, Adı ve Soyadı	İmza
Yrd. Doç. Dr. Gencer ELKILIÇ (Danışman)	……
Yrd. Doç. Dr. Erdinç PARLAK	……
Yrd. Doç. Dr. Bilal GENÇ	……

Bu tezin kabulü Sosyal Bilimler Enstitüsü Yönetim Kurulunun
….. / ….. / 200 tarih ve …… / …… sayılı kararı ile onaylanmıştır.

UYGUNDUR

……. / ……. / …….

Sosyal Bilimler Enstitüsü Müdürü

TABLE OF CONTENTS

ÖZET

Tezin Çeşidi:	Yüksek Lisans Tezi
Tezin Adı:	İngilizce Öğrenen Türk Öğrencilerin Anadil ve Yabancı Dil Seri Hatırlama Performansları Üzerine Yaşın ve Dil Kullanma Modlarının Etkisi
Tezi Hazırlayanın Adı:	Catherine AKÇA
Danışman:	Yrd. Doç. Dr. Gencer ELKILIÇ
Tezin Sunulduğu Yıl:	2009
Sayfa Sayısı:	82

İşler bellek sistemi için kullanılabilir bilişsel kaynaklar sınırlıdır. Bilişsel aşırı yükleme durumlarında geçici bilgi depolama ile bilgi işlem yapma talepleri arasında bir uzlaşma yapmaya gereksinim duyulabilir. Çocuklarda henüz tam olarak otomize olmamış grafik ve ortografik işlemler de ayrıca ilaveten bilişsel kaynaklar tüketebilir. Bu yüzden çocukların aynı görevi sözlü olarak yazılıdan daha iyi yapması beklenebilir.

Bu durumun Türk çocuklarında geçerli olup olmadığını ve yaşın hatırlama modu ile istatistiksel olarak anlamlı düzeylerde etkileşip etkileşmediğini belirlemek için 16'sı lisans (ortalama yaş: 19 yıl, 6 ay), 16'sı ilkokul beşinci sınıf (ortalama yaş: 11 yıl, 0 ay) ve 16'sı ilkokul üçüncü sınıf öğrencilerinden (ortalama yaş: 8 yıl, 11 ay) oluşan, anadil olarak Türkçe konuşan toplam 48 denek üç gruba ayrıldı. Katılımcıların, kendilerine verilen, sık kullanılan Türkçe kelime listelerini sırasıyla hatırlamaları istendi. Hatırlama hem sözlü ve hem de yazılı olarak yapıldı. Seri hatırlama verilerinin analizi hatırlama modunun yaş ile istatistiksel olarak anlamlı düzeylerde etkileşmediğini, fakat yetişkinler, çocukların tamamı ve yalnızca beşinci sınıf çocuklarında yazılı hatırlamanın sözlü hatırlamaya kıyasla istatistiksel olarak daha iyi olduğunu gösterdi. Bu bulgu Türkçe ortografinin şeffaflığına ve/veya katılımcı çocukların yaşlarına bağlanabilir.

Ayrıca, düzensiz ortografik sistemli yabancı dil olan İngilizce'de hazırlanmış seri hatırlama testi aynı katılımcılara verildi. Yine, yetişkinler yazılı modunda istatistiksel olarak anlamlı ölçülerde daha iyi performans sergilediler. Fakat çocuklarda hiç bir mod etki gözlenmedi. Çocuklarda anadil testinde tespit edilen yazılı hatırlama üstünlük etkisi burada gözlenmedi. Bu durum daha az otomize bir dili kullanmanın getirdiği ilave bilişsel yüke ortografik ve grafomotor yüklerin de eklenmesiyle bilgi depolamak için çocukların işler belleğindeki kullanılabilir rezervin azaltıldığını göstermektedir.

Elde edilen bulgular Türk eğitim sisteminde ilköğretim okullarında İngilizce'nin öğretimi bağlamında tartışıldı.

Anahtar Kelimeler: Bilişsel yük, yabancı dil, sözlü ve yazılı dil kullanım modu, seri hatırlama, işler bellek.

ABSTRACT

Type of Thesis: Master's Degree Thesis
Title: The Effects of Age and Production Mode on the Native and Foreign Language Serial Recall Performance of Turkish Students of English.
Author: Catherine AKÇA
Supervisor: Yrd. Doç. Dr. Gencer ELKILIÇ
Date: 2009
Number of Pages: 82

The cognitive resources available to the working memory system are finite. Under conditions of cognitive overload, a compromise may need to be made between the demands of storage and processing. In children, graphic and orthographic processes which are not yet fully automated may consume additional cognitive resources. Thus, children may be expected to perform the same task better orally than in writing.

In order to determine whether this is the case in Turkish children, and whether age would interact significantly with recall mode, 48 native speakers of Turkish were assigned to three groups, comprising 16 undergraduate students (mean age: 19 years, 6 months), 16 fifth grade primary school pupils (mean age: 11 years, 0 months) and 16 third grade pupils (mean age: 8 years, 11 months). The participants were required to recall lists of frequently used Turkish words in the serial order in which they were presented. Recall was performed both orally and in writing. Analysis of the serial recall data showed that recall mode did not interact significantly with age, but that in the adult sample, the combined samples of schoolchildren, and the fifth grade sample of children alone written recall was significantly better than oral recall. This finding may be attributable to the transparency of Turkish orthography and/or to the age of the participating schoolchildren.

In a separate trial, the serial recall test was administered to the same samples of subjects in English, a foreign language with an irregular orthographic system. Again, adults performed significantly better in the written mode. However, no mode effect was observed in children. The written recall superiority effect found in children in the native language test was not observed here. This suggests that when the additional cognitive costs of processing a less automated language are combined with orthographic and graphomotor transcription costs, the resources available to children in working memory for the storage of information are depleted.

The findings are discussed in the context of the Turkish education system, with reference to the teaching of English in primary schools.

Key Words: cognitive load, foreign language, oral and written production modes, serial recall, working memory.

FOREWORD

This Master's Degree thesis was written in the Department of English Language and Literature of the Institute of Social Sciences at Kafkas University.

The serial recall performance of French and German primary school children, but not of adults, for wordlists, has been found to be significantly better in the oral recall mode than in the written mode. This has been taken as evidence that the cognitive cost of processes such as spelling and handwriting, which have not yet been fully automated, adversely affects written language production in young children (Grabowski, 2005; Bourdin & Fayol, 1994). The current study sought to determine whether a similar interaction between age and recall mode would appear in Turkish.

The study also extended the scope of the research by repeating the experiment in English, with the same samples of participants, in order to determine what effects would emerge, in various age groups, when the test was conducted in a less automated foreign language, with a more irregular spelling system. A further aim of the study was to address the pedagogical implications of the findings, with specific reference to the teaching of English as a foreign language in Turkish primary schools.

I should like to thank my supervisor, Assistant Professor Dr. Gencer Elkılıç, for his invaluable assistance and guidance during the preparation of this thesis. I am also grateful to the authorities at Kafkas University and K.A.K.Ü.V. Private Primary School in Kars for granting me permission to undertake the experimental work required in the study. Many thanks too to all the students who gave up their precious time in order to participate in the research.

Finally, I should like to dedicate this work with all my love to my husband Atila and to our children Altay and Meryem, without whose patience and encouragement it could never have been completed.

Catherine AKÇA,
KARS, 2009

LIST OF ABBREVIATIONS

ANOVA:	analysis of variance
EFL:	English as a foreign language
FL:	foreign language
LTM:	long-term memory
M:	mean
NL:	native language
PSTM:	phonological short-term memory
SAS:	supervisory activating system
SE:	standard error
STM:	short-term memory
VSTM:	verbal short-term memory
WM:	working memory

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Introduction to the Study

The faculty of memory enables human beings to define and direct their existence. Without the capacity to store information over time and recollect it when required, it would be impossible to link the present with the past. Likewise, plans for the future are built upon the foundations of acquired and remembered knowledge. On a more immediate basis, all forms of cognitive activity and their behavioural outcomes depend upon the ability to store information temporarily in memory, while it is integrated with prior knowledge and processed in order to achieve a particular objective (Jarrold & Towse, 2006). The system which manages cognitive processing, including the short-term storage of new information and its integration with longer-term memory representations, has been defined as working memory (WM). The WM system is fundamental to all learning processes (Baddeley, 2003a; Baddeley, 1999).

It is principally through the medium of language that human beings acquire and transmit knowledge. Language itself relies upon memory. The ability to understand and to process language depends upon the ability to hold sequences of sounds temporarily in WM, while they are made meaningful (Gupta & MacWhinney, 1997). The theoretical model of WM proposed by Baddeley and Hitch comprises several components, one of which is the phonological loop (Baddeley, 1999). During cognitive processing, the phonological loop maintains sequences of sound in temporary mental storage. By means of a process of subvocal rearticulation, known as rehearsal, the phonological loop delays the loss of these sound traces from memory (Baddeley, 2003a). Repeated exposure to the same sequences of sound, accompanied by rehearsal, results in their eventual consolidation into long-term

memory (LTM) (Gathercole, 2006; Gupta & MacWhinney, 1997). Baddeley has therefore argued that the phonological loop evolved as a language learning device (2003a). A considerable amount of evidence has accumulated in support of this hypothesis, as a result of research carried out in a variety of populations. In particular, phonological short-term memory (PSTM) skills have been found to be associated with vocabulary development, in both native (NL) and foreign language (FL) contexts (Baddeley, 2003b). The strength of the association has been found to be most significant in young children, suggesting that phonological memory plays a particularly important role in the initial stages of language learning (Gathercole, 2006).

As vocabulary is acquired, it is used in oral and written communication. Liberman has observed that there are fundamental differences between spoken and written language production. Speech is a biological characteristic of human beings, whereas written language production depends upon education. Speech is acquired through exposure to a native language, whereas children are taught to read and write at school. Both speech perception and speech production are mediated phonologically, that is through sound. In contrast, written language production is less direct. Written texts must be encoded orthographically for output. In other words, written language production is mediated through an arbitrary system of alphabetic symbols or of characters, which must be learned. It is, therefore, the more difficult form of communication (Liberman, 1998).

The WM system is believed to be of limited capacity. Its available resources must be allocated between the functions of storing and processing information, and may be switched between the two. If a task places excessive demands upon the system, it is believed that a trade-off will be made between the two functions, as a result of which some information may be lost from memory and/or the quality of performance in the task may deteriorate (Gathercole, 1999). Writing is a complex cognitive activity, which involves a variety of processes (Hayes, 1996). Transcription processes, such as handwriting and spelling, are performed almost automatically once mastered, and are therefore considered to be low-level writing processes. On the other hand, the processes involved in planning, organising and reviewing the conceptual and linguistic content of a text require considerable mental resources and

have been defined as high-level writing processes (Kellogg, 2001; Kellogg, 1996; Bourdin & Fayol, 1994). However, in young children in whom the processes of handwriting and spelling have not yet been fully mastered, the orthographic-motor execution of a text may be so demanding that few resources remain in WM for higher level writing processes (Medwell & Wray, 2007; Bereiter, 1980). Thus, it is only after transcription processes have been at least partially automated that young writers start to divert cognitive resources towards improving the quality of their written language production (McCutchen, 1996).

Evidence for such a capacity theory of writing development has emerged from research which has focussed upon the cognitive cost of low-level writing processes during the childhood years (Grabowski, 2005; Bourdin & Fayol, 2000; 1996; 1994). In a series of experiments conducted with French speakers in their NL, the ability of children, but not of adults, to recall lists of words in their original serial order was found to be significantly poorer in the written recall mode than in the oral recall mode (Bourdin & Fayol, 1994). In other words, the written mode involves an additional cognitive load which adversely affects language production in children, but which becomes less significant with development. Similar results have also been obtained with German speakers in their NL, prompting the conclusion that written tasks should not be used in isolation if a comprehensive assessment of the knowledge and cognitive abilities of young schoolchildren is to be made (Grabowski, 2005).

Grabowski has recommended that cross-linguistic studies should be undertaken in order to establish whether the interaction between age and recall mode found in the French and German serial recall experiments is of universal significance, and whether it therefore represents a characteristic feature of cognitive development (2005). The current study, therefore, sought to determine whether Turkish schoolchildren would experience the same difficulty in the written serial recall mode as their French and German counterparts. If so, this would lend further weight to Bourdin and Fayol's thesis that the cognitive cost of low-level graphic and orthographic processes impairs language production in the written mode in children (1994).

A further, and distinctive, aim of this study was to compare the oral and written serial recall performance of the participants in their native Turkish with their

performance in English. As a FL, English should be a less automated language for these subjects than Turkish (Abutalebi & Green, 2007). Moreover, English has an extremely opaque orthography, with many irregularities in its spelling system (Landerl, 2005). The study, therefore, sought to establish what effects the presumed additional cognitive costs of performing the serial recall test in English would produce in the various samples of participants, in both the oral and the written output modes.

Working memory is a system of limited capacity, which is subject to developmental change during the childhood years. Moreover, there are differences in WM capacity between individuals (Gathercole, 1999). In a general educational context, researchers have recommended that teachers should avoid overstretching the capacity of young learners to store and process information. If a task makes excessive demands of the WM system, failure may occur. Repeated failure will undermine educational progress (Gathercole & Alloway, 2008; Gathercole *et al.*, 2006). In the light of such recommendations, the final objective of this study was to focus on the pedagogical implications of the current research, with specific reference to the teaching of English as a foreign language (EFL) in Turkish primary schools.

Pupils in the Turkish state school system start to study EFL in the fourth grade, at the age of around nine. In private schools, English forms a part of the curriculum from the first grade onwards, and is often introduced in the nursery class. Given that the phonological loop has been shown to be significantly associated with the development of FL vocabulary knowledge in young children (Masoura & Gathercole, 1999), it is clearly important to maximise its contribution to the learning process during the early school years. If the outcome of the current research lends support to the argument that the cognitive cost of written language production interferes with the short-term storage of information in memory in primary school children, then questions arise as to the balance which should be set between the skills of writing and speaking when teaching EFL to various age groups, and as to the way in which associated learning activities should be structured.

1.2 Literature Review

This Section explores the literature which underpins the current study. The theoretical construct of WM is described, with particular reference to its cognitive structure, its limited capacity and to developmental changes which affect it. Tests of WM are discussed in brief. Studies which have examined the role of the phonological loop component of WM in native and foreign language acquisition are summarised. A model of written language production, developed in the context of limited capacity accounts of WM, is outlined. Finally, an overview is given of the research conducted within this cognitive framework in order to test the hypothesis that the cognitive cost of written production impairs the serial recall performance of children (Grabowski, 2005; Bourdin & Fayol 2002; 2000; 1996; 1994).

1.2.1 Working Memory

Memory may be defined as the faculty which enables human beings to register information, store it over time, and retrieve it when required (Baddeley, 2004).

Memory was originally conceived of as a unitary system. During the 1960s, however, the theory that memory might be fractionated into separate long and short-term systems began to gain currency (Baddeley, 2004).

In 1960, the term "working memory" was applied by Miller and his colleagues to their theoretical construct of a memory system dedicated to short-term storage and to the executive control of cognition (Richardson, 2007). In 1974, Baddeley and Hitch, questioning the validity of earlier unitary models of short-term memory (STM), used Miller's terminology to denote their new three component model of WM (Baddeley, 2002).

Baddeley and Hitch's multi-component model of WM has not been universally accepted. For instance, Daneman, Carpenter and Just have described WM as a unitary system in which a limited memory capacity resource may be directed to either the storage or the processing of information (Gathercole, 1999). Engle, on the

other hand, has explained WM in terms of "executive attention", that is as a domain-free, limited capacity to inhibit potentially distracting interference during cognitive processing (Engle, 2002).

Whatever the lack of consensus regarding the structure of the WM system, the term WM itself is now widely used by researchers to refer to the human capacity to store information temporarily in the mind and to process it, despite interference, in order to achieve a particular cognitive goal (Jarrold & Towse, 2006). Thus, a distinction is generally made nowadays between STM and WM (Jarrold & Towse, 2006). Short-term memory refers to the ability to retain information in the mind over a limited period. On the other hand, the theoretical construct of WM posits a limited capacity system which involves not only this short-term storage of information but also its manipulation. The WM system is believed to enable interaction to take place between the products of perception and the LTM system (Baddeley, 2003a). Thus, by allowing a number of pieces of information to be held temporarily in the mind, and simultaneously manipulated and integrated, the WM system plays a critical role in the accomplishment of complex cognitive processes such as comprehension, reasoning and learning (Baddeley, 1999).

1.2.1.1 Components of Working Memory

Although the debate over the cognitive structure of WM has not yet been resolved, Baddeley and Hitch's multi-component model has long been influential in many areas of cognitive science, and has been validated by both behavioural and biological evidence (Baddeley, 2003a; Baddeley, 2000; Smith & Jonides, 1997). In terms of the current study, the Baddeley and Hitch model is of interest in that it provides the cognitive framework for ongoing psycholinguistic research into the acquisition of both native and foreign languages (Baddeley, 2003b).

Baddeley and Hitch's original model of WM comprised a limited capacity control system, the central executive, supported by two domain specific subsystems dedicated to storage: the phonological loop for speech-based information, and the visuospatial sketchpad for visual information (Baddeley, 2000).

The central executive was initially regarded as a general information processing resource. It was subsequently re-defined as a system able to control attention by one of two means. The routine control function manages habitual, automatised patterns of behaviour, stimulated by environmental triggers. Where routine control proves inadequate, the supervisory activating system (SAS) takes over. The central executive is believed to deploy its limited attentional resources between the storage, retrieval, processing and integration of information required in the accomplishment of all forms of human activity (Baddeley, 2003a; Gathercole, 1999; Baddeley, 1996).

The phonological loop is a slave system which serves the central executive in its information processing function by storing traces of sound in memory for several seconds until they decay. The duration of these traces may be prolonged by subvocal rearticulation, in a process known as rehearsal. However, since the storage system is of limited capacity and since rehearsal takes place in real time, increasing the load of items to be remembered renders it impossible to refresh memory traces in their entirety before irretrievable decay occurs (Baddeley, 2003a). This component of WM is of particular interest to this study in that a body of recent research, discussed below, has explored the hypothesis that the phonological loop evolved to enable the acquisition of language (Baddeley, 2003b).

Like the phonological loop, the visuospatial sketchpad is a domain-specific subsystem of WM. In this case, the slave system serves the central executive by storing the visual and spatial information required in cognitive processing. Its capacity is believed to be limited typically to 3 or 4 objects (Baddeley, 2003a). The function of the visuospatial sketchpad lies outside the scope of the current study.

Recently, Baddeley has added a fourth component to his original model: the "episodic buffer". Under the control of the central executive, the episodic buffer is assumed to allow interaction to take place between the domain-specific phonological and visuospatial subsystems, using a multi-dimensional processing code. The episodic buffer is also assumed to allow the WM subsystems to interact with the LTM system. Thus, on the one hand the buffer facilitates the construction of more durable memory stores, while on the other it allows information retrieved from LTM

in the form of conscious awareness to be manipulated in WM during thought processing (Baddeley, 2000).

1.2.1.2 Tests of Working Memory

A distinction may be made between simple tests of immediate memory and tests of complex WM. In simple memory span tasks, the subject is required to recall items which are held temporarily in STM, but are not processed. In complex span tasks, the capacity of the WM system to store and process information meaningfully is tested (Jarrold & Towse, 2006).

Verbal short-term memory (VSTM) capacity is usually tested by means of simple span tasks. Typically, these are tests of immediate serial recall, in which the subject is required to hold a series of items in memory and to reproduce them in their original serial order immediately after presentation has been completed. In 1956, Miller found a typical VSTM span of around 7 items in adults (Richardson, 2007). However, factors such as the degree of phonological similarity between the items to be recalled, and increases in word length have been found to affect span adversely. These phonological and word-length effects, respectively, indicate that storage in VSTM is phonologically mediated, and that STM span is influenced by the rate at which transient phonological memory traces may be refreshed by subvocal rehearsal (Baddeley, 2003a).

Short-term memory performance has been found to be assisted by LTM knowledge. For instance, superior levels of recall have been demonstrated for high frequency words compared to low frequency words (Gregg, Freedman, & Smith, 1989). Words which are linked semantically have been found to be recalled more easily than semantically unrelated words (Saint-Aubin & Poirier, 1999). Recall performance has also been found to be better for words in a native or dominant language rather than in a second language (Thorn *et al.*, 2002; Thorn & Gathercole, 1999). Likewise, higher levels of recall have been demonstrated for real words compared to invented "nonwords", through the lexicality effect; and for nonwords which contain high frequency rather than low frequency combinations of phonemes

(sounds), through the phonotactic frequency effect (Thorn *et al*., 2005; Gathercole *et al.*, 2001). In other words, through a process termed "redintegration", enduring memory representations retrieved from LTM may facilitate the reconstruction of decaying short-term memory traces during serial recall (Thorn *et al*., 2005).

In tasks designed to test complex WM span, the cognitive load born by the subject is higher than for simple span tasks. For instance, the archetypal test of complex WM is Daneman and Carpenter's reading span test. Here, subjects are required to read a series of unrelated sentences aloud, while at the same time trying to remember the last word of each sentence, which they must reproduce in serial order at the end of the test (Richardson, 2007). It is believed that due to the limited capacity of the WM system a trade-off must be made between the demands of storage and processing in order to accomplish this type of cognitively demanding task successfully (see Gathercole, 1999).

Thus, individual differences in WM performance reflect differences not only in the mental storage capacity and/or processing efficiency of individuals but also in their ability to manage the competing demands of storage and processing (for a review of work in this area, see Jarrold & Towse, 2006).

1.2.1.3 Aspects of the Development of Working Memory

Although there are significant differences in WM capacity between individuals, the WM system is also subject to predictable developmental changes (Gathercole, 2002; Gathercole, 1999).

In a study of children aged four to 15 years old, Gathercole and her colleagues established that the modular structure of WM, proposed by Baddeley and Hitch, is in place from at least age six, and that the functional capacity of each component increases over the school years to adolescence (Gathercole *et al.*, 2004, Baddeley, 2003a).

Typically, STM capacity increases two to threefold between the ages of four and 14, with a steep increase in memory performance up to age eight and a more gradual increase thereafter (Gathercole, 1999). In contrast, memory span has been

found to increase by only one item between the age of 13 and early adulthood (Fry and Hale, 2000).

This general pattern also holds true for the phonological component of short-term memory, with a variety of processes combining to produce a significant improvement in phonological storage capacity throughout childhood (Gathercole, 1999). These developmental changes include reduced rates of decay and improved phonological encoding, faster rates of articulation, the systematic use of rehearsal strategies after age seven, and increasingly efficient redintegration processes (Gathercole, 2002; Gathercole, 1999; Gathercole *et al.*, 1994).

Likewise, complex WM span increases over the childhood years. Thus, as children mature, they are able to accomplish more difficult cognitive tasks. This developmental improvement has been attributed to the faster and more efficient manipulation of information by older children, which enables the central executive to divert more of its limited resources towards storing the products of cognitive processing (for an overview, see Gathercole, 1999). Research has also demonstrated a greater capacity in older children to switch their attention between the demands of storage and of processing whilst engaged in cognitively complex tasks (Gavens & Barrouillet, 1994).

1.2.2 Working Memory and Language Learning

The WM system allows material to be stored temporarily in the mind and integrated with prior knowledge. Without memory, it would be impossible to understand, learn or manipulate information. Without memory for words and for order, it would be impossible to process language (Baddeley, 1999; Gupta & MacWhinney, 1997).

In order to comprehend and/or to produce spoken language, sequences of sound must be held in memory and made meaningful. In the Baddeley and Hitch model of WM, the phonological store allows the short-term storage of sequences of sound (Baddeley, 2003a). One view is that the process of subvocal rehearsal, which prolongs the duration of these phonological traces in STM, simultaneously promotes

the learning of language, since the rearticulation of phonological units combined into lexical or grammatical sequences causes them to be consolidated into LTM (see for instance, Ellis & Sinclair, 1996). Baddeley has therefore argued that the phonological loop component of WM evolved to facilitate the acquisition of language (Baddeley, 2003b).

1.2.2.1 The Phonological Loop and Vocabulary Acquisition

A substantial body of research has now contributed evidence in support of Baddeley's hypothesis that the biological function of the phonological loop is to enable language learning (Baddeley, 2003b).

In particular, a number of studies have shown that PSTM capacity is associated with vocabulary development. For instance, in a longitudinal study, Gathercole and Baddeley found that the ability of four year olds to repeat sets of unfamiliar nonwords in serial order predicted the extent of their native vocabulary knowledge at age five (cited in Baddeley, 2003b). Similarly, five year olds with good phonological memory skills, as measured by their ability to repeat lists of unfamiliar nonwords, were found to be faster at learning invented names for toys, and better at remembering them, than children with poor phonological memory capacity (Gathercole & Baddeley, 1990). Thus, the ability to reproduce unfamiliar sequences of sound correctly has been shown to be associated with the learning and retention of vocabulary in young children.

Furthermore, in a study involving 13 year old children, Gathercole and her colleagues have also established that the developmental association between phonological memory skills and native vocabulary knowledge persists into the early teenage years (Gathercole *et al.*, 1999). More recently, Gupta has shown that nonword repetition, serial recall and word learning abilities also correlate in adults (Gupta, 2003).

The association between phonological memory capacity and vocabulary development has also been found to apply in FL learning contexts, both in children and in adults. For instance, Service reported that the STM span of 9 year old Finnish

children for pseudowords that sounded like English words predicted their knowledge of English vocabulary two years later (cited in Gupta & MacWhinney, 1997). In a study of ten year old Greek children learning English as a FL, the STM performance of the subjects, measured by their ability to repeat nonwords accurately, was found to be significantly associated with both their native and their foreign vocabulary knowledge (Masoura & Gathercole, 1999). Likewise, in an experiment in which adult subjects were either required to repeat FL (Welsh) utterances aloud immediately after presentation, or prevented from using rehearsal, the group which used rehearsal was found to be better at acquiring, pronouncing, comprehending and translating the foreign words and phrases than the group in which rehearsal was suppressed (Ellis & Sinclair, 1996).

Although the existence of a developmental association between VSTM and vocabulary knowledge, in both native and foreign languages, has been widely recognised, the causal direction of the relationship remains open to question. Does PSTM capacity principally determine the extent of vocabulary growth? Or, is vocabulary knowledge itself the prime causal factor in the developmental increase in VSTM capacity?

Recent research suggests that the relationship is reciprocal. Gathercole has observed that every word is initially a "nonword" for the learner. The process of consolidating any new word into the long-term memory system begins with the attempt to repeat it (Gathercole, 2006). As discussed above, individuals with superior PSTM skills have been found to be better at learning new vocabulary than individuals with a more limited phonological memory resource. It is believed that the ability to form and rehearse accurate short-term representations of the sound structure of unfamiliar words enables the eventual formation of a stable representation of the new word in the course of repeated exposures to the item (Gathercole, 2006; Gupta & MacWhinney, 1997). The influence of this temporary phonological storage mechanism on vocabulary acquisition appears to be particularly important in the earlier stages of language learning (Gathercole, 2006). For instance, in a longitudinal study of children between the ages of four and eight, Gathercole and her colleagues found a highly significant correlation between nonword repetition and NL vocabulary scores at ages four, five and six respectively. On the other hand, by

age eight, the strength of the correlation had weakened considerably, although it remained statistically significant (cited in Gathercole, 2006). This finding suggests that factors other than temporary phonological storage capacity may play a role in facilitating the acquisition of NL vocabulary at various stages of development. For instance, reading may make an increasingly significant contribution to the rate of vocabulary growth as childhood progresses, compensating for basic limitations in PSTM capacity (Gathercole *et al.*, 1999). Moreover, as the size of the permanent lexicon expands over the childhood years, redintegration processes become increasingly efficient. Therefore, as the child develops, the reactivation of long-term phonological representations may begin to play an increasingly significant role in facilitating new word learning. For instance, Metsala has argued that NL vocabulary growth during childhood results in a shift away from holistic representations of words in STM towards a more segmental approach based upon phonemes, which in turn leads to greater flexibility in the arrangement of individual phonemes into new patterns, thus facilitating the learning of new words (cited in Gathercole, 1999).

In a FL learning context, Masoura and Gathercole found that the phonological memory performance of 11 year old Greek children, measured by their nonword repetition ability, was closely associated with their English vocabulary scores. In contrast, the rate at which the subjects were able to learn previously unfamiliar English words for a picture naming task was found to be entirely independent of their PSTM skills; rather, it was strongly related to the extent of the children's existing English vocabulary, described by the researchers as "substantial". The authors concluded that increasing familiarity with a language leads to a shift away from word learning mediated by the phonological loop, towards greater reliance on existing lexical knowledge - which might be phonological, semantic or conceptual - to support the acquisition of new vocabulary (Masoura & Gathercole, 2005). Elsewhere, in a study involving bilingual adults with variable levels of English-French proficiency, Majerus and his colleagues found that both serial order storage and rehearsal capacities and the extent of previous exposure to the phonological structure of a language influenced the rate of vocabulary acquisition of the participants. The researchers highlighted their finding that serial order STM performance predicted lexical learning even in bilingual participants with broad

phonological knowledge, and concluded that their results favoured an integrative model of word learning (Majerus *et al.*, 2008).

In sum, both short-term phonological storage and rehearsal mechanisms and long-term lexical knowledge contribute to vocabulary acquisition, although the degree of involvement of the latter factor appears to increase in relation to the degree of familiarity with a language. For instance, Thorn and Gathercole tested STM performance in bilingual and monolingual children with differing levels of exposure to French and/or English. The researchers found that the phonological loop functioned in a highly language-specific way, with the participants displaying better levels of recall for lexical and nonword stimuli which were typical of their dominant language(s) (Thorn & Gathercole, 1999). It follows that the phonological loop may play a more enduring role in the process of acquiring FL vocabulary, since in this context long-term lexical knowledge is less likely to be available to support the learning of new words than is the case in NL acquisition (Masoura & Gathercole, 1999).

1.2.3 Working Memory and Language Production

1.2.3.1 Speech and Writing

As new language is learned, it becomes available for communication. Language may be communicated by speech or in writing. In practical and general terms, speech and writing differ in respect of their social and functional applications (Grabowski, 1996). However, regardless of pragmatic considerations, each of these methods of communication involves the participants, to various degrees, in the cognitive processes of perception, comprehension, planning, execution or review (Kellogg, 1996). The key difference, in cognitive terms, between speech and writing - as between listening and reading - is one of processing. Research indicates that in a normal population written language, like speech, involves the temporary storage of phonological representations in verbal WM (Kellogg *et al.*, 2007). However, whereas speech is by nature phonetic and therefore immediately processable, written language

requires further cognitive manipulation to make it so. In other words, before it can be processed cognitively, alphabetic script must be encoded phonologically, as inner sound. Likewise, for a planned text to be output in written form, it must be encoded orthographically (Kellogg, 1996). Writing has therefore been described as a less natural and more difficult way of communicating than speech (Liberman, 1998).

1.2.3.2 A Limited Capacity Theory of Writing Development

In 1980, Hayes and Flower proposed a model of writing which involved three principal cognitive processes: planning, translating and revision (Hayes, 1996). In planning, ideas are generated and organised. Translation is the process whereby this conceptual content is encoded grammatically and orthographically into sentences, which are then transcribed as text. Revision involves monitoring, correcting and improving the plan or text (Kellogg, 2001; Kellogg, 1996; Hayes, 1996).

In other words, writing is a complex cognitive task. The planning and organisation of the conceptual content of a text, its linguistic realisation and its successful motor execution rely upon the storage and processing capacities of WM.

The executive capacity of WM is believed to be limited. Thus, if the cognitive load placed upon the system by a task exceeds its executive resources, a trade-off will occur between the demands of storage and processing, resulting in a deterioration in performance (Baddeley, 2003a; Baddeley, 1996). In a probed study of college students, who were required to write narrative, descriptive or persuasive texts in longhand or on a word processor, Kellogg established that the planning, translating and reviewing processes competed with each other for such a common, general-purpose resource of WM. The lower planning demands of the more highly-practised narrative composition form, and the lower execution cost of writing by longhand each released capacity which was then redirected towards other writing processes (Kellogg, 2001).

Kellogg's study provided further evidence that the motor execution of written text makes negligible demands upon the capacity of WM in adults, for whom handwriting has become automated after years of practise (Kellogg, 2001; Bourdin &

Fayol, 1994). As a consequence of the automation of transcription processes, the practised writer is able to divert WM resources towards the processes of conceptualizing, formulating, and reviewing text, and is likely to produce higher quality output. In young writers, on the other hand, planning and reviewing skills begin to develop only after the cognitive demands imposed by the transcription process have become manageable (McCutchen, 1996; Bereiter, 1980).

McCutchen has argued that the findings from research on transcription processes in young children provide compelling evidence for a capacity theory of writing development (McCutchen, 1996). Transcription means forming and physically executing written representations of text, and involves spelling and handwriting processes which young children must learn (McCutchen, 1996). During spelling, order and identity information about letters is stored in WM while output is programmed (Service & Turpeinen, 2001). Young children, who are still in the process of learning how to spell, construct spellings "online" in WM, drawing upon their knowledge of phonology, orthography and morphology; whereas, with practise, retrieval of spellings becomes automated in older children and adults (McCutchen, 1996). Similarly, the process of physically writing text demands conscious attention in young children. Research indicates that in young writers, the orthographic-motor integration of handwriting, which is generally considered to be a low-level process, may make place such a load upon WM that relatively few resources remain for the high-level processes of conceptualization, linguistic planning and review which determine the quality of written production (Medwell & Wray, 2007). Furthermore, there is evidence that the attentional demands which the processes of spelling and handwriting make upon children's WM capacity may affect storage to the extent that material which has been formulated in WM may be forgotten before it can be output. Bereiter and his colleagues, who compared children's forecasts of what they were about to write with what they actually wrote, estimated that memory failures resulted in some content loss from one in every ten phrases written by young children (cited in McCutchen, 1996). Likewise, King and Rental found that both the length and the standard of texts produced by six to eight year old children increased when they dictated rather than wrote the material themselves (cited in Bourdin & Fayol, 2000).

Similarly, young children have been found to produce better narratives orally than in written form (Applebee, cited in Bourdin & Fayol, 1994).

However, with development, transcription processes start to become automatised. Therefore, by around the age of nine or ten, children are able to divert cognitive resources towards higher level writing processes, with a consequent improvement in the quality of their written language production (McCutchen, 1996).

1.2.3.3 Mode Effects in Language Production

Fundamental research into the impact transcription processes have upon language production in the childhood years was conducted by Bourdin and Fayol. The participants in their study were asked to recall lists of words in their native French language, in serial order, either orally or in writing. Thus, the task involved the temporary storage of item and order information in WM, and its output in phonetic or graphic form. Since the word lists required neither conceptual input nor linguistic planning from the participants, the task was considered to correspond to a low-level process: the maintenance of a pre-planned message in WM during its execution. The participants in the initial study were two groups of schoolchildren, aged seven or nine years old, and a group of adult students. The researchers hypothesized that the serial recall performance of the children would be better in the oral mode than in the written mode. They speculated that the cognitive load imposed by transcription processes which had not yet been fully mastered would limit the amount of resources available to the schoolchildren for the retention of word lists in memory. As they had anticipated, Bourdin and Fayol found that the serial recall performance of children was significantly better in the oral mode than in the written mode. Moreover, this finding applied to both groups of schoolchildren, regardless of age. In contrast, there was no significant effect of mode on the recall performance of the adult group (Bourdin & Fayol, 1994).

Bourdin and Fayol then conducted a series of control experiments in order to clarify their findings. By limiting the pace of oral recall to that of written recall, the researchers established that the slowness of writing compared to speaking did not

account for their initial results. The oral recall performance of schoolchildren remained significantly better than their written recall performance, even under these fixed rate recall conditions. In further control experiments, the researchers established that orthographic difficulties and the motor execution of text accounted, at least in part, for their initial findings. Thus, the recall performance of the children improved to a level comparable to that in the oral mode when they dictated their recall responses for transcription by the experimenter, that is when difficulties associated with handwriting and spelling were eliminated. Interestingly, when adults were asked to use an unusual form of writing, cursive upper-case longhand, their serial recall performance deteriorated significantly in relation to their performance in both the oral and normal longhand modes. In other words, even in educated adults, an unfamiliar mode of transcription may drain limited WM resources, leading to a decline in the quality of output. Similarly, when adults were presented with irregularly spelt and/or unfamiliar words, their written recall performance was significantly worse than their oral performance, suggesting that cognitive resources were diverted away from other cognitive processes in order to cope with orthographic difficulties. In children too, the level of recall for unfamiliar words was significantly higher in the oral mode than in the written mode. The overall level of recall by schoolchildren was significantly better for regularly spelt words than for irregular words. However, contrary to expectations, it was not significantly better in the oral mode than in the written mode, under this condition. The researchers speculated that the eight year old subjects may simply have ignored spelling irregularities when writing these more complex words, and recommended that further studies should be undertaken to clarify this problem (Bourdin & Fayol, 1994).

This series of experiments demonstrated that the cognitive cost of low-level transcription activities, which have not been fully automated, may adversely affect the capacity of the central executive system to maintain items in WM and to retrieve them during serial recall, and by extension may limit its capacity to manage higher level activities such as planning and review during writing. In particular, the poorer level of recall displayed by schoolchildren in the written mode, when compared to their level of spoken recall for the same task, suggested that output mode may be a determining factor in the allocation of resources between cognitive processes during

language production in the developmental years, with significant implications for the quality of written output.

In a subsequent study, Bourdin and Fayol demonstrated that in young children non-automated writing processes are equivalent to a secondary task which demands additional cognitive resources, over and above those required in oral language production. Two groups of schoolchildren were asked to complete a serial recall task under five different conditions: oral recall, written recall, or oral recall combined with a concurrent tapping, drawing or sound categorisation task. The tapping task was not found to tax WM resources to any significant degree in either group of participants. Drawing represented a secondary graphomotor execution task. Sound categorisation represented a secondary planning task, assumed to consume attentional resources in WM. In seven year old children, recall in the oral mode was found to be significantly better than written recall and than oral recall with drawing or sound categorisation. In other words, as cognitive resources were diverted towards the secondary tasks of spelling, handwriting, drawing or categorising sounds, fewer resources remained available in WM for the storage and retrieval of words, and performance deteriorated. In nine year old children, oral recall was found to be significantly better than written recall and than oral recall with sound categorisation. Thus, the graphomotor activity of drawing did not incur an additional cognitive cost in this group. The researchers suggested that the lack of a significant difference in performance between oral recall alone and oral recall with drawing in the nine year old children might indicate that by this age handwriting has been at least partially automated. They further speculated that in this age group the superiority of the oral mode over the written mode might, therefore, be due to spelling rather than graphomotor difficulties (Bourdin & Fayol, 2000).

This study lent weight to Bourdin and Fayol's hypothesis that the written mode involves an additional cognitive load which adversely affects language production in children, but which becomes less significant with development.

Bourdin and Fayol further investigated whether their finding that output mode interacted with age could be extended from the simple serial recall of word lists to the production of more complex language in the form of sentences and texts. Children and adults were given a speaking span test and a writing span test, in which

they were required to remember lists of words and then to produce a meaningful sentence for each word in the order in which it had originally been presented. Output was required in either the written or the oral mode. The speaking span performance of seven year old children was found to be significantly higher than their writing span, whereas for adults this position was reversed. No significant difference was found between the oral and written spans of nine year old children, indicating that writing processes which had begun to be automated by this age demanded fewer cognitive resources than in the younger group. Nonetheless, at the longest word list length, the oral span of this group was found to be significantly better than its writing span, suggesting that as development proceeds the cognitive resources consumed by low-level writing processes become significant only when WM is pushed to the limits of its capacity (Bourdin & Fayol, 1996). Moreover, Bourdin and Fayol have recently demonstrated that even in adults the quality of written output may be affected under conditions of cognitive overload. In a text production span test, adult participants were required to remember lists of words and then to generate texts using the words in the set, in the written or oral mode. The researchers found that the quality of the texts produced in writing was significantly lower than that of the texts produced orally, but only when the words in the sets were not semantically linked. Bourdin and Fayol concluded that when the cognitive cost of maintaining unrelated elements in WM and organising them into a coherent text was added to the cost of managing its written output, the central executive was left with insufficient residual capacity to meet the level of quality attained when the same task was performed in the oral mode. In other words, even in adults the cost of managing the mechanics of written production may not always be insignificant (Bourdin & Fayol, 2002).

The cognitive framework within which Bourdin and Fayol carried out their studies presupposes a WM system of limited capacity, and is assumed to be of universal application (Baddeley, 2003a). However, it is possible that differences between languages may influence the degree to which limited cognitive resources are consumed by low-level processes in developing writers. On this basis, the researcher Grabowski recommended that cross-linguistic studies should be undertaken in order to substantiate the hypothesis that orthographic and graphomotor processes which have not been fully automated adversely affect written language production in

children. Using a serial recall paradigm with German subjects in their native language, Grabowski found that age interacted with recall mode: that is, the word span of seven and ten year old schoolchildren, but not of adults, was significantly better in the oral recall mode than in the written mode. In other words, the results of Grabowski's study were wholly consistent with the hypothesis and findings of the French researchers (Grabowski, 2005).

1.2.4 Working Memory and Education

Grabowski drew attention to the educational implications of the research instigated by Bourdin and Fayol. He observed that if the performance of the same simple serial recall task by primary schoolchildren is significantly worse in the written mode than it is in the oral mode, due to difficulties with handwriting and spelling, then the limited resources of the WM system are likely to be depleted still further by more complex tasks in which planning and redintegration processes must be combined with the management of written output. In view of this shortfall in the quality of written, but not of spoken language production, Grabowski concluded that the evaluation of written performance alone may not be the best way to assess the knowledge or cognitive abilities of young schoolchildren (Grabowski, 2005).

Grabowski's conclusions are in accord with current thinking on the importance of managing the limitations of the WM system in the classroom. In a normal population, there is a general, developmental increase in the capacity of WM over the formative primary and early secondary school years (Gathercole, 1999). At the same time, in an average classroom of nine year old children, individual differences in WM capacity may account for variations in performance equivalent to around five years of cognitive development (Gathercole *et al.*, 2006). The WM system stores information while it is being processed during the learning activities which form the basis for the acquisition of knowledge and for continuing academic development. As might be expected, measures of complex WM span in young children have been found to be closely associated with their levels of literacy and with their mathematical skills (Gathercole & Alloway, 2008). During everyday

classroom activities, such as writing or arithmetic, schoolchildren must store and process information, remember instructions, and produce the required output (Gathercole *et al.*, 2006). Clearly, if WM is a system of limited capacity, teachers must take care not to overtax the mental resources of young children in whom the system is not yet fully developed, particularly those pupils in whom the system may be underdeveloped. If the cognitive load placed upon the WM system by a task is excessive, the child may not be able to accomplish it. Thus, a learning opportunity is lost. If such failure occurs regularly, the rate of learning will be adversely affected (Gathercole & Alloway, 2008; Gathercole *et al.*, 2006).

Gathercole and Alloway have recommended that the learning environment should be controlled so as to minimise the likelihood of WM overload. In particular, they have suggested that learning activities should be organised in such a way as to reduce the load on WM, while maintaining the intended learning outcomes; that ways should be found to support WM demands in the classroom, for example breaking activities down into simple steps; and that children should be helped to evelop strategies of their own, such as reearsal, to enable them to manage their WM resources more effectively (Gathercole & Alloway, 2008).

1.3 Objectives of the Study

In the light of the above, the current study was undertaken with the following objectives:

1. In previous research using a serial recall paradigm, an interaction was observed between the age of the subjects tested and the recall mode used. In both French and German primary school children, but not in adults, the level of recall was significantly higher in the oral mode than in the written mode. It was concluded that in developing writers low-level transcription processes, such as spelling and handwriting, consume limited WM resources to such an extent that the quality of written language production suffers (Grabowski, 2005; Bourdin & Fayol, 1994). The current study used a serial recall paradigm in order to determine whether an

interaction between age and recall mode would be found in native speakers of Turkish.

2. A second objective of the study was to extend the scope of the research by repeating the experiment with the same samples of participants in English. It was anticipated that serial recall in a foreign and therefore less automated language would involve additional cognitive costs which would adversely affect the recall performance of all participants. Furthermore, if age were to interact with recall mode, this negative effect should be most apparent in the younger subjects in the written recall mode. This study therefore sought to determine whether age would in fact interact with recall mode when the test was conducted in English as a FL.

3. A third objective of the study was to compare the outcomes of the NL and FL tests in order to determine whether there might be any association between the two languages with respect to the effects of recall mode and age upon performance.

4. The final objective of the study was to draw attention to the pedagogical implications of the research, with particular reference to the teaching of EFL in Turkish primary schools.

CHAPTER TWO

METHODOLOGY

2.1 Subjects

Forty-eight subjects participated in the experiment. All participants were native speakers of Turkish. The subjects were selected from three different age groups. At the time of the study, all the members of a given age group were at a similar stage in the process of learning English as a FL.

The adult group comprised 16 undergraduate student volunteers (13 female and 3 male) from the Preparatory Class of the Department of English Language and Literature at Kafkas University, Kars, Turkey. Their mean age was 19 years, 6 months (range: 17 years 6 months to 21 years 3 months). All but one of the students had attended a preparatory class in English at their various state secondary schools throughout Turkey, and all had achieved a similar level of success in the Turkish Foreign Language University Entrance Exam. The general level of proficiency in English of the undergraduate group may be categorised as upper intermediate/advanced.

The members of the other two groups, each comprising 16 participants, were pupils from the fifth (8 female and 8 male) and third (8 female and 8 male) grades of KAKÜV Private Primary School in Kars. The mean age of the fifth grade group was 11 years, 0 months (range: 10 years 5 months to 11 years 7 months), while the mean age of the third grade group was 8 years, 11 months (range: 8 years 7 months to 9 years 5 months). The participating children were selected from amongst their classmates on the basis that they had been taking EFL since the first year of primary school (age 6). Two of the children in each class had joined the school in the second grade and therefore had one year less experience of studying English. However, as all four children were considered by their English teachers to be amongst the better students in their respective classes, it was considered appropriate to include them in the study.

2.2 Materials and Design

2.2.1 Design of the Study

The design of the study was based on a procedure devised by Bourdin and Fayol in their research into written and oral language production and working memory in adult and school age native speakers of French (1994). However, the current study extended the scope of the earlier research by using its original serial recall paradigm with native speakers of Turkish in both a native and a foreign language context.

2.2.2 Foreign Language (English) Materials

As the range of material available for use in the FL component of the experiment was constrained by the relatively limited FL vocabulary of the younger primary school children, the words for the English serial recall task were selected first. The Turkish words for the NL serial recall task were then chosen following similar principles, as described in the next section.

Eighty-four frequent bisyllabic nouns were selected for the English serial recall task (See Appendix 1). To ensure that the words chosen would be familiar to the participants, they were selected initially on the basis of the vocabulary content of the English syllabus which had been followed by the participating primary school pupils during their first to third grades (Ashworth & Clark, 1996; House & Scott, 2003; Mitchell & Parker, 2002a; Mitchell & Parker, 2002b, Mitchell & Parker, 2002c; Worrall, 2005). It was taken for granted that these nouns would be familiar to the undergraduate level students of English. The words thus selected were then cross-checked against the British National Corpus (Leech *et al.*, 2001; Davies, 2004), to ensure that they could be considered to be nouns of frequent occurrence (mean frequency: 9829 per 100 million word tokens).

Each noun chosen belonged to one of the following semantic categories: food, nature, animals, time, occupations, family, classroom, appearance, or house.

The 84 words were next grouped into two sets of 42 nouns. Each set of 42 nouns was further divided into a series of seven lists, increasing in length from three to nine words long. Thus, each noun was used once only (2x [3+4+5+6+7+8+9] =84). To avoid task facilitation by semantic linking during serial recall, each of the words at any given list length was selected at random from a different semantic category. At each list length, the semantic categories were placed in random order so as to obviate any effect of cumulative learning due to the linking of position and meaning during serial recall. Words allocated to lists of the same length in the two main word sets were matched in respect of their semantic categories, but not in respect of the order of their presentation, again to prevent learning effects during recall.

The adult group was tested at list lengths increasing from four to nine words, while the primary school participants were tested using word lists ranging from three to seven words long. Thus, although the lists four to seven words long were used for all age groups, the adult group was required to recall a longer series of word lists overall, while the two primary school groups began with a shorter word list than their university counterparts. In this way, the experimental conditions were equalised for all participants, since a two to three-fold increase in STM capacity typically occurs between the ages of four and 14 (Gathercole, 1999).

2.2.3 Native Language (Turkish) Materials

Eighty-four frequent bisyllabic Turkish nouns were selected for the NL serial recall task (See Appendix 2). The Turkish words were chosen, at random, from the same semantic categories used in the selection of the English words. However, in order to prevent learning effects during serial recall, direct translations of the English nouns were avoided.

The 84 Turkish nouns were then sub-divided into word sets and word lists following the same procedure used for the English nouns.

Although the words allocated to lists of the same length were matched for semantic category within each individual language, word lists of the same length

were not matched for semantic category between languages, so as to avoid semantic linking effects during serial recall.

Mirroring the FL phase of the experiment, NL lists increasing in length from four to nine words were used for serial recall by the adult sample, while NL lists increasing in length from three to seven words were used with the two groups of primary school participants.

2.3 Procedure

Permission to conduct the experimental work was obtained from the respective university and school authorities.

Age-appropriate information about the purpose of the study and the procedures to be followed was given to each group of participants in separate sessions prior to commencement of the experimental work. Instruction was given in both English and Turkish. Testing was then carried out in a quiet room at the university or primary school, as applicable, over a four week period beginning in early April, 2008.

Participants were tested individually over two sessions, each approximately 15 minutes long, held at an interval of between one and two weeks apart. In each session, testing was completed in both the written and the oral mode for either English or Turkish. For each participant, the first testing session began with a trial run in the oral mode, using a sample three-word list in the language in which the first test would be conducted. At the start of their second testing session, participants were reminded briefly of the procedure to be followed.

Regardless of whether serial recall was to be carried out in the oral or written mode, all word lists were read aloud by the experimenter, a native speaker of English with proficiency in Turkish. All material was presented orally in order to ensure consistency of encoding throughout the testing, since input is thought to be processed differently according to whether it is presented via auditory or visual channels (Baddeley, 2003a), and since the focus of the research was on phonological short-term memory capacity.

Words were presented at a rate of approximately one word per second. Testing began with the shortest list length (3 words for the primary school groups and 4 words for the adult group) and continued in ascending order until all lists had been presented (concluding with 7 words for the primary school groups and 9 words for the adult group). Participants had to recall each list in serial order immediately after its presentation. No time limit was placed on recall.

In the oral mode, the subjects recalled the words aloud. They were instructed to state "No" (or the Turkish equivalent) if they were unable to recall a particular word, and then to continue in sequence. Each participant's oral responses were marked by the experimenter against a pre-prepared checklist. Oral responses were also simultaneously recorded using a digital recorder (Cenix VR-W600 Series). Upon completion of the testing session, the experimenter's checklist was double-checked against the recording.

Prior to testing in the written mode, the participants were advised that spelling would not be taken into account: a word written incorrectly but recognisably and placed in its original serial position would be considered to be correct. The subjects wrote their responses in pre-prepared notebooks containing separate pages for each list length for each subject. Upon completion of the testing session, the experimenter checked the written responses against the original word lists.

For each language, subjects were tested on their serial recall of two sets of words, subdivided into two sets of word lists of increasing length. At each list length, regardless of the language in which the testing session was being conducted, subjects were asked to recall one list in the oral mode and the other list of the same length in the written mode. Bourdin and Fayol required their subjects to alternate between modes for each list length (1994). In the current study, however, the participants completed recall of all list lengths, from shortest to longest, in either the oral or the written mode, before moving on to recall all lists, in succession, in the alternative mode. This is in line with Grabowski, who argued that switching between recall modes for each list length might incur additional cognitive costs particularly in the children (2005).

The design of the experiment was counterbalanced in order to avoid the same word set being used continually in the same mode, and so as to enable analyses to be

undertaken to determine whether there might be any effects of the order in which language, mode or word set were used. Therefore, half of the subjects from each sample of participants began with a testing session in Turkish and the other half began with a session in English. For each language, half of the subjects started the test in the oral mode, and the other half in the written mode. Half of the subjects who started the test in the oral mode began with the first set of words in the oral mode and continued with the second set in the written mode. The remaining subjects who started recall in the oral mode began with the second set of words in the oral mode and continued with the first set of words in the written mode. Likewise, half of the subjects who started the test in the written mode began recall with the first set of words in the written mode and continued with the second set in the oral mode. The remaining half began with the second set of words in the written mode and continued with the first set of words in the oral mode.

2.4 Statistical Analysis

The proportion of words recalled in their correct serial position at each list length was calculated for all subjects and for both languages. This statistic was used as the dependent variable in the analyses undertaken, thus allowing comparison to be made between serial recall performance results for lists of varying length.

Statistical analyses were performed by analysis of variance (ANOVA), using the General Linear Model Repeated Measures function of the SPSS software package (SPSS 16.0).

CHAPTER THREE

DATA ANALYSIS

3.1 Introduction

Using the proportion of words recalled in their correct serial position at each list length as the dependent variable, a series of ANOVA tests was performed in order to determine whether the hypothesis that writing involves a cognitive load which adversely affects the serial recall performance of children but not of adults would hold valid in a Turkish context.

The data obtained in this study enabled not only the NL but also the FL serial recall performance of the participants to be evaluated. Following separate analysis of the data for each language, a comparative analysis of the results for the Turkish and English serial recall tests was performed.

An alpha level of .05 was used for all statistical tests.

3.2 Analysis of the Native Language (Turkish) Results

3.2.1 Analysis of Variance in the Assignment of Word Sets to Recall Modes and in the Sequencing of Experimental Conditions

A repeated measures ANOVA was performed on the overall mean proportions of Turkish words recalled in their correct serial position, in order to determine whether the allocation of word set to recall mode and/or the sequencing of experimental conditions might yield statistically significant effects. As these control

factors were found to be statistically non-significant ($p > .05$), they were not taken into consideration in the subsequent analyses.

3.2.2 Overview

The mean proportions of Turkish words recalled in their correct serial position are shown in Figure 3.1 as a function of recall mode and age group. The data given here are comprehensive, summarising the results for all subjects, averaged over all list lengths (3 to 7 word lists for primary school subjects; 4 to 9 word lists for adult subjects).

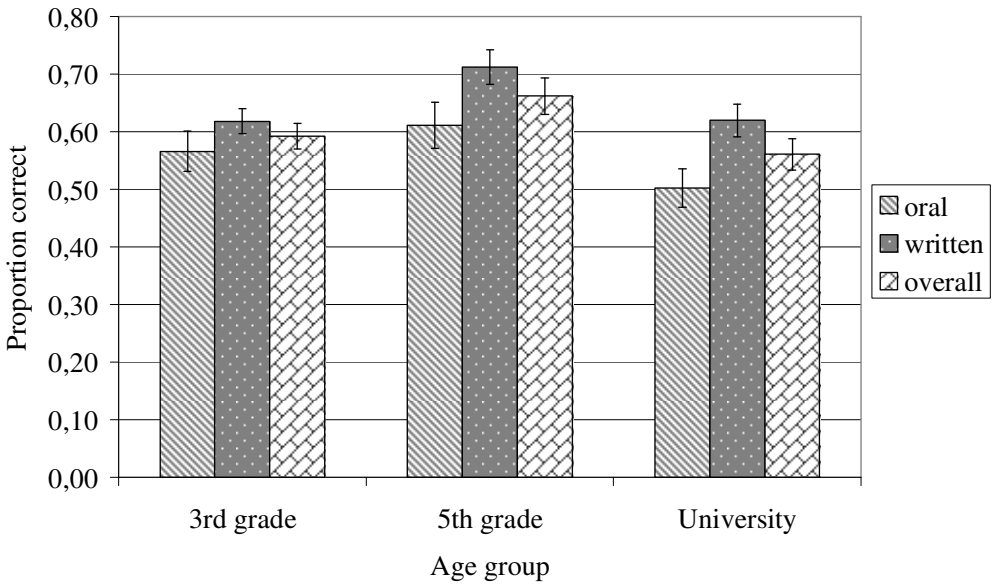


Fig. 3.1 Mean proportions of Turkish words recalled in their correct serial position, averaged over all list lengths, as a function of recall mode and age group. Error bars represent the standard error.

In the NL study as a whole, the overall proportion of words correctly recalled was 0.60 (standard error = 0.02), with a significantly higher level of recall in the written mode (mean = 0.65, SE = 0.02) than in the oral mode (M = 0.56, SE = 0.02), $F(1, 45) = 20.43$, $p < .001$.

Analysis of the data as a function of age group revealed that the overall proportion of words recalled in their correct serial position was higher in the written mode than in the oral mode for all of the samples. Thus, as shown in Figure 3.1 and Table 3.1, the mean proportion of Turkish words correctly recalled by the third grade pupils, averaged over all list lengths, was 0.62 (SE = 0.02) in the written mode, and 0.57 (SE = 0.04) in the oral mode; for the fifth grade pupils it was 0.71 (SE = 0.03) in the written mode, and 0.61 (SE = 0.04) in the oral mode; and for the university students it was 0.62 (SE = 0.03) in the written mode, and 0.50 (SE = 0.03) in the oral mode. In other words, the interaction between recall mode and age was not statistically significant, $F(2, 45) = 0.93$, $p > .05$.

3.2.3 Analysis of Variance in Turkish Word Lists with a Recall Proportion of Approximately 0.60

The use of the proportions of words recalled in their correct serial position as the dependant variable allowed comparisons to be made between word lists of varying length. However, the current analysis also needed to take account of the fact that the primary school subjects began with a shorter word list than their adult counterparts, and that the latter were also required to recall a longer series of word lists overall. Moreover, in an analysis inclusive of data from the shortest to the longest word lists, the presence of floor and ceiling effects may disguise significant effects or interactions (Bourdin & Fayol, 1994). For these reasons, only those Turkish word lists for which the recall proportion was close to the overall mean proportion of recall for the NL phase of the research (0.60), and which had been used with all three samples, were tested for a mode effect in relation to age. Thus, lists of five words (overall recall proportions: 0.63 and 0.65 for the third and fifth grade samples, respectively) and six words (overall recall proportion: 0.61 for the adult sample) were entered as a function of recall mode (oral vs. written) into a repeated measures ANOVA, with age group (adult vs. 3rd grade vs. 5th grade) as the between subjects factor.

Table 3.1 Mean proportions and standard errors of words correctly recalled, as a function of language, age group and recall mode

		Turkish			English		
Age Group		Oral	Written	Overall	Oral	Written	Overall
3rd Grade	Mean	0,57	0,62	0,59	0,46	0,44	0,45
		(0,04)	(0,02)	(0,02)	(0,03)	(0,02)	(0,02)
5th Grade	Mean	0,61	0,71	0,66	0,50	0,52	0,51
		(0,04)	(0,03)	(0,03)	(0,03)	(0,03)	(0,02)
University	Mean	0,50	0,62	0,56	0,38	0,44	0,41
		(0,03)	(0,03)	(0,03)	(0,03)	(0,03)	(0,03)
Total	Mean	0,56	0,65	0,60	0,45	0,47	0,46
		(0,02)	(0,02)	(0,02)	(0,02)	(0,02)	(0,01)

Standard errors in parentheses

The analysis revealed a significant main effect of age, $F(2, 45) = 6.21$, $p = .004$, indicating that the level of performance in the serial recall of NL words, averaged over the five and six word lists, increased significantly with age. As shown in Figure 3.2, the overall mean proportion of Turkish words correctly recalled from these lists was 0.49 (SE = 0.05) for the third grade sample; 0.57 (SE = 0.05) for the fifth grade sample; and 0.71 (SE = 0.05) for the adult sample. Post hoc analyses using the Scheffé test indicated that the mean proportion of words correctly recalled by the adult sample was significantly higher than that of the third grade sample, $p < .05$.

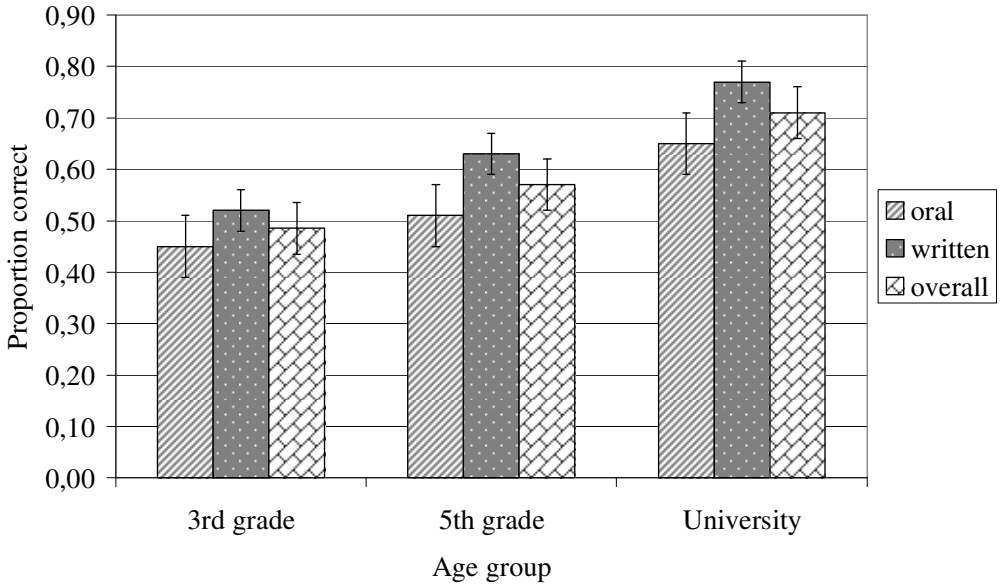


Fig. 3.2 Mean proportions of Turkish words recalled in their correct serial position, averaged over lists of five and six words, as a function of recall mode and age group. Error bars represent the standard error.

The ANOVA also revealed a significant main effect of recall mode, $F(1, 45) = 8.95$, $p = .004$, indicating that a significantly higher proportion of NL words were recalled in their correct serial position in the written mode (M = 0.64, SE = 0.03) than in the oral mode (M = 0.54, SE = 0.04), averaged over all three samples and over both the five and the six word lists.

In fact, for all three samples the mean level of written recall was higher than the mean level of oral recall. However, split analysis of the data revealed that the

effect of recall mode was non-significant at the level of each sample individually ($p > .05$). Thus, as shown in Figure 3.2, for the adult sample the mean proportion of words recalled was 0.77 (SE = 0.04) in the written mode and 0.65 (SE = 0.06) in the oral mode, with the effect of recall mode marginally non-significant, $F(1, 15) = 4.43$, $p = .053$. For the fifth grade sample, it was 0.63 (SE = 0.04) in the written mode and 0.51 (SE = 0.06) in the oral mode, with the effect of recall mode marginally non-significant, $F(1, 15) = 4.22$, $p = .058$. For the third grade sample the mean proportion of words recalled was 0.52 (SE = 0.04) in the written mode and 0.45 (SE = 0.06) in the oral mode, with the effect of recall mode non-significant, $F(1, 15) = 1.19$, $p = .292$.

The interaction between mode and age was non-significant, $F(2, 45) = 0.24$, $p > .05$.

3.2.4 Analysis of Variance in the Serial Recall of Turkish Words in Relation to Recall Mode and List Length

The ANOVA performed on Turkish word lists with a recall proportion of close to 0.60 found significant main effects of age and mode. It did not reveal a statistically significant interaction between mode and age. In order to test further the significance of the effects found here, a second phase of analysis was carried out.

Following the procedure of Bourdin and Fayol (1994), two separate ANOVAs were conducted: one on the data obtained from the adult sample and the other on the data obtained from the primary school samples, in each case examining serial recall performance in relation to recall mode and list length.

3.2.4.1 Analysis of Variance in the Adult Sample

The proportions of Turkish words recalled in their correct serial position by the adult participants were entered into a 2 (recall mode: oral vs. written) X 5 (list length: 5 to 9 words) ANOVA with repeated measures on both factors.

The ANOVA revealed a significant main effect of recall mode, $F(1, 15) = 14.85$, $p = .002$, indicating that a significantly higher proportion of NL words were recalled in their correct serial position in the written mode ($M = 0.54$, $SE = 0.03$) than in the oral mode ($M = 0.42$, $SE = 0.04$), as shown in Figure 3.3.

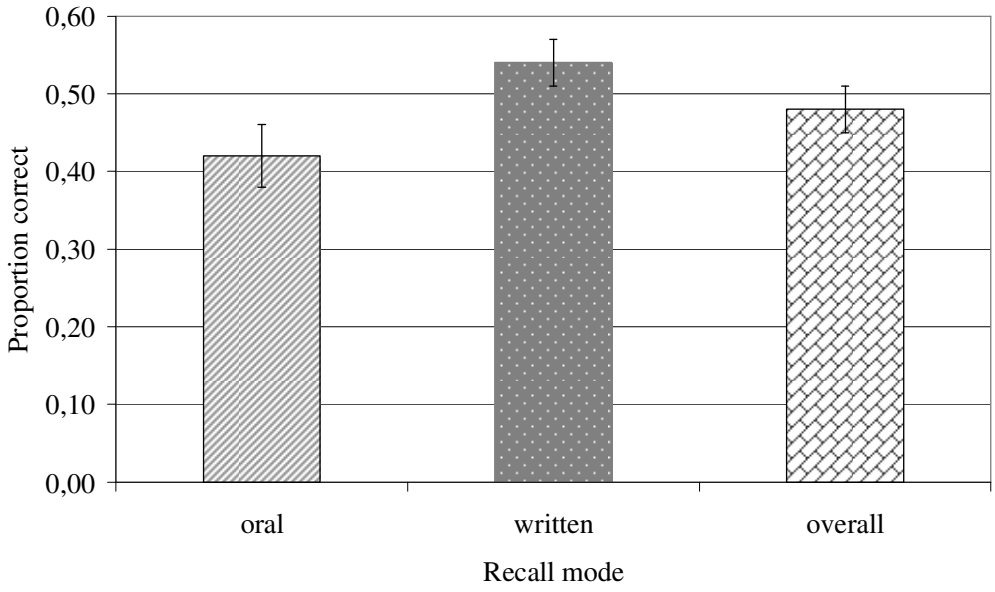


Fig. 3.3 Mean proportions of Turkish words recalled in their correct serial position by the adult sample, averaged over lists of five to nine words, as a function of recall mode. Error bars represent the standard error.

As shown in Figure 3.4, the ANOVA also revealed a significant main effect of word list length, $F(4, 60) = 37.74$, $p < .001$. The mean proportion of words recalled in their correct serial position, averaged over both recall modes, decreased from 0.81 ($SE = 0.05$) for five-word lists to 0.32 ($SE = 0.05$) for nine-word lists. However, there was no significant interaction between list length and recall mode, $F(4, 60) = 0.71$, $p > .05$. At each length, the proportion of recall was higher in the written mode.

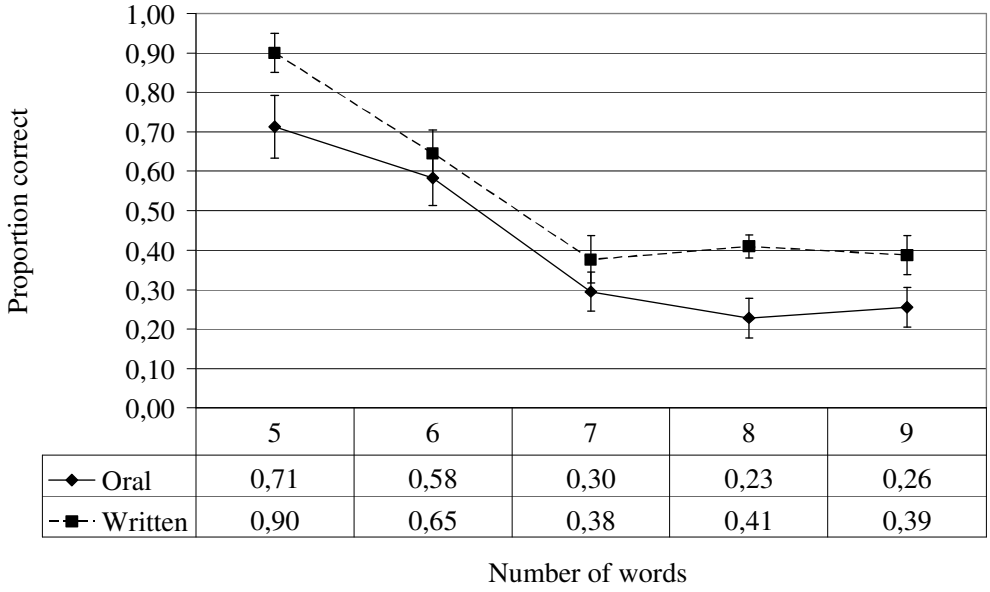


Fig. 3.4 Mean proportions of Turkish words recalled in their correct serial position by the adult sample, for lists of five to nine words, as a function of recall mode and list length. Error bars represent the standard error.

3.2.4.2 Analysis of Variance in the School Age Samples

The proportions of Turkish words recalled in their correct serial position by the two school age samples of participants were entered into a 2 (grade: 3rd vs. 5th) X 2 (recall mode: oral vs. written) X 5 (list length: 3 to 7 words) ANOVA with repeated measures on the last two factors.

As shown in Figure 3.5, the overall level of performance in the serial recall of NL words was higher in the fifth grade (M = 0.66, SE = 0.03) than in the third grade (M = 0.59, SE = 0.03). However, the difference between the means for the two age groups was not statistically significant, $F(1, 30) = 3.30$, $p > .05$.

The ANOVA revealed a significant main effect of recall mode, $F(1, 30) = 9.09$, $p = .005$, indicating that a significantly higher proportion of NL words were recalled in their correct serial position in the written mode (M = 0.67, SE = 0.02) than in the oral mode (M = 0.59, SE = 0.03), by the children overall.

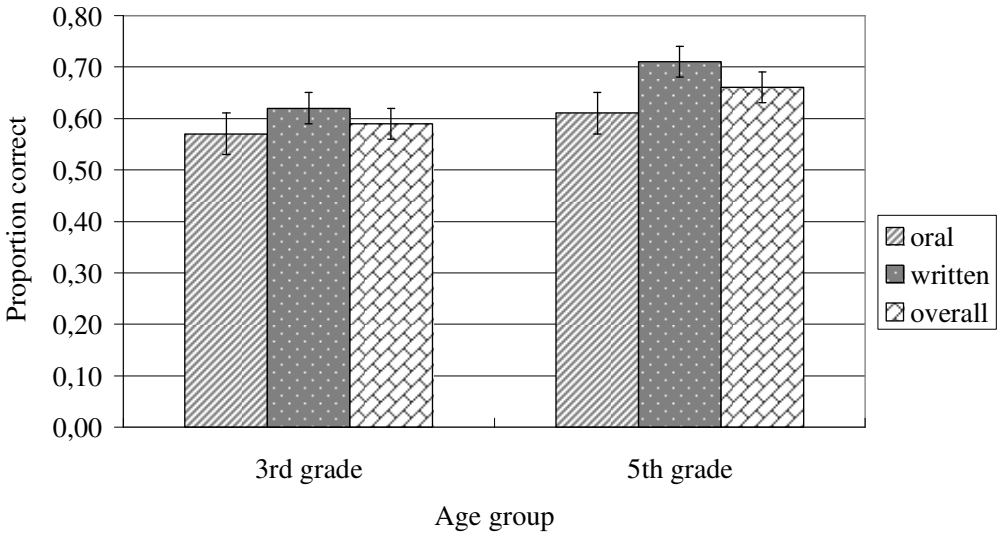
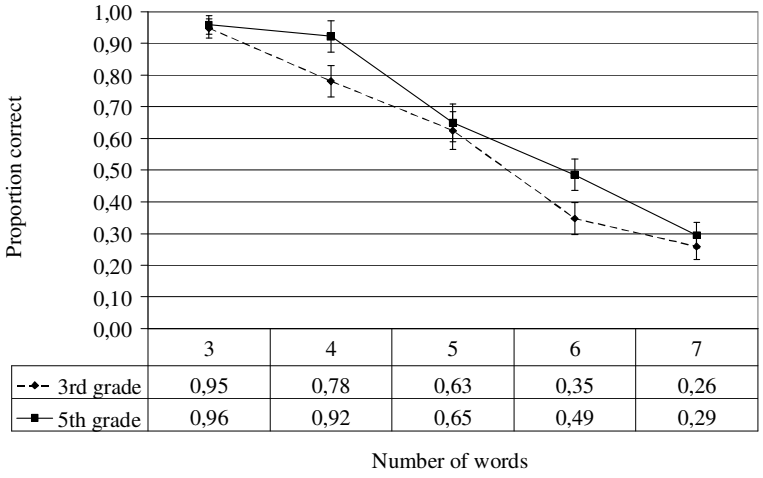


Fig. 3.5 Mean proportions of Turkish words recalled in their correct serial position by the third and fifth grade samples, averaged over lists of three to seven words, as a function of recall mode. Error bars represent the standard error.

Separate analysis of the data for the two groups revealed that the level of serial recall was higher in the written mode than in the oral for both samples. As shown in Figure 3.5, the fifth grade sample recalled a mean proportion of 0.71 (SE = 0.03) of the words in the written mode and 0.61 (SE = 0.04) in the oral mode, while the third grade sample recalled a mean proportion of 0.62 (SE = 0.03) of the words in the written mode and 0.57 (SE = 0.04) in the oral mode. However, while the effect of recall mode was statistically significant for the fifth grade sample, $F(1, 15) = 9.13$, $p = .009$, it was non-significant for the third grade sample, $F(1, 15) = 1.86$, $p > .05$.

The interaction between age and recall mode was non-significant, $F(1, 30) = 0.92$, $p > .05$.

The ANOVA revealed a significant main effect of word list length, $F(4, 120) = 104.13$, $p < .001$. The mean proportion of words recalled in their correct serial position, averaged over both of the recall modes and both of the primary school samples, decreased from 0.95 (SE = 0.02) for three-word lists to 0.28 (SE = 0.03) for seven-word lists.

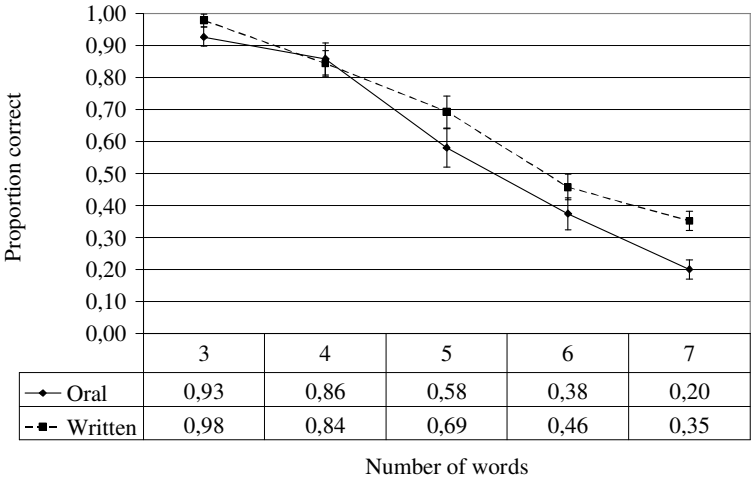


	3	4	5	6	7
3rd grade	0,95	0,78	0,63	0,35	0,26
5th grade	0,96	0,92	0,65	0,49	0,29

Fig. 3.6 Mean proportions of Turkish words recalled in their correct serial position by the third and fifth grade samples, for lists of three to seven words, as a function of age. Error bars represent the standard error.

Moreover, as shown in Figure 3.6, this pattern was consistent for each sample when considered separately. In other words, list length did not interact significantly with age, $F(4, 120) = 1.31$, $p > .05$.

Likewise, as shown in Figure 3.7, the interaction between list length and recall mode was non-significant, $F(4, 120) = 1.44$, $p > .05$.



	3	4	5	6	7
Oral	0,93	0,86	0,58	0,38	0,20
Written	0,98	0,84	0,69	0,46	0,35

Fig. 3.7 Mean proportions of Turkish words recalled in their correct serial position for lists of three to seven words, averaged over the third and fifth grade samples, as a function of recall mode. Error bars represent the standard error.

3.3 Analysis of the Foreign Language (English) Results

3.3.1 Analysis of Variance in the Assignment of Word Sets to Recall Modes and in the Sequencing of Experimental Conditions

A repeated measures ANOVA was performed on the overall mean proportions of English words recalled in their correct serial position, in order to determine whether the allocation of word set to recall mode and/or the sequencing of experimental conditions might yield statistically significant effects. As these control factors were found to be statistically non-significant, they were not taken into consideration in the subsequent analyses.

3.3.2 Overview

The mean proportions of English words recalled in their correct serial position are shown in Figure 3.8 as a function of recall mode and age group.

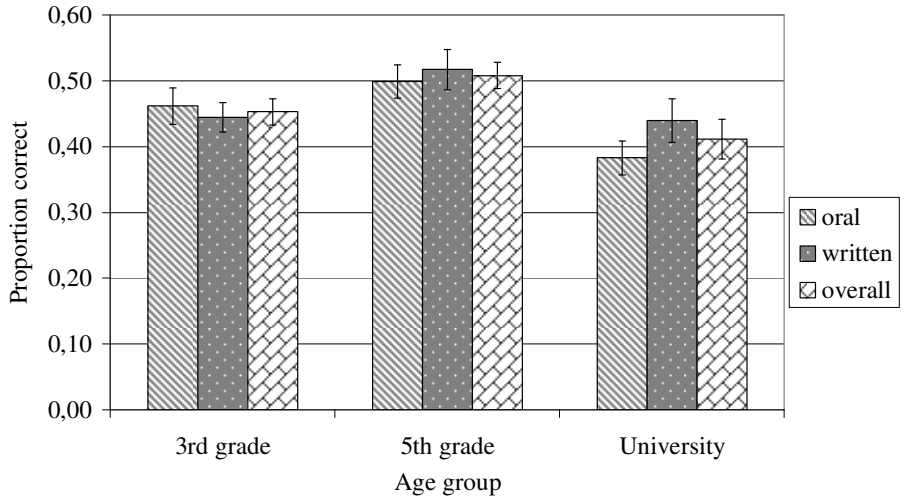


Fig. 3.8 Mean proportions of English words recalled in their correct serial position, averaged over all list lengths, as a function of recall mode and age group. Error bars represent the standard error.

The data given here are comprehensive, summarising the results for all subjects, averaged over all list lengths (3 to 7 word lists for primary school subjects; 4 to 9 word lists for adult subjects).

The overall proportion of words correctly recalled was 0.46 (SE = 0.01), with a slightly higher level of recall in the written mode (M = 0.47, SE = 0.02) than in the oral mode (M = 0.45, SE = 0.02). Thus, the effect of recall mode was statistically non-significant, $F(1, 45) = 0.93$, $p > .05$.

Analysis of the data as a function of age group revealed that the overall proportion of words successfully recalled remains higher in the written mode than in the oral mode for the adult and fifth grade samples. Thus, as shown in Figure 3.8 and Table 3.1, the mean proportion of English words correctly recalled by the fifth grade pupils, over all list lengths, was 0.52 (SE = 0.03) in the written mode, and 0.50 (SE = 0.03) in the oral mode; while for the university students it was 0.44 (SE = 0.03) in the written mode, and 0.38 (SE = 0.03) in the oral mode. On the other hand, the third grade pupils demonstrated a somewhat higher level of recall in the oral mode than in the written mode, with 0.46 (SE = 0.03) of words correctly recalled in the oral mode and 0.44 (SE = 0.02) in the written mode. Nonetheless, the interaction between recall mode and age was found to be non-significant, $F(2, 45) = 1.20$, $p > .05$.

3.3.3 Analysis of Variance in English Word Lists with a Recall Proportion of Approximately 0.50

The initial FL analysis included data from the shortest to the longest word lists. Therefore, the presence of floor and ceiling effects might have masked an underlying recall mode effect or a significant interaction between mode and age. Moreover, account needed to be taken of the fact that the primary school subjects began with a shorter word list than their adult counterparts, and that the latter were also required to recall a longer series of word lists overall. Therefore, a further ANOVA was performed. To avoid floor and ceiling effects, only those English word lists for which the recall proportion was relatively close to 0.50, and which had been used with all three samples, were tested for a mode effect in relation to age. Thus,

lists of four words (overall recall proportion: 0.63 for the third grade sample) and five words (overall recall proportions: 0.43 and 0.54 for the fifth grade and adult samples, respectively) were entered as a function of recall mode (oral vs. written) into a repeated measures ANOVA, with age group (adult vs. 3rd grade vs. 5th grade) as the between subjects factor.

The analysis revealed a significant main effect of age, $F(2, 45) = 8.03$, $p = .001$, indicating that the level of performance in the serial recall of FL words, averaged over the four and five word lists, increased significantly with age. As shown in Figure 3.9, the overall mean proportion of English words correctly recalled from these lists was 0.51 (SE = 0.03) for the third grade sample; 0.56 (SE = 0.03) for the fifth grade sample; and 0.70 (SE = 0.03) for the adult sample. Post hoc analyses using the Scheffé test indicated that the mean proportion of words correctly recalled by the adult sample was significantly higher than that of both the third and fifth grade samples, $p < .05$.

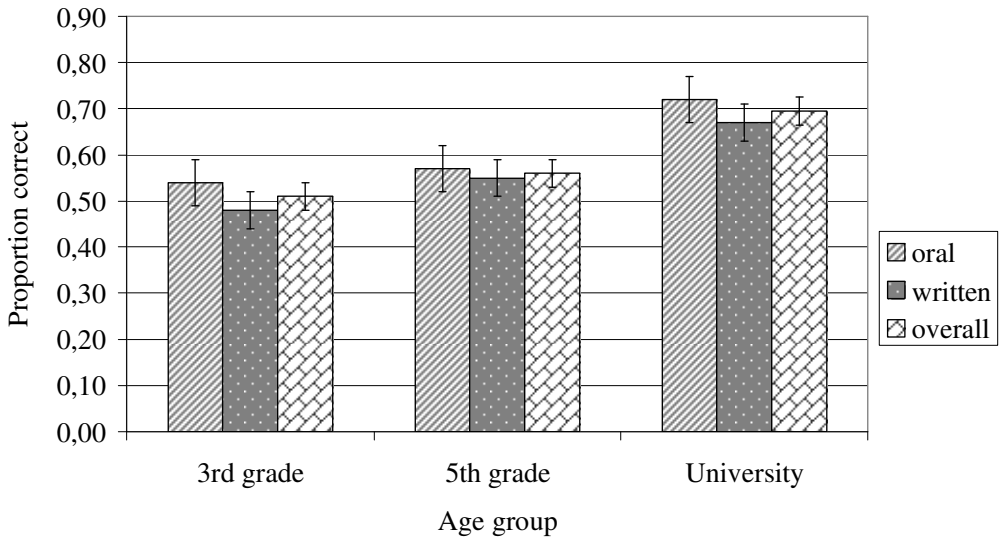


Fig. 3.9 Mean proportions of English words recalled in their correct serial position, averaged over lists of four and five words, as a function of recall mode and age group. Error bars represent the standard error.

As shown in Figure 3.9, all three samples recalled a higher proportion of FL words in their correct serial position in the oral mode than in the written mode, over the four and five word lists as a whole. Thus, for the third grade sample the mean proportion of English words recalled was 0.54 (SE = 0.05) in the oral mode and 0.48 (SE = 0.04) in the written mode; for the fifth grade, it was 0.57 (SE = 0.05) in the oral mode and 0.55 (SE = 0.04) in the written mode; and for the adult sample it was 0.72 (SE = 0.05) in the oral mode and 0.67 (SE = 0.04) in the written mode. However, the effect of recall mode was not statistically significant, $F(1, 45) = 1.40$, $p > .05$.

Likewise, the interaction between mode and age in this FL context was non-significant, $F(2, 45) = 0.08$, $p > .05$.

3.3.4 Analysis of Variance in the Serial Recall of English Words in Relation to Recall Mode and List Length

The ANOVA performed on English word lists with a recall proportion of close to 0.50 did not reveal a statistically significant interaction between mode and age, or a significant effect of recall mode. To investigate these results further, a second phase of analysis was undertaken.

As in the NL study, two separate ANOVAs were performed: one on the FL data obtained from the adult sample for lists five to nine words long, and the other on the FL data obtained from the third and fifth grade samples for lists three to seven words long. In each case serial recall performance was examined in relation to recall mode.

3.3.4.1 Analysis of Variance in the Adult Sample

The proportions of English words recalled in their correct serial position by the adult participants were entered into a 2 (recall mode: oral vs. written) X 5 (list length: 5 to 9 words) ANOVA with repeated measures on both factors.

As shown in Figure 3.10, the ANOVA revealed a significant main effect of recall mode, $F(1, 15) = 10.37$ $p = .006$, indicating that a significantly higher proportion of FL words were recalled in their correct serial position in the written mode ($M = 0.37$, $SE = 0.03$) than in the oral mode ($M = 0.28$, $SE = 0.03$).

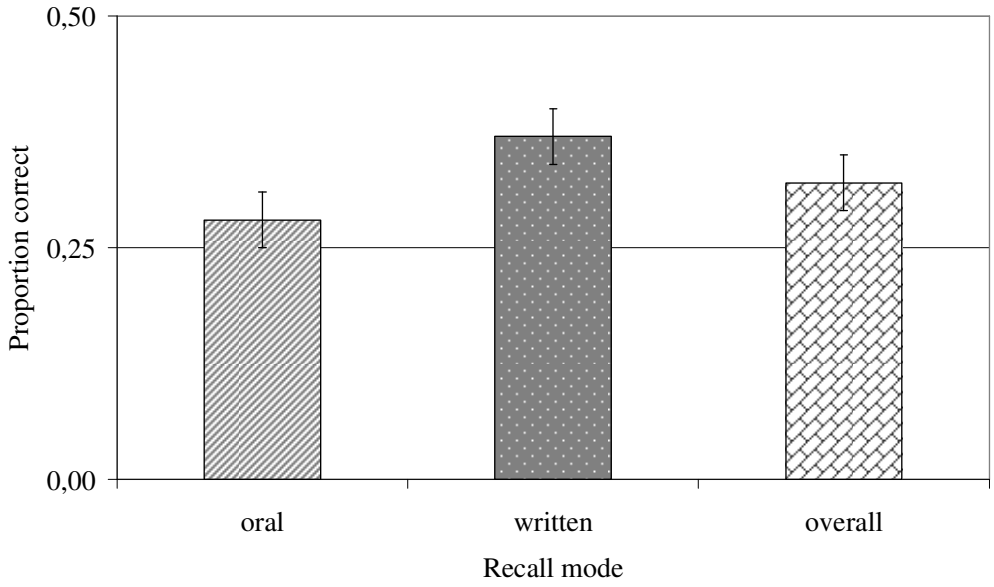
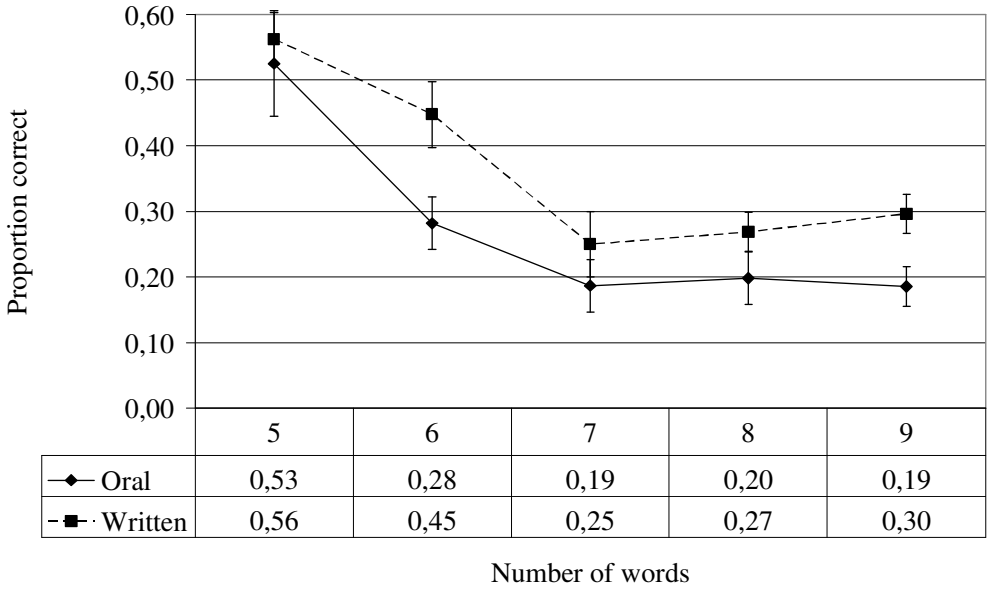


Fig. 3.10 Mean proportions of English words recalled in their correct serial position by the adult sample, averaged over lists of five to nine words, as a function of recall mode. Error bars represent the standard error.

The analysis also revealed a significant main effect of word list length, $F(4, 60) = 31.43$, $p < .001$. The mean proportion of words recalled in their correct serial position, averaged over both recall modes, decreased from 0.54 ($SE = 0.04$) for five-word lists to 0.24 ($SE = 0.03$) for nine-word lists.

However, as shown in Figure 3.11, the interaction between list length and recall mode was non-significant, $F(4, 60) = 0.63$, $p > .05$. At each list length, the proportion of recall was higher in the written mode.



	5	6	7	8	9
Oral	0,53	0,28	0,19	0,20	0,19
Written	0,56	0,45	0,25	0,27	0,30

Number of words

Fig. 3.11 Mean proportions of English words recalled in their correct serial position by the adult sample, for lists of five to nine words, as a function of recall mode and list length. Error bars represent the standard error.

3.3.4.2 Analysis of Variance in the School Age Samples

The proportions of English words recalled in their correct serial position by the 2 school age samples of participants were entered into a 2 (grade: 3rd vs. 5th) X 2 (recall mode: oral vs. written) X 5 (list length: 3 to 7 words) ANOVA with repeated measures on the last two factors.

Although the overall level of performance in the serial recall of FL words was higher in the fifth grade ($M = 0.51$, $SE = 0.02$) than in the third grade ($M = 0.45$, $SE = 0.02$), as shown in Figure 3.12, the difference between the means for the two age groups was not statistically significant, $F(1, 30) = 3.49$, $p > .05$.

The ANOVA revealed that the effect of recall mode was non-significant, F(1, 30) = 0.00, p > .05. In fact, the same proportion of FL words was correctly recalled in the written mode (M = 0.48, SE = 0.02), as in the oral mode (M = 0.48, SE = 0.02), by the children overall. Similarly, split analysis of the data revealed that the effect of recall mode was non-significant at the level of each sample individually.

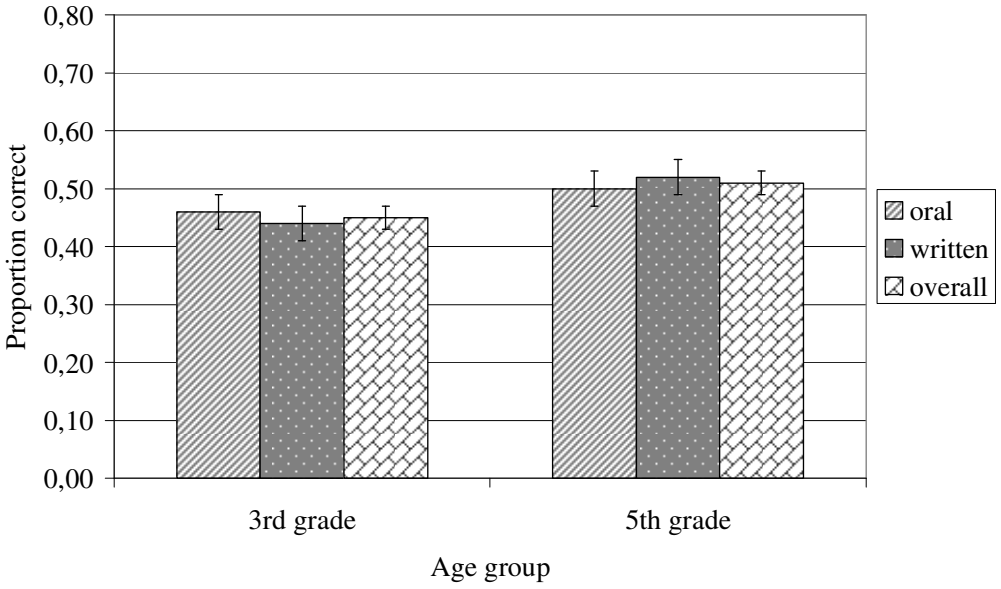
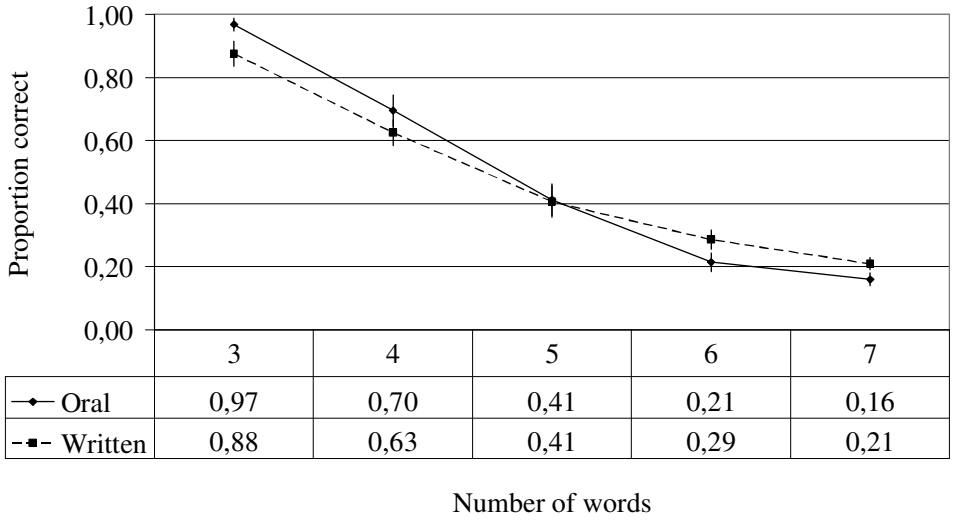


Fig. 3.12 Mean proportions of English words recalled in their correct serial position by the third and fifth grade samples, averaged over lists of three to seven words, as a function of recall mode. Error bars represent the standard error.

Thus, as shown in Figure 3.12, for the fifth grade sample the mean proportion of words recalled was 0.52 (SE = 0.03) in the written mode and 0.50 (SE = 0.03) in the oral mode, with the effect of recall mode non-significant, F (1, 15) = 0.34, p > .05. For the third grade sample the mean proportion of words recalled was 0.44 (SE = 0.03) in the written mode and 0.46 (SE = 0.03) in the oral mode, with the effect of recall mode non-significant, F (1, 15) = 0.24, p > .05. Likewise, the interaction between age and recall mode was found to be non-significant, F(1, 30) = 0.56, p > .05.

The analysis revealed a significant main effect of word list length, $F(3.39, 101.60) = 169.72$, $p < .001$. The mean proportion of FL words recalled in their correct serial position, averaged over both of the recall modes and both of the primary school samples, decreased from 0.92 (SE = 0.02) for three-word lists to 0.16 (SE = 0.02) for seven-word lists.

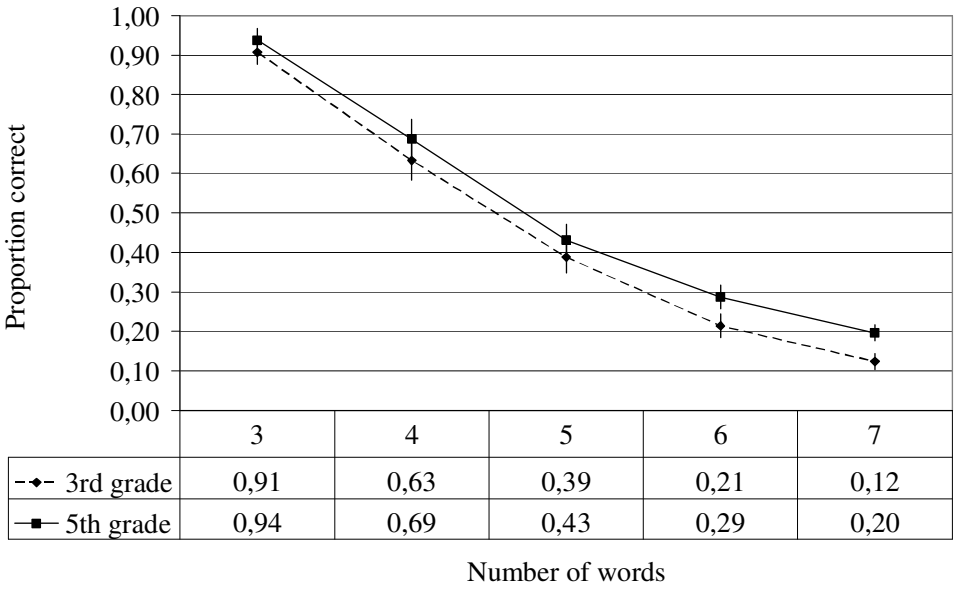
As shown in Figure 3.13, the ANOVA revealed a significant interaction between recall mode and list length, $F(4, 120) = 3.03$, $p = 0.02$. Averaged over the third and fifth grade samples, a higher proportion of FL words were recalled in their correct serial position in the oral recall mode at shorter list lengths (3 to 5 words), whereas at longer list lengths (6 and 7 words), a higher proportion of words were recalled in their correct serial position in the written recall mode.



	3	4	5	6	7
Oral	0,97	0,70	0,41	0,21	0,16
Written	0,88	0,63	0,41	0,29	0,21

Fig. 3.13 Mean proportions of English words recalled in their correct serial position for lists of three to seven words, averaged over the third and fifth grade samples, as a function of recall mode. The interaction between recall mode and list length is significant ($p < .05$). Error bars represent the standard error.

List length did not interact significantly with age, $F(3.39, 101.60) = 0.15$, $p > .05$, as can be seen in Figure 3.14.



	3	4	5	6	7
3rd grade	0,91	0,63	0,39	0,21	0,12
5th grade	0,94	0,69	0,43	0,29	0,20

Fig. 3.14 Mean proportions of English words recalled in their correct serial position by the third and fifth grade samples, for lists of three to seven words, as a function of age. Error bars represent the standard error.

3.4 Comparative Analysis of the Native Language (Turkish) and Foreign Language (English) Results

The serial recall performance of all participants in the study was tested in both Turkish, their native language, and English, their foreign language of study. This enabled a comparative analysis of the results to be undertaken between languages.

3.4.1 Overview

In the study as a whole, the overall proportion of words correctly recalled was 0.53 (SE = 0.01), with a significantly higher level of recall in the written mode (M = 0.56, SE = 0.02) than in the oral mode (M = 0.50, SE = 0.02), $F(1, 45) = 12.46$, $p = .001$. These data are comprehensive, summarising the results for all subjects, averaged over all list lengths (3 to 7 word lists for primary school subjects; 4 to 9 word lists for adult subjects), for both languages.

Analysis of the data as a function of language, as shown in Figure 3.15 and Table 3.1, revealed that the overall proportion of words successfully recalled was significantly higher in the NL (M = 0.60, SE = 0.02) than in the FL (M = 0.46, SE = 0.02), $F(1, 45) = 155.11$, $p < .001$.

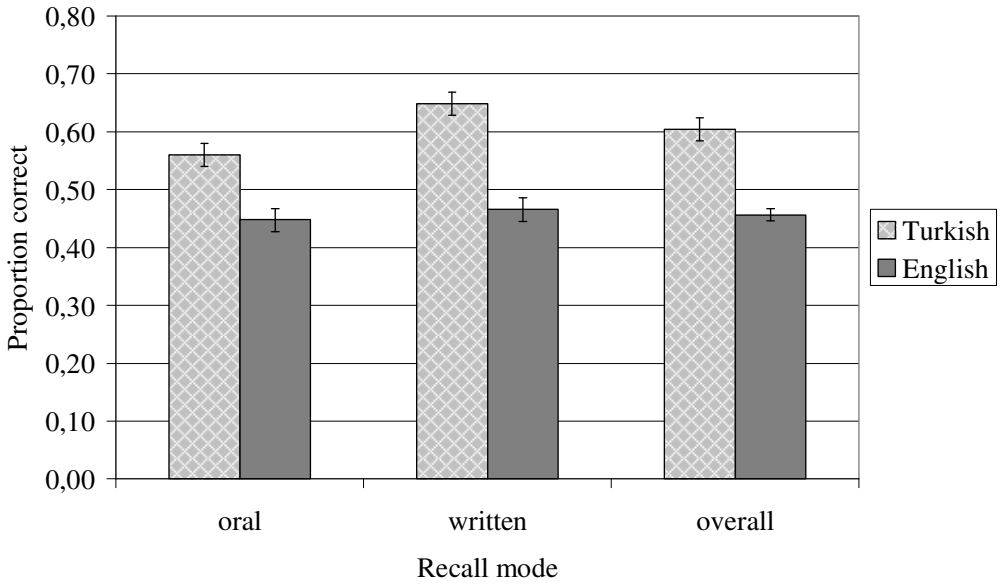


Fig. 3.15 Mean proportions of Turkish and English words recalled in their correct serial position, averaged over all list lengths, as a function of recall mode. Error bars represent the standard error.

For both languages, the level of successful serial recall was higher in the written mode than in the oral mode. Thus, the mean proportion of Turkish words correctly recalled was 0.65 in the written mode (SE = 0.02), and 0.56 (SE = 0.02) in the oral mode, while the mean proportion of English words correctly recalled was 0.47 in the written mode (SE = 0.02), and 0.45 (SE = 0.02) in the oral mode. However, while this effect of recall mode was statistically significant for the Turkish data, $F(1, 45) = 20.43$, $p < 0.001$, it was non-significant for the English data, $F(1, 45) = 0.93$, $p > .05$.

3.4.2 Analysis of Variance in the Serial Recall of Native and Foreign Language Words in Lists Four to Seven Words Long

The comparative analysis reported in the previous section included data from the shortest to the longest word lists in both languages, and may therefore have been subject to floor and ceiling effects. Moreover, account needed to be taken of the fact that the primary school participants began with a shorter word list than their adult counterparts and also recalled a shorter series of word lists overall. Therefore, in order to clarify the effects of recall mode on serial recall in relation to age group and language, only the data from those word lists which had been presented to all three samples, i.e. lists of four to seven words, were used in the next analysis.

The proportions of words recalled in their correct serial position were entered into a 3 (age group: adult vs. 3rd grade vs. 5th grade) X 2 (language order: Turkish 1st vs. English 1st) X 2 (language: Turkish vs. English) X 2 (recall mode: oral vs. written) X 4 (list length: 4 to 7 words) ANOVA with repeated measures on the last three factors.

The mean proportion of words correctly recalled for lists four to seven words long, averaged over both languages and recall modes and all age groups, was 0.50 (SE = 0.01).

The ANOVA revealed that the overall level of serial recall increased significantly with age, $F(2, 42) = 11.74$, $p < .001$. Moreover, post hoc analyses using the Scheffé test indicated that the level of recall was significantly higher for each

single age group in relation to any younger group(s), $p < .05$. Thus, as shown in Figure 3.16, the overall mean proportion of words correctly recalled from these lists was 0.59 (SE = 0.02) for the adult sample; 0.49 (SE = 0.02) for the fifth grade sample; and 0.42 (SE = 0.02) for the third grade sample.

Analysis of the data as a function of language revealed that the overall proportion of words successfully recalled was significantly higher in the NL (M = 0.59, SE = 0.02) than in the FL (M = 0.41, SE = 0.01), $F(1, 42) = 172.83$, $p < 0{,}001$.

Moreover, split analysis of the data revealed that recall performance was also significantly better in Turkish than in English at the level of each sample individually. Thus, as shown in Figure 3.16, for the adult sample the mean proportion of words recalled was 0.68 (SE = 0.03) in Turkish and 0.49 (SE = 0.02) in English, $F(1, 14) = 49.74$, $p < .001$. For the fifth grade sample, it was 0.59 (SE = 0.04) in the Turkish and 0.40 (SE = 0.02) in English, $F(1, 14) = 50.69$, $p < .001$. For the third grade sample, it was 0.50 (SE = 0.03) in Turkish and 0.34 (SE = 0.02) in English, $F(1, 14) = 94.11$, $p < .001$. In other words, the interaction between language and age was not statistically significant, $F(2, 42) = 0.29$, $p > .05$.

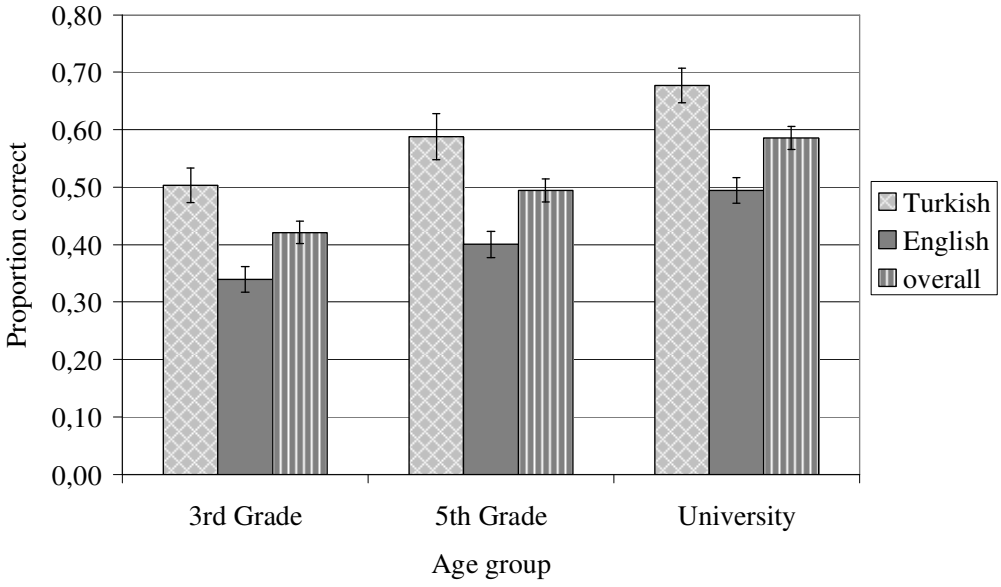


Fig. 3.16 Mean proportions of words recalled in their correct serial position, averaged over lists of four to seven words, as a function of language and age group. Error bars represent the standard error.

The ANOVA also revealed a significant main effect of recall mode, F(1, 42) = 9.07, p = .004, indicating that a significantly higher proportion of words were recalled in their correct serial position in the written mode (M = 0.53, SE = 0.02) than in the oral mode (M = 0.47, SE = 0.02), averaged over both languages and all three samples, for lists of four to seven words.

In fact, for all three samples the mean level of written recall was higher than the mean level of oral recall, averaged over both languages. Thus, the interaction between recall mode and age was non-significant, F(2, 42) = 0.45, p > .05.

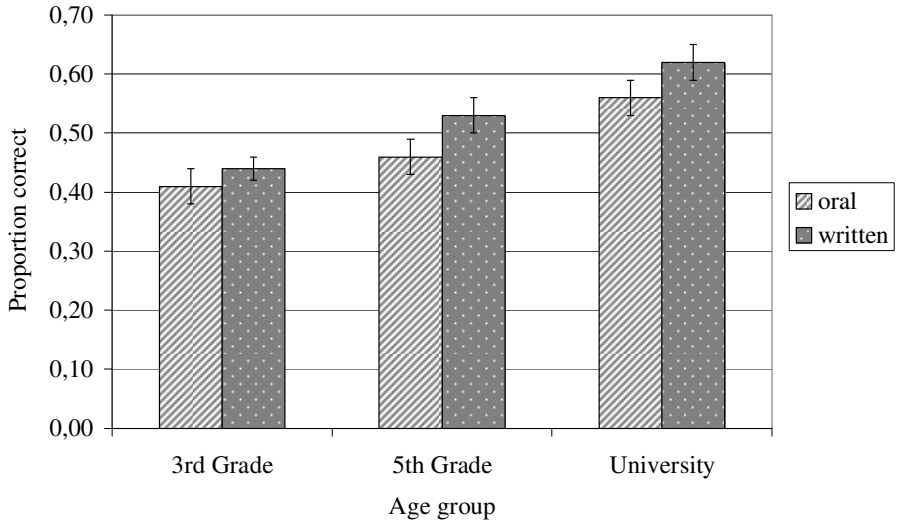
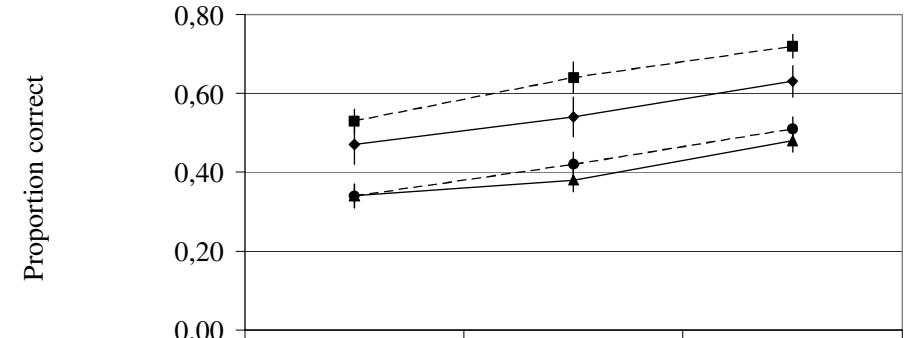


Fig. 3.17 Mean proportions of words recalled in their correct serial position, averaged over lists of four to seven words for Turkish and English, as a function of recall mode and age group. Error bars represent the standard error.

Split analysis of the data revealed that, the effect of recall mode was statistically significant for only one of the samples, when considered individually. Thus, as shown in Figure 3.17, the mean proportion of words correctly recalled by the fifth grade sample was 0.53 (SE = 0.03) in the written mode and 0.46 (SE = 0.03) in the oral mode, with the effect of recall mode statistically significant, F (1, 14) = 6.41, p = .02. The mean proportion of words recalled by the adult sample was 0.62 (SE = 0.03) in the written mode and 0.56 (SE = 0.03) in the oral mode, with the effect of recall mode marginally non-significant, F (1, 14) = 4.55, p = .051. The

mean proportion of words recalled by the third grade sample was 0.44 (SE = 0.02) in the written mode and 0.41 (SE = 0.03) in the oral mode, with the effect of recall mode non-significant, F (1, 14) = 0.73, p > .05.

The interaction between language and mode was found to be marginally non-significant, F(1, 42) = 4.05, p = .051. Further analysis revealed that the interaction between language and mode was statistically non-significant at the level of each sample individually (p > .05). In other words, as can be seen in Figure 3.18, the interaction between language, mode and age was statistically non-significant, F (2, 42) = 0.004, p > .05.



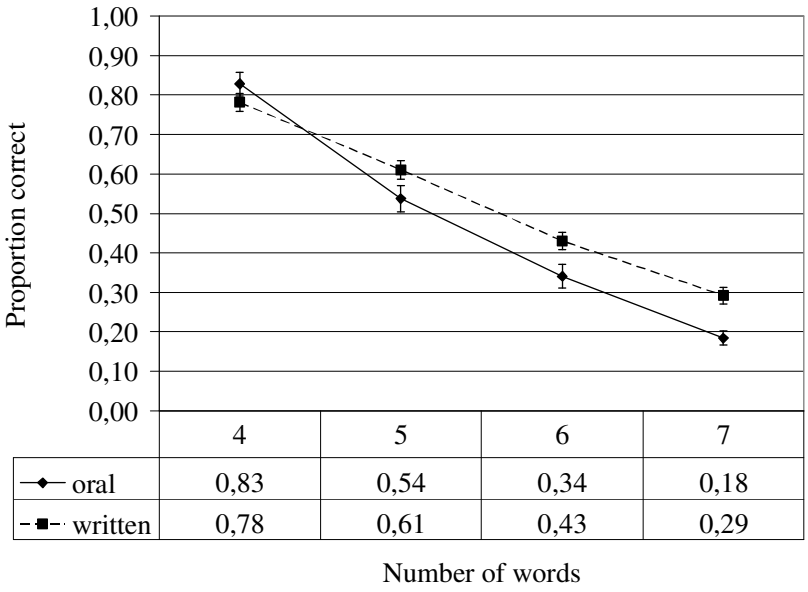
	3rd Grade	5th Grade	University
Turkish oral	0,47	0,54	0,63
Turkish written	0,53	0,64	0,72
English oral	0,34	0,38	0,48
English written	0,34	0,42	0,51

Age group

Fig. 3.18 Mean proportions of words recalled in their correct serial position, averaged over lists of four to seven words, as a function of language, recall mode and age group. Error bars represent the standard error.

The ANOVA revealed a significant main effect of word list length, F(3, 126) = 221.82, p < .001. The mean proportion of words recalled in their correct serial position, averaged over both languages and all age groups, decreased from 0.81 (SE = 0.02) for four-word lists to 0.24 (SE = 0.02) for seven-word lists. List length did not interact significantly with age, F(6, 126) = 1.26, p > .05, indicating that his pattern was consistent across all age groups.

As shown in Figure 3.19, there was a significant interaction between recall mode and list length, $F(3, 126) = 6.51$, $p < .001$. Averaged over both languages and all age groups, a higher proportion of words were recalled in their correct serial position in the written recall mode at the longer list lengths (5 to 7 words), whereas at the shortest list length (4 words), a higher proportion of words was recalled correctly in the oral mode.



	4	5	6	7
oral	0,83	0,54	0,34	0,18
written	0,78	0,61	0,43	0,29

Fig. 3.19 Mean proportions of words recalled in their correct serial position for lists of four to seven words, averaged over all age groups for Turkish and English, as a function of recall mode. Error bars represent the standard error.

There were no significant interactions between mode, length and age; language, mode and length; or language, mode, length and age, ($p > .05$).

As shown in Figure 3.20, the ANOVA revealed a significant interaction between language and length, $F(3,126) = 3.69$, $p = .014$, indicating that serial recall performance was significantly higher in the NL at each list length. The interaction between language, length and age was non-significant, $F(6, 126) = 1.24$, $p > .05$, indicating that this pattern was consistent across all samples.

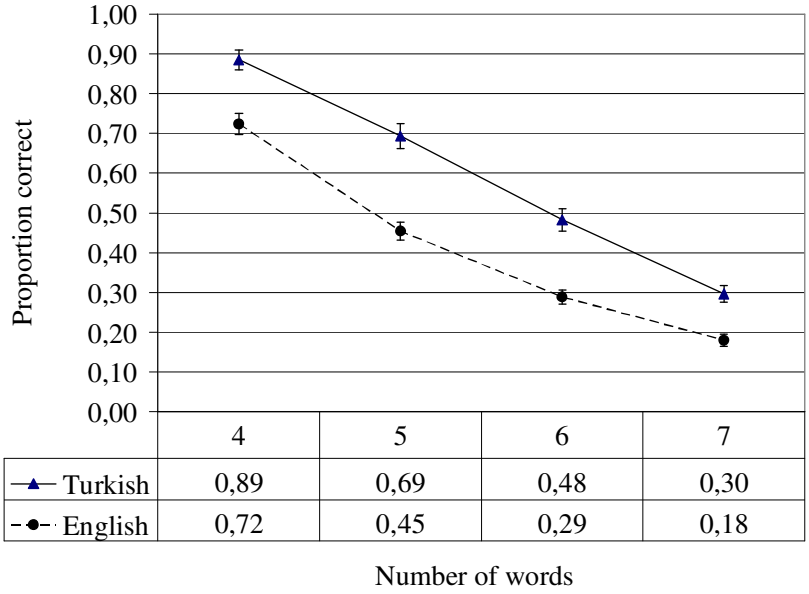


Fig. 3.20 Mean proportions of words recalled in their correct serial position for lists of four to seven words, averaged over all age groups, as a function of language. Error bars represent the standard error.

The analysis also addressed the question of whether the order in which the languages were presented to the subjects might have affected their serial recall performance. Averaged over all samples, the mean proportion of words recalled in their correct serial position was found to be higher for those subjects who started with the Turkish test ($M = 0.53$, $SE = 0.02$), than for those who started with the English test ($M = 0.47$, $SE = 0.02$). However, this effect of language order was found to be marginally non-significant, $F(1, 42) = 3.86$, $p = .056$.

In contrast, the interaction between language order and age was found to be marginally significant, $F(2, 42) = 3.31$, $p = .046$. However, split analysis of the data revealed that the effect of language order only reached statistical significance for the

adult group. Thus, the mean proportion of words correctly recalled by the subjects from the adult sample who started with the Turkish test was 0.64 (SE = 0.02), while for those who started with the English test it was 0.53 (SE = 0.04), with the effect of language order marginally significant, $F(1, 14) = 4.80$, $p = .046$. For those adults who started in the NL, the mean proportion of Turkish words recalled in the first trial was 0.71 (SE= 0.04), while the mean proportion of English words recalled in the second trial was 0.56 (SE= 0.03). For those adults who started in the FL, the mean proportion of English words recalled in the first trial was 0.43 (SE= 0.03), while the mean proportion of Turkish words recalled in the second trial was 0.64 (SE= 0.04). The mean proportion of words recalled by subjects from the fifth grade sample who started with the Turkish test was 0.55 (SE = 0.03), while for those who started with the English test it was 0.44 (SE = 0.05), with the effect of language order approaching significance, $F(1, 14) = 4.05$, $p = .064$. For those fifth grade subjects who started in the NL, the mean proportion of Turkish words recalled in the first trial was 0.66 (SE= 0.04), while the mean proportion of English words recalled in the second trial was 0.44 (SE= 0.03). For those fifth grade subjects who started in the FL, the mean proportion of English words recalled in the first trial was 0.37 (SE= 0.03), while the mean proportion of Turkish words recalled in the second trial was 0.52 (SE= 0.04). On the other hand, the proportion of words recalled by subjects from the third grade sample who started with the Turkish test ($M = 0.40$, $SE = 0.03$) was lower than that of those who started with the English test ($M = 0.45$, $SE = 0.04$). Nonetheless, the effect of language order was also statistically non-significant for this sample, $F(1, 14) = 1.13$, $p > .05$.

All other interactions with language order were found to be statistically non-significant ($p > .05$).

3.4.3 Correlation between Measures of Serial Recall Performance in the Native and Foreign Languages

A highly significant correlation was found between the mean proportions of words recalled in their correct serial position in the written mode in Turkish and in English, $r = .70$ (46), $p < .01$.

Similarly, a highly significant correlation was found between the mean proportions of words recalled in their correct serial position in the oral mode for the two languages, $r = .53$ (46), $p < .01$.

The written recall means for the NL and FL tests were also found to correlate significantly at the level of each sample individually: adult, $r = .59$ (14), $p < .05$; fifth grade, $r = .58$ (14), $p < .05$; third grade, $r = .69$ (14), $p < .01$.

However, the oral recall means for the two languages were only found to correlate significantly for the third grade sample, $r = .57$ (14), $p < .05$.

CHAPTER FOUR

CONCLUSION

4.1 Discussion

In previous research using a serial recall paradigm, the effect of recall mode upon performance was reported to vary in relation to age. Whereas the level of serial recall achieved by primary schoolchildren was significantly higher in the oral mode than in the written mode, no effect of output mode was observed in adults (Bourdin & Fayol, 1994). These findings are consistent with a capacity theory of writing development (McCutchen, 1996). Each of the various processes involved in writing is believed to carry a cognitive cost. The extent to which a process consumes resources from the limited pool available in WM depends upon the degree to which it has been automated. Bourdin and Fayol have argued that low-level processes associated with the mechanics of writing are cognitively costly in young children. In other words, since spelling and handwriting are not yet fully automated in young primary schoolchildren, they require a greater level of conscious control than is the case in adults. As a result, less residual capacity remains in WM for the storage of information. This accounts for the poorer performance of children in the written mode when the same recall task is performed both orally and in writing (Bourdin & Fayol, 1994). Moreover, by implication, until such low-level transcription processes have been automated, developing writers may possess insufficient residual WM capacity to attend to higher level writing processes, such as conceptual and linguistic planning and review. Again, this may have a negative impact upon the quality of written language production (Bourdin & Fayol, 2000; 1996; 1994).

Bourdin and Fayol's study was undertaken with native speakers of French (1994). The research was subsequently replicated with native speakers of German in an attempt to investigate the cross-linguistic validity of the findings. The outcome of the German study was entirely consistent with the French results. The effect of recall mode varied as a function of age: primary schoolchildren performed significantly better in the oral recall mode than in the written mode; there was no mode effect in adults (Grabowski, 2005). The current study was undertaken with native speakers of Turkish. One objective of the research was to explore further the question of whether there might be differences between languages in terms of the relative cognitive burden imposed by low-level writing processes upon developing writers.

In the present study, in contrast to the earlier research, no significant interaction was observed between the age of the Turkish participants and the mode in which serial recall was performed. In the Bourdin and Fayol study, although adults performed systematically better in the written recall mode, the difference was not statistically significant in relation to their oral performance (1994). In the current study, however, the recall performance of adults was significantly better in the written mode than in the oral mode. More strikingly, whereas the French and German primary schoolchildren performed significantly better in the oral recall mode regardless of age group, analysis of the Turkish data revealed that the children as a whole, and also the fifth grade children when sampled separately, performed significantly better in the written recall mode than in the oral mode. The question therefore arises as to whether the finding of a written mode superiority effect in Turkish children undermines the validity of Bourdin and Fayol's thesis that graphomotor and orthographic difficulties adversely affect language production in young writers. In fact, two explanations may be offered for the Turkish results, neither of which challenges the assumption that Bourdin and Fayol's findings reflect universal cognitive processes.

Firstly, the outcome of the Turkish study is consistent with the developmental framework within which the French researchers interpreted their results. As schoolchildren mature, transcription processes start to require less conscious control. As handwriting and spelling processes begin to be executed more automatically, WM resources are released and may be diverted towards maintaining items in storage

and/or towards controlling the quality of written output (Bourdin & Fayol, 1994). For methodological reasons discussed in Section 4.3 below, the average age of the primary school participants in the Turkish study was one year older than that of their counterparts in the earlier research. Thus, the mean age of the fifth grade participants in the current study was exactly 11 years, whereas for the older primary school participants in the French experiment it was 9 years and 7 months. Bereiter and Scardamalia have reported that the superiority of the oral mode over the written mode found in young children disappears at around the age of nine or ten (cited in Bourdin & Fayol, 1996). Thus, it is arguable that the additional year of development may have accounted, at least in part, for the lack of a superiority effect for the oral recall mode in the older sample of Turkish schoolchildren.

This additional year of development may also explain why recall performance did not vary significantly as a function of age between the Turkish primary school samples. Short term memory performance is known to increase steeply up to age eight, and more gradually thereafter (Gathercole, 1999). The average age of the younger sample of Turkish schoolchildren was 8 years and 11 months. Thus, the fact that the mean age of each sample of Turkish children was higher than the threshold of eight may account for the lack of any significant difference between their levels of recall. In contrast, the mean age of the younger sample of French children was only 7 years and 6 months (Bourdin & Fayol, 1994), below the threshold of eight, which may explain why its level of recall was significantly lower than that of the older sample of French children.

Although the absence of a superiority effect for the oral mode over the written mode in the older Turkish children may be attributable to their extra year of cognitive development, the same lack of an oral superiority effect in the younger Turkish primary school participants requires further explanation. These children were on average eight months younger than the older primary school participants in the French study, in whom a significant oral superiority effect was nonetheless observed (Bourdin & Fayol, 1994). Likewise, the performance of the older children in the German study was significantly better in the oral recall mode than in the written mode (Grabowski, 2005). The second possible explanation for the distinctive nature of the Turkish findings may be raised here. Unlike French and German, in

which the matching of sounds to letters may be irregular, Turkish has a highly transparent orthographic system (Aro, 2005; Durgunoğlu, 2005; Jaffré & Fayol, 2005; Landerl, 2005; Bourdin & Fayol, 1994). As the mapping of phonemes to graphemes follows a consistent pattern in Turkish, the process of constructing spellings is relatively straightforward. It is therefore arguable that spelling processes become automated more quickly in young Turkish children than in French or German children who must learn to manipulate more complex orthographic systems. Thus, the younger Turkish children may have had to expend less mental effort upon the conscious construction of spellings than the older children in the French and German research, making the mechanics of writing relatively less difficult for them. Elsewhere, Bourdin and Fayol have reported that handwriting processes were at least partially automated in French children by the age of nine. The researchers suggested that the superiority effect which they found for oral over written recall in this age group might, therefore, have been due to difficulties with spelling rather than graphomotor execution (Bourdin & Fayol, 2000). The lack of an oral superiority effect in the younger sample of children in the current study is consistent with this hypothesis, reflecting both the partial automation of handwriting processes and the relative straightforwardness of spelling in Turkish.

In sum, the absence of an oral superiority effect in children in the current study may reflect increased levels of automaticity in low-level writing processes due to the orthographic characteristics of Turkish, or due to the age of the participating schoolchildren, or indeed due to the combined impact of these factors. However, a final question remains as to why, in the Turkish study, the recall performance of adults, and even of children, in the written mode was significantly better than their performance in the oral mode, rather than simply comparable with it. This is in marked contrast to the French and German studies, in which there was no evidence of a superiority effect for written recall, even in adults (Bourdin & Fayol, 1994; Grabowski, 2005). In fact, although the question of the impact which recall mode has upon VSTM has not yet been answered conclusively, the current finding is consistent with those reports which indicate a superiority effect for written recall over spoken recall (see Kopp *et al.*, 2006). Various explanations have been offered for this effect. Craik suggested that the superiority effect which he found for the written mode might

be due to the maintenance of rehearsal during writing, which would enable the last items in a list to be retained in memory for longer. In contrast, overt articulation during spoken recall may disrupt the subvocal rehearsal process, leading to the faster decay of phonological traces from STM and poorer oral performance (Craik, 1970, cited in Kopp *et al.*, 2006). An alternative hypothesis is that access to unrecalled items in memory may be triggered when subjects review their written output (Brimer & Mueller, 1979, cited in Kopp *et al.*, 2006). Kopp and her colleagues reported data from a delayed serial recall paradigm in which adults not only demonstrated a higher level of recall in the written mode than in the spoken mode but also claimed to find the former easier than the latter. Patterns of neural activity reported in the same study indicated that there might be a difference in the way that items are retained in written as compared to spoken recall, although this finding may have been a consequence of the visual presentation mode adopted in the experiment (Kopp *et al.*, 2006). Elsewhere Bourdin and Fayol, reporting a superiority effect for written sentence span over oral sentence span in adults, have also suggested that different encoding processes and/or mnemonic strategies may be used by adults in relation to recall mode (1996). In the latter study, in which both visual and auditory presentation were used, whereas an oral superiority effect was reported in seven year old children, no mode effect was observed in nine year olds (Bourdin & Fayol, 1996). In other words, there was progression towards a written superiority effect with increasing maturity. It appears that, subsequent to the automation of low-level writing processes, aspects of written production may assist written recall, as discussed above. The same pattern was observed in the present study, although significant changes occurred at earlier stages of development - perhaps as a consequence of the relatively straightforward orthography of Turkish. Thus, in the younger primary school sample, written recall was systematically but not significantly better than spoken recall, but by the fifth grade and into adulthood there was a consistent mode effect in favour of written performance.

A further objective of the current study was to apply the serial recall paradigm in a FL context, using English. The use of a less automated language is known to incur additional cognitive costs. Having to identify words in a FL may consume additional resources in WM; FL words function as low frequency words

and language processing is slower (Clahsen & Felser, 2006; Ardila, 2003). It was therefore anticipated that the requirement to carry out serial recall in a FL would have an adverse effect upon overall performance, regardless of production mode. This proved to be the case, in that proportion of items recalled in their correct serial position in the FL study was significantly lower than that in the NL study. This NL superiority effect was found when the overall performance of the subjects was assessed in each language; when each sample was analysed individually; and in relation to each word list, irrespective of its length. It is consistent with reports of superior recall for words in a native or dominant language, and may reflect lower levels of redintegrative support available to the less familiar language from knowledge in LTM (Thorn *et al.*, 2002; Thorn & Gathercole, 1999).

The English orthographic system is highly irregular (Landerl, 2005). In the light of Bourdin and Fayol's research, it was expected that difficulties with spelling processes might drain limited cognitive resources, leading to a decline in written recall performance, particularly in children, in whom graphomotor and orthographic processes have not been fully automated, but potentially even in adults should conditions of cognitive overload occur in the FL context (Bourdin & Fayol, 1994; 1996; 2002). However, the outcome of the study was not entirely consistent with these expectations.

As in the Turkish study, no interaction was observed between the age of the participants and recall mode when the experiment was performed in English. However, in contrast to the overall superiority effect found for written compared with spoken recall in the NL experiment, in the FL study no overall effect of either recall mode was observed. Only in the adult sample was a consistent mode effect found, in favour of written recall. The factor of age did not produce any significant difference in recall performance between the two samples of schoolchildren. As discussed previously in relation to the Turkish study, this finding might be attributable to the stage of development already reached by both samples. In the FL experiment, no effect of recall mode was observed in the schoolchildren, regardless of whether the third and fifth grade samples were analysed separately or in combination. In other words, the cognitive demands of performing the serial recall task in a less automated language with a particularly difficult orthographic system

did not result in the anticipated superiority effect for oral recall, even in the younger sample of schoolchildren.

Nonetheless, by comparing these results with data from the NL experiment, it is possible to reconcile this finding with Bourdin and Fayol's thesis that in young children the cognitive cost of managing low-level transcription processes adversely affects written as opposed to oral language production (1994). In the Turkish study, the schoolchildren as a whole and the fifth grade sampled individually demonstrated significantly higher levels of recall in the written mode. In the FL experiment, in contrast, no written superiority effect was found in children. On the contrary, the performance of children as a whole, and of the fifth grade children when sampled separately, was at the same level in the written recall mode as in the oral mode. This reduction in written recall performance cannot be attributed solely to reduced automaticity associated with undertaking the task in a FL. Had this been the case, alongside the overall drop in performance compared with the NL results, the gap between written and oral performance should also have persisted. Rather, the greater reduction found in the level of written recall suggests that when the additional costs of processing a less automated language were combined with graphomotor and orthographic costs, the resources available to children in WM for the storage of information were depleted, leading to the more marked decline in written recall performance. This interpretation is consistent with Bourdin and Fayol's argument that with increasing maturity the cognitive cost of the written mode only becomes significant when the total load weighing on WM at a particular moment exceeds the amount of cognitive resources available for deployment (1996; 1994).

Surprisingly, despite the FL context and the difficulty of English orthography, there was no significant difference between the written and oral recall performance of the younger sample of children in the FL experiment, as was also the case in the NL study. Since all participants were advised at the outset that words written incorrectly but recognisably in their original serial position would be marked as correct, it is possible that the younger subjects simply ignored irregularities in English spelling. This interpretation is borne out by the presence of numerous spelling errors in the written output from this sample. Similarly, Bourdin and Fayol have suggested that eight year old French children, who did not perform better in the

oral recall mode than in the written mode when writing irregular words in their NL, may have disregarded spelling irregularities (1994).

In the present study, as predicted and discussed above, the overall recall performance of the adult sample was weaker in English than in native Turkish. However, the finding of a consistent mode effect in favour of written recall in this group suggests that English orthography did not impose an additional cognitive cost here. Since the FL words were selected from the primary school English syllabus, it is probable that their spellings were highly familiar to university students who had been studying English for at least seven years. Thus, the superiority effect for written over oral recall observed here may derive from the same factors to which it was putatively attributed in the NL study: a lack of articulatory interference, and/or reliance upon visual output to cue recall, and/or the use of different encoding processes or mnemonic strategies in written recall. It is also feasible that the oral recall performance of the adult sample in the FL experiment may have been adversely affected by affective factors. Unlike the younger participants in the study, who were following a balanced skills, communicative English syllabus at a private school, the university students were graduates of the Turkish state school system. The nature of the FL component of the Turkish University Entrance Examination makes it inevitable that in FL teaching emphasis is placed upon the acquisition of grammatical and lexical knowledge, and upon reading comprehension and translation skills. Thus, speaking skills tend to be neglected. Most of the adult participants in the study had therefore had relatively little opportunity to speak English before they entered university. Their performance in the oral recall task may therefore have been disrupted by affective interference, such as anxiety about pronunciation (Stevick, 1999).

The analysis of serial recall performance in relation to word list length revealed that, for all samples, increasing list length resulted in a significant decrease in the proportion of accurate recall, regardless of recall mode or language. This replicates Bourdin and Fayol's finding of a list length effect in adults and children (1994). The list length effect is well-established and may be interpreted in the light of the word length effect, whereby the level of recall is higher for words containing few syllables and for words of short articulatory duration, since they may be rehearsed

more quickly than longer words (Baddeley, 2003). It follows that here the slower rehearsal of longer word lists caused some of the information held in transient memory to decay before it could be refreshed or output.

At each list length, the proportion of recall achieved by the adults in each language, and by the children in Turkish, was higher in the written mode than in the oral mode. These findings underlie the superiority effects for written recall discussed earlier. However, in the FL study, an interaction was observed between recall mode and list length, in the combined samples of children. Here, the proportion of accurate recall for shorter list lengths (3 to 5 words) was higher in the oral mode, whereas the proportion of recall for longer lists (6 and 7 words) was higher in the written mode. In the study as a whole, the overall level of recall was at its lowest in the FL, specifically for the longer list lengths. Therefore, this interaction between mode and list length suggests that, under conditions of cognitive overload associated with the FL context, relatively few children had sufficient cognitive reserves to be able to achieve recall at the longest list lengths, but that those who did took advantage of characteristics of the writing process in order to prolong rehearsal and/or to facilitate redintegration.

Finally, comparative analysis of the level of serial recall between languages yielded the anticipated evidence of a superiority effect for the NL Turkish, as discussed above. Alongside this finding, the overall mean proportion of words correctly recalled in the written mode in Turkish was found to correlate highly significantly with the same parameter in English. Likewise, the mean proportions of accurate recall in the oral mode were found to correlate highly significantly between the two languages. Furthermore, the written recall means for the NL and FL tests correlated significantly at the level of each sample individually, while the oral recall means for the two languages correlated significantly for the third grade sample. This pattern of correlation is consistent with previous neurological and biological evidence of a general capacity for bilingual processing, and suggests that the significantly superior NL recall performance observed in the present study reflects differences between the languages in the degree to which cognitive processing had been automated, rather than language-specific differences in processing (Abutalebi &

Green, 2007; Perani, 2005; Gutiérrez-Clellen *et al.*, 2004; Xue *et al.*, 2004; Service *et al.*, 2002).

4.2 Pedagogical Implications of the Study

The results of this study indicated that low-level writing processes may make no significant impact upon tasks involving language production in native Turkish by as early as the third grade at primary school. Moreover, by the fifth grade, Turkish children appeared to be able to benefit from aspects of the writing process in such a way as to enhance the quality of their task performance. Nonetheless, the FL data showed that under conditions of cognitive overload induced by working in a less automated language, the written performance of the fifth grade students deteriorated to a greater extent than their oral performance. It follows that teachers should take care that the cognitive load placed on WM by a specific activity is not excessive, if unwarranted failure is to be avoided, particularly in tasks involving written language production. This recommendation applies regardless of language context. Clearly, individual differences in WM capacity should also be taken into careful consideration when planning educational tasks (Gathercole & Alloway, 2008).

A specific objective of the current study was to elicit the implications of the findings for the teaching of EFL in Turkish primary schools. The data showed that the performance of the serial recall task in English involved additional cognitive costs, regardless of mode or age group. Furthermore, the study provided evidence that, in children up to at least fifth grade, when the additional costs of processing in English were combined with graphomotor and orthographic costs, the resources available in WM for the storage of information were depleted, leading to a decline in written performance. Again, it follows that in planning writing activities in EFL for primary school children, care should be taken to avoid unnecessary complexity. Based on the principles of the WM intervention devised in a general educational context by Gathercole and Alloway (2008), an EFL writing task might be first modelled with the teacher and then broken down into consecutive steps for implementation by the pupils; instructions should be straightforward and should be repeated where necessary; excessive use of new or unfamiliar language should be

avoided, until it has been well-practised in simpler activities; aids to memory should be made available, for instance the provision of word lists to mitigate spelling difficulties; finally, the progress of the child through the task should be continually monitored and help made available where required. Of course, the design of writing activities in many modern EFL course books is consistent with these recommendations.

In principle, the teaching of EFL requires that all four skills of reading, writing, listening and speaking should be taught. In the Turkish state school system, EFL is taught with effect from the fourth grade. However, in practice relatively little time is devoted to spoken English at any level of EFL teaching in the state school system. The outcome of the FL experiment in the current study suggests that the balance of time allocated to the teaching of specific skills at various stages of development should be reconsidered. For instance, the third grade children in the study appeared to simply ignore FL spelling irregularities, suggesting that it may be unproductive to teach English orthography at this age. On the other hand, difficulty with low-level transcription processes may have adversely affected the written performance of the fifth grade children. Therefore, if children are to fulfil their learning potential in the early stages of EFL learning at primary school, it may be more productive to focus upon oral language activities, at least until the end of the fifth grade. This does not mean that spoken English should be taught to the exclusion of the written language, but rather that the balance of emphasis should be changed. Moreover, the FL findings also reinforce the view that the tendency to rely solely or largely upon assessment by written tests may be misleading at this stage of EFL learning. A more comprehensive system of evaluation, including a significant oral component based upon an expanded oral syllabus, might more accurately reveal the extent of a child's progress in EFL in the early primary school years.

Finally, the question of whether the fourth grade is the best time to start teaching EFL may also be considered in relation to the FL findings in the present study. It has been argued here that cognitive costs associated with low-level writing processes make it advisable that greater emphasis should be placed upon oral activities in primary EFL learning. If writing skills were to be given less weight in the EFL curriculum, this would make it easier to introduce spoken English at an

earlier stage in primary education. This is the practice in the private school sector in Turkey, which is generally considered to be more successful than the state school sector in teaching EFL. Learning in general builds upon previously acquired and assimilated knowledge structures (Gathercole & Alloway, 2008). Age is known to be a key variable affecting the ultimate level of achievement in FL acquisition, particularly in relation to phonology (Oyama, 1976, cited in Miyake, 1998). Newport's "less is more" hypothesis postulates that the relatively limited WM capacity of children simplifies the internal analysis of encoded elements, making it easier for children to apprehend the internal structure of a language (cited in Miyake, 1998). Gathercole and her colleagues have demonstrated that there may be a causal relationship between PSTM and the acquisition of vocabulary in four to six year old children (cited in Gathercole, 2006). Although it remains significant, the strength of the relationship weakens by age eight, suggesting that other factors, such as long-term lexical knowledge, may play an increasingly important role in facilitating vocabulary growth at this stage of development (Gathercole, 2006). On the other hand, PSTM may continue to play a significant role in FL vocabulary acquisition over a longer period, since there may be less support available from long-term lexical knowledge (Masoura & Gathercole, 1999). Nonetheless, the repeated rehearsal of FL words and sequences of words in STM enables their eventual consolidation into LTM (Gupta & MacWhinney, 1997; Ellis & Sinclair, 1996). Furthermore, there is evidence that once a sufficient FL lexicon has been established in LTM, it is used to mediate new vocabulary learning (Gathercole, 2006, Masoura & Gathercole, 2005). Thus, the introduction of spoken English lessons at the outset of primary education in the state system would maximise the contribution of PSTM to vocabulary development and accelerate the consolidation into LTM of the lexical and sequential information upon which future language attainment depends. It would also obviate the negative impact that affective constraints arising from insufficient spoken practise may have upon oral language production in older Turkish students of EFL, including at university level.

4.3 Limitations of the Study

The current study followed the procedure of Bourdin and Fayol, who presented one word list for serial recall at each list length in each condition (1994). Arguably, the use of the standard calculation of memory span, which requires the presentation of three word lists for recall at each list length, would have minimised any potential distortion of scores due to loss of attention (Grabowski, 2005). On the other hand, it has also been reported that the repetition of tests may itself result in a loss of attention and consequent drop in performance in young subjects (Kail, 2002). Whatever the case, in the current study the single list methodology was adopted, for the specific reason that the extension of the research into a FL context narrowed the amount of vocabulary available for use with the youngest sample of participants in the English serial recall test, such that it was not feasible to run each list length in each mode in triplicate here.

A further potential limitation of the present study again derives from the age of the youngest participants. In order to find subjects with a sufficient knowledge of English vocabulary to be able to take part in the FL component of the experiment, the youngest participants were one school grade older than their counterparts in previous studies (Grabowski, 2005; Bourdin & Fayol, 1994). Since the age difference between the primary school samples was maintained, each group of Turkish children was at a higher level of cognitive development than its respective counterpart. As discussed above, this factor in combination with the relatively transparent orthography of Turkish may explain the absence of an interaction between the ages of the participants and recall mode here. It would therefore be interesting to repeat the NL trial with younger subjects in order to determine the stage of development at which the upper threshold for a superiority effect for oral recall in Turkish children is reached.

Finally, the fact that the school age samples were all being educated at a private school may mean that they are not typical of Turkish primary schoolchildren in general. It would therefore be useful to include a more mixed population of schoolchildren in any future trial.

4.4 Conclusion

In conclusion, the key findings of this study may be summarised as follows:

1. When a serial recall paradigm was used with Turkish subjects in their NL, recall mode did not interact significantly with age. Rather, the recall performance of both adults and children was found to be significantly better in the written mode than in the oral mode. In contrast, in previous research conducted with native speakers of French and German in their respective languages, the oral recall performance of children was found to be significantly better than their written performance in the same task. This led to the conclusion that, in children, graphic and orthographic processes which are not yet fully automated may deplete limited cognitive resources in WM (Bourdin & Fayol, 1994; Grabowski, 2005). The current findings are not inconsistent with such a capacity theory of writing development, since they may be attributed to the relative transparency of Turkish orthography and/or to the fact that the participating schoolchildren were one school grade older than their counterparts in the earlier research.

2. The finding of a superiority effect for written recall in this experiment suggests that, after low-level writing processes have been automated, recall may be facilitated by features of writing which are not present or available during speech.

3. When the serial recall paradigm was used with the same Turkish subjects in English, which has an irregular orthographic system, adults were again found to perform significantly better in the written mode. No mode effect was observed in children. In other words, the superiority effect for written recall found in children in the NL test was not observed here. This suggests that when the additional cognitive costs of processing a less automated FL are combined with orthographic and graphomotor transcription costs, the resources available to children in WM for the storage of information may be depleted, to the detriment of written language production.

REFERENCES

Abutalebi, J. and Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. Journal of Neurolinguistics, 20, 242-275.

Ardila, A. (2003). Language representation and working memory with bilinguals. Journal of Communication Disorders, 36, 233–240.

Aro, M. (2005). Learning to read: The effect of orthography. In R.M Joshi and P.G. Aaron (Eds.), Handbook of Orthography and Literacy (pp. 531-550). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Ashworth, J. and Clark, J. (1996). I-Spy 1. Oxford: OUP.

Baddeley, A.D. (2004). Essential Handbook of Memory Disorders for Clinicians. Chichester: John Wiley and Sons.

Baddeley, A.D. (2003a). Working memory: Looking backward and looking forward. Nature Reviews Neuroscience, 4(10), 829-839.

Baddeley, A.D. (2003b).Working memory and language: an overview. Journal of Communication Disorders, 36, 189–208.

Baddeley, A.D. (2002). The psychology of memory. In: A. D. Baddeley, B. A. Wilson & M. Kopelman (Eds.) Handbook of Memory Disorders, (2nd ed., pp. 3-15). Hove: Psychology Press.

Baddeley, A.D. (2000). The episodic buffer: A new component of working memory? Trends in Cognitive Sciences, 4(11), 417-423.

Baddeley, A.D. (1999). Essentials of Human Memory. Hove: Psychology Press.

Baddeley, A.D. (1996). Exploring the central executive. The Quarterly Journal of Experimental Psychology, 49A(1), 5-28.

Bereiter, C. (1980). Development in writing. In L.W.Gregg & E.R.Steinberg (Eds.) Cognitive Processes in Writing (pp. 73-93). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Bourdin, B. and Fayol, M. (2002). Even in adults, written production is still more costly than oral production. International Journal of Psychology, 37(4), 219-227.

Bourdin, B. and Fayol, M. (2000). Is graphic activity costly? A developmental approach. Reading and Writing: An Interdisciplinary Journal, 13, 183-196.

Bourdin, B. and Fayol, M. (1996). Mode effects in a sentence production span task. European Bulletin of Cognitive Psychology, 15(3), 245-264.

Bourdin, B. and Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. International Journal of Psychology, 29(5), 591-620.

Clahsen, H. and Felser, C. (2006). How native-like is non-native language processing? Trends in Cognitive Sciences, 10(12), 564-570.

Davies, M. (2004-). BYU-BNC: The British National Corpus. Available online at http://corpus.byu.edu/bnc

Durgunoğlu, A.Y. (2005). How language characteristics influence Turkish literacy development. In R.M Joshi and P.G. Aaron (Eds.), Handbook of Orthography and Literacy (pp. 219-230). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Ellis, N.C. and Sinclair, S.G. (1996). Working memory in the acquisition of vocabulary and syntax: putting language in good order. The Quarterly Journal of Experimental Psychology, 49A (1), 234-250.

Engle, R.W. (2002). Working memory capacity as executive attention. Current Directions in Psychological Science, 11, 19-23.

Fry, A.F. and Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. Biological Psychology, 54, 1–34.

Gathercole, S.E. (2006). Nonword repetition and word learning: The nature of the relationship. Applied Psycholinguistics, 27, 513-543.

Gathercole, S.E. (2002). Memory development during the childhood years. In A.D. Baddeley, M.D. Kopelman and B.A. Wilson (Eds.), The Handbook of Memory Disorders (pp. 475-494). New York: Wiley.

Gathercole, S.E. (1999). Cognitive approaches to the development of short-term memory. Trends in Cognitive Sciences, 3(11), 410-419.

Gathercole, S.E., Adams, A.M. and Hitch, G.J. (1994). Do young children rehearse? An individual differences analysis. Memory and Cognition, 22(2), 201-207.

Gathercole, S.E. and Alloway, T.P. (2008). Working memory and classroom learning. In K. Thurman and C.A. Fiorello (Eds.), Applied Cognitive Research in K-3 Classrooms (pp. 15-37). New York: Routledge /Taylor and Francis.

Gathercole, S.E. and Baddeley, A.D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. British Journal of Psychology, 81, 439-54.

Gathercole, S.E., Lamont, E. and Alloway, T.P. (2006). Working memory in the classroom. In S. Pickering (Ed.), Working Memory and Education (pp. 219-240). Burlington, MA: Academic Press.

Gathercole, S.E., Pickering, S.J., Ambridge, B. and Wearing, H. (2004).The structure of working memory from 4 to 15 years of age. Developmental Psychology, 40(2), 177–190.

Gathercole, S.E., Pickering, S.J., Hall, M. and Peaker, S.M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. The Quarterly Journal of Experimental Psychology, 54A (1), 1-30.

Gathercole, S.E., Service, E., Hitch, G.J., Adams, A.M., and Martin, A.J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. Applied Cognitive Psychology, 13, 65-77.

Gavens, N. and Barrouillet, P. (2004). Delays of retention, processing efficiency, and attentional resources in working memory span development. Journal of Memory and Language 51, 644–657.

Grabowski, J. (2005). Speaking, writing, and memory span performance: Replicating the Bourdin and Fayol results on cognitive load in German children and adults. In L. Allal & J. Dolz (Eds.), Proceedings Writing 2004. Geneva (CH): Adcom Productions.

Grabowski, J. (1996). Writing and speaking: Common grounds and differences toward a regulation theory of written language production. In C.M. Levy & S. Ransdell (Eds.), The Science of Writing: Theories, Methods, Individual Differences, and Applications (pp. 73-91). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Gregg, V.H., Freedman, C.M. and Smith, D.K. (1989). Word frequency, articulatory suppression and memory span. British Journal of Psychology, 80, 363-374.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. The Quarterly Journal of Experimental Psychology, 56A (7), 1213–1236.

Gupta, P. and Macwhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. Brain and Language, 59, 267–333

Gutiérrez-Clellen, V.F., Calderón, J., and Ellis Weismer, S. (2004). Verbal working memory in bilingual children. Journal of Speech, Language, and Hearing Research, 47(4), 863-876.

Hayes, J.R. (1996). A new framework for understanding cognition and affect in writing. In C.M. Levy & S. Ransdell (Eds.), The Science of Writing: Theories, Methods, Individual Differences, and Applications (pp. 1-27). Mahwah, New Jersey: Lawrence Erlbaum Associates.

House, S. and Scott, K. (2003). Story Magic 2. Oxford: Macmillan Education.

Jaffré, J-P. and Fayol, M. (2005). Orthography and literacy in French. In R.M Joshi and P.G. Aaron (Eds.), Handbook of Orthography and Literacy (pp. 81-104). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Jarrold, C. and Towse, J.N. (2006). Individual differences in working memory. Neuroscience, 139, 39-50.

Kail, R. (2002). Developmental changes in proactive interference. Child Development, 73(6), 1703-1714.

Kellogg, R.T. (2001). Competition for working memory among writing processes. The American Journal of Psychology, 114(2), 175-191.

Kellogg, R.T. (1996). A model of working memory in writing. In C.M. Levy & S. Ransdell (Eds.), The Science of Writing: Theories, Methods, Individual Differences, and Applications (pp. 57-71). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Kellogg, R.T., Olive, T. and Piolat, A. (2007). Verbal, visual, and spatial working memory in written language production. Acta Psychologica 124, 382-397.

Kopp, F., Schröeger, E. and Lipka, S. (2006). Synchronized brain activity during rehearsal and short-term memory disruption by irrelevant speech is affected by recall mode. International Journal of Psychophysiology, 61, 188- 203.

Landerl, K. (2005). Reading acquisition in different orthographies: Evidence from direct comparisons. In R.M Joshi and P.G. Aaron (Eds.), Handbook of Orthography and Literacy (pp. 513-530). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Leech, G., Rayson, P. and Wilson, A. (2001). Word Frequencies in Written and Spoken English: Based on the British National Corpus. London: Longman.

Liberman, A.M. (1998). When theories of speech meet the real world. Journal of Psycholinguistic Research, 27(2), 111-122.

Majerus, S., Poncelet, M., Van der Linden, M. and Weekes, B.S. (2008). Lexical learning in bilingual adults: The relative importance of short-term memory for serial order and phonological knowledge. Cognition, 107(2), 395-419.

Masoura, E.V. and Gathercole, S.E. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. Memory, 13(3/4), 422-429.

Masoura, E.V. and Gathercole, S.E. (1999). Phonological short-term memory and foreign language learning. International Journal of Psychology, 34(5/6), 383-388.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. Educational Psychology Review, 8(3), 299-325.

Medwell, J. and Wray, D. (2007). Handwriting: what do we know and what do we need to know? Literacy, 41(1), 10-15.

Mitchell, H.Q. and Parker, S. (2002a). Zoom In 1. London: MM Publications.

Mitchell, H.Q. and Parker, S. (2002b). Zoom In 2. London: MM Publications.

Mitchell, H.Q. and Parker, S. (2002c). Zoom In 3. London: MM Publications.

Miyake, A. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A.F. Healy and L.E. Bourne (Eds.), Foreign Language Learning: Psycholinguistic Studies on Training and Retention (pp. 339-361). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Perani, D. (2005). The neural basis of language talent in bilinguals. Trends in Cognitive Sciences, 9(5), 211-213.

Richardson , J.T.E. (2007). Measures of short-term memory: A historical review. Cortex, 43, 635-650.

Saint-Aubin, J. and Poirier, M. (1999). Semantic similarity and immediate serial recall: Is there a detrimental effect on order information? The Quarterly Journal of Experimental Psychology, 52A, 367–394.

Service, E., Simola, M., Metsänheimo, O. and Maury, S. (2002). Bilingual working memory span is affected by language skill. European Journal of Cognitive Psychology, 14 (3), 383–408.

Service, E. and Turpeinen, R. (2001). Working memory in spelling: Evidence from backward typing. Memory, 9(4/5/6), 395-421.

Smith, E.E. and Jonides, J. (1997). Working memory: A view from neuroimaging. Cognitive Psychology 33, 5-42.

Stevick, E.W. (1999). In J. Arnold (Ed.), Affect in Language Learning (pp. 43-57). Cambridge: Cambridge University Press.

Thorn, A.S.C., Gathercole, S.E., Frankish, C.R. (2005). Redintegration and the benefits of long-term knowledge in verbal short-term memory: An evaluation of Schweickert's (1993) multinomial processing tree model. Cognitive Psychology, 50, 133–158.

Thorn, A.S.C., Gathercole, S.E., Frankish, C.R. (2002). Language familiarity effects in short-term memory: The role of output delay and long-term knowledge. The Quarterly Journal of Experimental Psychology, 55A (4), 1363-1383.

Thorn, A.S.C. and Gathercole, S.E. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. The Quarterly Journal of Experimental Psychology, 52A (2), 303-324.

Worrall, A. (2005). English Adventure 2. Harlow: Pearson Longman.

Xue, G., Dong, Q., Jin, Z., Chen, C. (2004). Mapping of verbal working memory in nonfluent Chinese-English bilinguals with functional MRI. NeuroImage 22, 1-10.

APPENDIX 1

ENGLISH WORD LISTS

SET 1			
3 Word List	**4 Word List**	**5 Word List**	**6 Word List**
pizza	mountain	zebra	morning
flower	orange	ice-cream	brother
lion	Sunday	season	popcorn
	rabbit	country	pilot
		doctor	starfish
			garden

7 Word List	8 Word List	9 Word List
grandma	dolphin	dentist
farmer	autumn	midnight
jungle	children	people
water	trousers	lizard
Friday	cherry	ocean
classroom	waiter	number
hamster	city	curtain
	notebook	glasses
		coffee

SET 2			
3 Word List	**4 Word List**	**5 Word List**	**6 Word List**
monkey	forest	salad	olive
apple	tiger	Monday	planet
river	sandwich	teacher	student
	winter	spider	giraffe
		weather	summer
			baby

7 Word List	8 Word List	9 Word List
pencil	sunhat	penguin
island	turtle	breakfast
parrot	science	moustache
sister	minute	crayon
biscuit	nature	bathroom
August	lemon	uncle
dancer	woman	weekend
	driver	robber
		village

APPENDIX 2

TURKISH WORD LISTS

SET 1			
3 Word List	**4 Word List**	**5 Word List**	**6 Word List**
hafta	etek	sınav	sebze
çorap	dolap	vakit	inek
silgi	ocak	yüzük	tahta
	şeker	armut	yıldız
		güneş	sene
			vücut

7 Word List	**8 Word List**	**9 Word List**
mangal	kebap	dede
küpe	şafak	servis
yaprak	adam	kuzu
sıra	timsah	beyin
halı	soba	mutfak
köpek	köşe	kasım
yüzyıl	yağmur	mimar
	ağız	pirinç
		çeşme

SET 2			
3 Word List	**4 Word List**	**5 Word List**	**6 Word List**
masa	saat	sakal	dergi
salı	karpuz	yemek	havuz
ayak	toka	cetvel	bahar
	kitap	dünya	atkı
		eylül	hayvan
			ekmek

7 Word List	**8 Word List**	**9 Word List**
parmak	tepe	öğle
tilki	böcek	bölge
yoğurt	tören	asker
kıta	salon	tatil
mayıs	hırka	çilek
yatak	meyve	kuzen
karne	hala	bacak
	gece	civciv
		kapı

CURRICULUM VITAE

Name: Catherine

Surname: AKÇA

Birthplace: Barking, England

Date of Birth: 19th September, 1961

Marital Status: Married, with two children.

Foreign Languages: French, Spanish and Turkish.

Education:

Secondary School: Plessington High School, Rock Ferry, England.

University: BA in Business Studies (First Class), Liverpool John Moores University, England.

Chartered Institute of Personnel Management: Graduate Membership.

Postgraduate: Postgraduate Certificate in Teaching English as a Foreign Language and English for Specific Purposes, Liverpool University, England.

Employment History:

1986-1999: Merseyside County Council/ Merseyside Waste Disposal Authority, Liverpool, England (Personnel Management).

2000 to date: KAKÜV Primary School, Kars, Turkey (English Teacher).

2004 to date: Kafkas University, Kars, Turkey (Lecturer in English Language and Literature).