UNIVERSIDADE FEDERAL DE SANTA CATARINA

PÓS-GRADUAÇÃO EM LETRAS/INGLÊS E LITERATURA CORRESPONDENTE

WORKING MEMORY CAPACITY AND L2 SPEECH PRODUCTION: AN EXPLORATORY STUDY

por

Mailce Borges Mota Fortkamp

Tese submetida à Universidade Federal de Santa Catarina em cumprimento parcial dos requisitos para obtenção do grau de

DOUTOR EM LETRAS

FLORIANÓPOLIS

Agosto 2000

Esta Tese de Mailce Borges Mota Fortkamp, intitulada Working Memory Capacity and L2 Speech Production: An Exploratory Study, foi julgada adequada e aprovada em sua forma final, pelo Programa de Pós-Graduação em Letras/Inglês e Literatura Correspondente, da Universidade Federal de Santa Catarina, para fins de obtenção do grau de

DOUTOR EM LETRAS

Área de concentração: Inglês e Literatura Correspondente
Opção: Língua Inglesa e Lingüística Aplicada

Anelise Corseuil
Coordenadora

Barbara Oughton Baptista
Orientadora e Presidente

BANCA EXAMINADORA:

Laura Stella Miccoli
Examinadora

Marco Antônio Esteves da Rocha
Examinador

John Robert Schmitz
Examinador

Loni Kreis Taglieber
Examinadora

To my mother and father,
with gratitude

ACKNOWLEDGEMENTS

The completion of this research project was made possible through the help of many great souls who went a mile further just to show me the way. I am indebted to:

CAPES, for the 12-month grant at the University of Minnesota.

My colleagues in the DLLE, for allowing me to go. Marta, you're the greatest boss!

The staff at the University of Minnesota for making sure I was comfortable.

Dr. Barbara Oughton Baptista, an incredible advisor, for always believing I was able to carry out this study. Thank you for the letters of encouragement in times of trouble, for the long phone calls, for caring about my health, for seeing me to talk about this study at your place, during your summer vacation, on your birthday.

Dr. Andrew D. Cohen, for his willingness to supervise my work in the Department of ESL at the University of Minnesota and to give me the chance to learn and grow in many ways.

Dr. Elaine Tarone, for lending me her office to collect the data.

Dr. Dennis Cook, from the School of Statistics, for making me realize that if I had all the answers, I wouldn't need to carry out the study.

Dr. Marco Antonio Esteves da Rocha, for listening to me, helping with the statistics, and giving me hope back. You don't know that, but when we first met, I was about to give up this study.

Dr. Lêda Maria Braga Tomitch, for writing day in day out while I was at the U of MN, for caring about my well-being and my feeling alone, for sending me valuable material from Carnegie-Mellon University, for being company at AAAL, TESOL, for the dancing parties, the flowers when I came back...

Andre Luis, my youngest brother, for spending two months with me in the Minnesotan winter and telling me ten times a day, every day, that I would be able to finish my data collection. Thank you for being there.

Rose, Anup, and Dominic, brothers and dear sister, for making my stay in Minneapolis the most memorable. Thank you for driving me to school under –25 F, for preparing soup when I had mono, for helping with the computer, for the day I came back home. You're a lesson on friendship and I'll never forget you.

Cristina Salgado, dear sister in Minneapolis, for endless hours of talk, reading, bread, Marisa Monte, Buena Vista Social Club...

Paulo, Eduardo, Flavia, Marilena, Brazilian friends at the U of MN, for the fun we had together.

Malcolm, Blair, Frank, Janelle, American friends, for dinner, movies, music, and books. Malcolm, thank you for the homemade Christmas cake and for Sinclair Lewis.

Drs. Gloria Gil and Maria Lucia Vasconcellos, for trying to explain what it feels like to be a doctor. You're great!

Ana Cecilia and Maria Inêz, for their time, encouragement, and friendship.

Dorotheia e Karin, for forgiving my absence at important moments in their lives and being present when I needed them. Dorotheia, thank you for staying up until midnight to call me every two weeks in Minneapolis.

Patricia McGuire, for encouraging students to come participate in my study.

The students who participated in the present study, for showing up in the two data collection sessions, during the winter term. That was a hard winter, almost painful, but they were always there. No money, no pizza, only true interest. Many of them became friends.

My family in Brasilia, for the long distance love. You're always in my mind.

Finally, my husband, Eduardo, the greatest example of love and respect, for supporting me in my decision to go pursue my dream and showing me we would be able to spend ten months apart. Thank you for writing every day, for calling every Wednesday and Sunday, for making sure I had everything I needed to finish this study, for accepting me the way I am. Ours is the most beautiful story of love.

July 21, 2000

ABSTRACT

WORKING MEMORY CAPACITY AND L2 SPEECH PRODUCTION: AN EXPLORATORY STUDY

MAILCE BORGES MOTA FORTKAMP

UNIVERSIDADE FEDERAL DE SANTA CATARINA

2000

Supervising Professor: Dr. Barbara Oughton Baptista

This study investigated whether there was a relationship between working memory capacity and L2 speech production and whether this relationship was task-specific or domain-free. The participants were 13 advanced learners of English as a second language at the University of Minnesota. Participants' working memory capacity was assessed by means of the speaking span test (Daneman, 1991) and the operation-word span test (Turner and Engle, 1989), both designed in the L2. L2 speech production was elicited by means of a picture description task and a narrative task. Four aspects of speech production were assessed: fluency, accuracy, complexity, and weighted lexical density. Statistical analyses revealed that, in both tasks, working memory capacity, as measured by the speaking span test, correlates positively with fluency, accuracy, and complexity, as predicted, and negatively with weighted lexical density, contrary to predictions. The analyses also revealed that the speaking span test is a significant predictor of fluency, accuracy, and complexity in L2 speech and that it partially accounts for variation in L2 oral performance. The analyses further indicate a tendency for an interaction between pauses and hesitations, and between fluency, accuracy, complexity, and weighted lexical density. Finally, the analyses indicated that the operation-word span test suffered a methodological error and yielded inadequate data to discuss whether the relationship between working memory capacity and L2 speech production is task-specific or domain free. To explain the relationship between working memory capacity, as measured by the speaking span test, and the measures of L2 speech production, it is proposed that L2 grammatical encoding is a complex subtask of L2 speech production that requires the control and regulation of attention.

No. de páginas: 230

No. de palavras: 59,275

RESUMO

CAPACIDADE DA MEMÓRIA OPERACIONAL E PRODUÇÃO ORAL EM L2: UM ESTUDO EXPLORATÓRIO

MAILCE BORGES MOTA FORTKAMP

UNIVERSIDADE FDERAL DE SANTA CATARINA

2000

Professor Orientador: Dra. Barbara Oughton Baptista

Este estudo investiga se há relação entre a capacidade da memória operacional e produção oral em L2 e se esta relação é específica à tarefa de produção da fala ou de natureza geral, independente da tarefa que está sendo desempenhada. Os participantes deste estudo foram 13 alunos de inglês como segunda língua na Universidade de Minnesota. A capacidade de memória operacional foi medida através *do speaking span test* (Daneman, 1991) e do *operation-word span test* (Turner & Engle, 1989), ambos aplicados em inglês. Duas tarefas foram usadas para elicitar a produção oral em L2: descrição de uma gravura e narrativa. Quatro aspectos da produção oral foram medidos: fluência, precisão, complexidade e densidade lexical. Análises estatísticas mostram que a capacidade de memória operacional, quando medida pelo *speaking span test,* se correlaciona de forma positiva com fluência, precisão e complexidade e, de forma negativa, com a densidade lexical, em ambas as tarefas. As análises revelam, também,

que o *speaking span test* pode prever o desempenho oral em L2 nos aspectos de fluência, precisão e complexidade gramatical, explicando parcialmente diferenças de desempenho nestes aspectos. As análises revelam, ainda, que há uma tendência para uma interação entre pausas e hesitações, e entre fluência, precisão, complexidade e densidade lexical durante a produção oral em L2. Por fim, as análises mostram que o *operation-word span test* sofreu um erro metodológico na sua aplicação, comprometendo, assim, os dados gerados pelo teste. Consequentemente, este estudo não apresenta dados adequados para determinar se a relação entre a capacidade de memória operacional e produção oral em L2 é específica à tarefa em questão ou se é de caráter geral. Para explicar a relação entre a capacidade de memória, quando medida pelo *speaking span test*, e produção oral em L2, propõe-se que a codificação gramatical é uma sub-tarefa complexa no processo hierárquico de produção da fala que exige o controle e regulação da atenção.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Preliminaries

As an L2 learner and instructor[1], I have always been intrigued by the fact that speaking is seen, by most learners and teachers, as a hard skill to deal with. Over the years, I have heard comments, by both students and instructors, that underscore the complexities of L2 speech production. On the part of learners, typical comments include "I know the rules, and can read and write well, but when I have to speak...", and, on the part of teachers, "I'd like to make them speak faster, with fewer mistakes, the right vocabulary and intonation..." or " X speaks really well, s/he's really gifted for languages". Although this anecdotal evidence is, in great part, related to L2 learners and teachers' beliefs about speaking in an L2, it reveals, nevertheless, the feeling that those involved in the L2 learning process seem to have: that speech production in the L2 is, to a great extent and in many ways, beyond our power. Thus, intellectually motivated by my perceptions of classmates' and colleagues' attitudes towards L2 speaking, and by a personal interest in human cognition, I decided to pursue further the intricacies of L2 oral production.

Being able to convey thoughts and ideas into overt speech in a second or foreign language (L2) is the objective of most L2 learners around the world (Guillot, 1999; Hieke, 1985; Lucena, Fortkamp, Braga, & Almeida, 2000; Wiese, 1984). Speaking is the

[1] In the present study, both second and foreign language will be referred to as L2. Where necessary, a distinction will be made as to whether the context is that of second or foreign language.

primary objective of most L2 instructional programs and stands as one of the major (if not the major) factor in the evaluation of L2 competence (Lennon, 1990; Riggenbach, 1989, 1991). Yet, little is currently known about L2 oral production. Although the study of L2 speech performance has gained increased attention over the past two decades, researchers in the field of L2 acquisition and use have not reached consensus on the best ways to approach L2 speaking as an object of study, or at least in a way that yields results that prove relevant from both a theoretical and pedagogical perspective. In general, L2 speech production is poorly understood, poorly taught, poorly learned, and poorly tested.

The reason for the lack of systematic research on L2 speech production is part of a more general phenomenon stemming from research in the area of first language (L1) acquisition and use. Speaking, a core human skill that, for its uniqueness, is taken as a gift from evolution to mankind (Levelt, 1995), has not been the main focus of attention within the research program of first language processing, with studies on comprehension being far more numerous than studies on production (Bock, 1996; Crookes, 1991; Levelt, 1989). This imbalance is well justified on the grounds that, in general, comprehension can be more easily assessed than production. It seems simpler to design and apply tasks that measure comprehension processes than production processes because, for the former, the experimenter's level of control over the input is much greater. In other words, it is easier to manipulate material that will be processed for comprehension than material that will be processed for language production (Bock, 1996; Christiansen, 1999, personal communication; Engle, 1999, personal communication; Just, 1999, personal communication; Ratner & Menn, 2000; Richardson, 1999, personal communication).

The field of L2 acquisition and use tends to reproduce the overall pattern found in the L1 literature (Crookes, 1989). Thus, studies of L2 speech performance are scattered around several areas, including sociolinguistics (Butler-Wall, 1986; Ejzenberg, 1992; Olynyk, Sankoff, & d'Anglejan, 1983; Riggenbach, 1989, 1991), psycholinguistics (De Bot, 1992; Dechert, 1984 and elsewhere; Poulisse, 1999; Poulisse & Bongaerts, 1994; Raupach, 1984), testing (O'Loughlin, 1995; Shohamy, 1988, 1994), and, within the field of task-based approaches, pretask planning time (Crookes, 1989; Ellis, 1987; Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999). Due largely to differences in focus and scope, these studies have investigated different aspects of L2 speech production, from temporal variables through task structure to the effects of various planning times on oral performance. These studies have shown, among other things, that compared to L1, L2 speech presents a greater number of pauses, greater pause time, increased hesitation phenomena, and decreased speech rate (Deschamps, 1980; Olynyk, Sankoff, & d'Anglejan, 1983; Raupach, 1980); that speech performance is sensitive to context, task structure, and to the level of cognitive difficulty of the task (Ejzenberg, 1992; Foster & Skehan, 1996); and that pretask planning time improves speech performance (Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999; Skehan, 1998).

One of the perspectives from which to approach L2 speech performance is the information processing theory, the theory which guides much of the work developed in contemporary cognitive psychology. Information processing theory conceptualizes human beings as autonomous, active, and limited-capacity processors (McLaughlin & Heredia, 1996; McLaughlin, Rossman, & McLeod, 1983) who possess a multicomponent memory system (Ashcraft, 1994) consisting of at least three standard systems: sensory memory, short-term memory, and long-term memory (e.g. Atkinson & Shiffrin, 1968). An impressive amount of research carried out over the past three decades has greatly refined each of these systems and one outcome has been the proposal that human beings possess a *working memory system*, a limited-resource memory system in charge of the online processing (the work) and temporary maintenance (the memory) of information in the performance of complex tasks such as problem solving, reading, writing, and speaking (Baddeley, 1981, 1990, 1992a, 1992b, 1992c, 1999; Baddeley & Hitch, 1974; Carpenter & Just, 1989; Carpenter, Miyake, & Just, 1994; Daneman, 1991; Daneman & Carpenter, 1980, 1983; Just & Carpenter, 1992; Engle, Cantor, & Carullo, 1992; Shah & Miyake, 1999, among many others).

Working memory is at the heart of the human cognitive system. It is a computational arena in which our mental processes take place (Harrington, 1992; Just & Carpenter, 1989, 1992). These processes involve the manipulation of information as well as the temporary storage of the partial products of this manipulation for subsequent integration and completion of a goal in the performance of complex tasks, that is, of tasks that involve various sequences of goals (McLaughlin, 1987, 1998). The mental processes involved in the performance of complex tasks compete for the limited capacity of working memory, which has to be shared among the various processes and the storage of intermediate products. The limit in working memory capacity refers to how much work can be done and how much material can be temporarily maintained at a time (Ashcraft, 1994).

Researchers have consistently shown that the limited capacity of working memory differs among individuals. Thus, evidence accumulates demonstrating that individual differences in working memory capacity are related to several aspects of L1 reading comprehension (Daneman & Carpenter, 1980, 1983; Daneman & Green, 1986; Masson & Miller, 1983; Miyake, Just, & Carpenter, 1994; Tomitch, 1995, among many others), L1 writing (Benton, Kraft, Glover, & Pale, 1984), complex learning (Shute, 1991), learning to spell (Ormrod & Cochran, 1988), expert performance (Ericsson & Delaney, 1998, 1999), and L1 speech production (Daneman, 1991). The interpretation of these findings has been that individuals with a higher working memory capacity tend to demonstrate better performance on the complex tasks than individuals with a lower working memory capacity. Although there is a massive body of research investigating the role of working memory in first language processing, this research has been limited, to the best of my knowledge, to language comprehension, with only one published study, Daneman (1991), dealing entirely with speech production.

Although there seems to be a consensus on the fact that individual differences in working memory capacity can account for variation in performance in complex tasks, researchers do not agree on whether this capacity is specific to the task to which it is being correlated or a general capacity that remains the same across several tasks. Presently, there is evidence in favor of both views (Cantor & Engle, 1993; Daneman, 1991; Daneman & Carpenter, 1980, 1983; Daneman & Green, 1986; Engle & Oransky, 1999; Kylonnen & Christal, 1990; Tirre & Pena, 1992; Turner & Engle, 1989).

In the area of L2 acquisition and use, research focusing on the role of working memory is scarce, but the field is starting to acknowledge that a better understanding of the relationship between working memory capacity and L2 performance might help explain the wide range of individual differences in the level of L2 proficiency attained by adult learners[2] (Miyake & Friedman, 1998). Thus, research is starting to be developed which investigates the relationship between working memory capacity and the acquisition of native-like sensitivity to L2 linguistic cues (Miyake & Friedman, 1998), reading comprehension (Harrington & Sawyer, 1992; Berquist, 1998; Torres, 1998), and speech production (Fortkamp, 1998, 1999). The findings in the L2 area tend to reproduce those of L1, with individuals with a higher working memory capacity performing better in the tasks to which working memory capacity is correlated. The present research project draws on existing research on both working memory and speech production, in L1 and L2, to advance the proposal that one of the factors driving L2 speech performance is working memory capacity.

The main assumption of the present study is that to gain insights on the complexities of L2 oral performance, it is necessary to conceptualize L2 speaking as a cognitive activity that has to be carried out within the constraints of a limited-capacity cognitive system. In this sense, the concepts of working memory and of individual differences in working memory capacity serve as a window through which to inspect L2 speaking as a cognitive action.

1.2 Statement of purpose

The objective of the present study is to investigate whether there is a relationship between working memory capacity and L2 speech production and whether this relationship is task-specific or domain-free. Following mainstream research, working

[2] In the present study, proficiency in the L2 means "the degree of skill with which a person can use a language" (Richards, Platt, & Platt, 1993, p. 204).

memory capacity is defined as the capacity to process and store information during the performance of complex cognitive tasks (Daneman, 1991; Daneman & Carpenter, 1980, 1983; Shah & Miyake, 1999). Working memory capacity was assessed by means of the speaking span test, developed by Daneman and Green (1986) and Daneman (1991), and by the operation-word span test, developed by Turner and Engle (1989), both adapted to English as a second language. L2 speech performance was elicited by means of a picture description task and a narrative task and four aspects were assessed: fluency, accuracy, complexity, and lexical density. In the present study, speaking is defined as the ability to perform orally a picture description task and a narrative task (Bachman & Palmer, 1996). The study was carried out with 13 speakers of English as a second language, within the psychometric correlational approach.

1.3 Significance of the study

The present study contributes to existing research on individual differences in working memory capacity and on L2 speech performance in three major ways. First, except for Fortkamp (1998, 1999), no studies either in the field of working memory or L2 research have investigated the relationship between working memory capacity and L2 speech performance. Second, in the area of L2 working memory research, the present study is the first to address the debate over whether working memory capacity is task-specific or domain-free. Third, in the area of L2 speech performance, the present study is the first to address L2 speech production from an individual difference perspective. Thus, the results obtained in the present study might shed light on the discussion of how a central cognitive system, working memory, and individual variation in the capacity of this system relate to an important cognitive task, L2 speaking. The present study might also contribute to theories of L2 speech production in that it specifies the cognitive mechanisms constrained by the capacity of working memory. Finally, the present study might provide methodological suggestions as to how to best assess working memory capacity and L2 speech performance.

1.4 Organization of the thesis

In addition to the introduction (Chapter 1), the thesis consists of 6 chapters. In Chapter 2, the literature on working memory is reviewed. It starts with the historical background to the distinction between short-term and long-term memory leading to Baddeley & Hitch's (1974) seminal work that gave rise to the notion of working memory as presently conceptualized. The Chapter then focuses on the research on individual differences in working memory capacity, the discussion on the nature of this capacity, and the studies carried out in the L2 area.

In Chapter 3, the literature on L2 speech production is reviewed. The Chapter first focuses on models of speech production then on the studies judged relevant for the present study, both from a theoretical and empirical point of view.

In Chapter 4, the method used to assess working memory capacity and L2 speech production and the statistical techniques used to analyze the data are presented. This Chapter also poses the research questions and specific hypotheses investigated in the present study.

In Chapter 5, the results of the descriptive and inferential statistical analyses are presented.

In Chapter 6, the results of the present study are discussed in relation to the research questions and hypotheses posed in the method section and in light of existing research.

Finally, in Chapter 7, the main findings of the study are summarized and a final reflection on the relationship between working memory capacity and the aspects of L2 speech production investigated in the present research project is made. The Chapter also presents the limitations of the study, outlines suggestions for further research and concludes with the pedagogical implications that can be inferred from the results obtained and that can provide some answers as to why L2 speech production may, at times, seem beyond our power.

CHAPTER 2

Review of the Literature on Working Memory

The human memory system has been studied for over one hundred years now (Just & Carpenter, 1992; Ericsson & Kintsch, 1995) and the idea that it consists of an alliance of complex systems working together is one of the oldest in contemporary cognitive psychology (Baddeley, 1999). The fragmentation of human memory into two major systems--long-term memory and short-term memory--has now become, as Engle & Oransky (1999) suggest, part of our cultural knowledge. The impressive amount of research conducted on human memory has yielded a great number of cognitive models, including the working memory model.

The focus of this chapter is on the empirical and theoretical work on working memory. The chapter is divided into five sections. The first section offers the background to the distinction made by most models (Cantor & Engle, 1993) between long-term memory and short-term memory, along with an overview of the research on these two systems, outlining how research on short-term memory led to the development of the construct of working memory. The second section focuses on the seminal work of Baddeley & Hitch (1974), which is a precursor of the concept of working memory as presently understood. Section three focuses on working memory from an individual differences perspective and reviews studies found relevant to the present research project. Section four presents an overview of the debate over whether working memory capacity limitations are task-specific or domain-free. Finally, section five reviews the few studies carried out on the relationship between working memory capacity and L2 acquisition/use.

2.1 Long-term memory and short-term memory: background and overview

Evidence for the existence of distinct memory systems was first presented by Brown (1958, cited in Baddeley, 1990) and Peterson and Peterson (1959, cited in Baddeley, 1990), who showed that information is forgotten within seconds if rehearsal is prevented. This finding stimulated a great deal of further research, including the development of the short-term forgetting paradigm, based on the Brown-Peterson distractor technique (Baddeley, 1999, 1990) and on the assumption that certain kinds of task reflect the work of a short-term memory with limited capacity, somewhat distinct from the system responsible for long-term learning (Broadbent, 1958, cited in Baddeley, 1990; Baddeley, 1999, 1990).

Evidence for this major distinction comes from a variety of sources. First, there are the classic studies carried out under the free-recall paradigm, in which participants were presented with lists of unrelated words and asked to recall as many words as possible in any order. When recall was immediate, these studies consistently showed a recency effect[1]. When recall was delayed, the recency effect disappeared, which led researchers to suggest that for immediate recall, items were maintained in a kind of temporary storage, while for delayed recall, items were retrieved from a long-term store. Second, there are the experiments in coding during immediate and delayed recall (Baddeley, 1992). Conrad and Hull (1964), for instance, showed that acoustic similarity of the material to be learned severely interfered with immediate recall, whereas Baddeley (1966a) showed that this interference seemed to decrease if material was similar in meaning. Baddeley (1966b) also showed that under delayed recall conditions, similarity

---

[1] According to Baddeley (1990:31), recency effect refers to the enhanced recall of the most recently presented items.

in meaning becomes a crucial factor affecting performance. Thirdly, early neuropsychological research (e.g., Baddeley & Warrington, 1970) showed that patients suffering from classic amnesia with severe long-term learning deficits demonstrated no impairment in short-term learning. This research showed that the reverse may occur as well: Shallice and Warrington (1970) for example, described patients with defective short-term memory, but normal long-term learning.

It was with the work of Atkinson and Shiffrin (1968), however, that the distinction between long-term memory and short-term memory became classical and consequently a major feature of information-processing theories (Cantor & Engle, 1993; Ericsson & Kintsch, 1995; Harrington, 1992). Technically speaking, long-term memory is the term used by experimental psychologists to refer to previously learned knowledge that is stored for considerable periods of time (Baddeley, 1999) and can be retrieved during performance (Ashcraft, 1994; Baddeley, 1990 and elsewhere; Cantor & Engle, 1993; Haberlandt, 1994; Searleman & Herrmann, 1994, among many others). Over the past 30 years, an enormous amount of research has been developed on the ways long-term memory can be conceptualized, but there are currently three main approaches to the study of the system responsible for long-term retention of information.

In the first approach, long-term memory is dichotomized into episodic memory and semantic memory, a distinction introduced by Tulving (1972, 1983, and 1985). Semantic memory is our storage of generic information about the world and contains our knowledge about facts of nature, things we have learned at school, and our mental models of the world. Episodic memory, on the other hand, stores the information we acquire through personal events and is thus the store of our autobiographical data.

In the second approach, long-term memory is subdivided into declarative memory and procedural memory. The declarative-procedural distinction was based on the work of the philosopher Ryle (1949, cited in Ellis, 1993), who first noted that our knowledge may be distinguished in terms of "knowing about" and "knowing how". Within modern experimental psychology, these terms were widely adopted after the work of Anderson (1983, 1995). Declarative memory is a static fact-like kind of memory, generally characterized as dealing with knowledge that can be acquired explicitly, is accessible to consciousness, and is most times verbalizable (Haberlandt, 1994; Searleman & Herrmann, 1994; Stillings, Feinstein, Garfield, Rissland, Rosenbaum, Weisler, & Baker-Ward, 1987).

In contrast, procedural memory consists of our knowledge of how to do things, that is, the operations involved in carrying out activities, including mental processes. As a rule, procedural knowledge is not available to consciousness and thus can not be verbalized. Declarative and procedural memory are also distinct in terms of the speed with which we use them. Procedures are retrieved and used much faster than declarative memory, as shown by studies in recognition and sentence verification. For instance, it takes about 1/10 of a second for us to recognize a letter and over 1 second to verify whether a sentence like 'a canary is a bird' is true or false (Haberlandt, 1994).

Finally, the third approach to long-term memory conceptualizes the system in terms of explicit and implicit memory and originated in neuropsychological research (e.g. Graf & Schacter, 1985, 1987). Explicit memory consists of knowledge that can be consciously recollected (Graf & Schacter, 1985, p. 501), whereas implicit memory consists of knowledge that, when used, does not allow a conscious experience of knowing, perceiving, or remembering this knowledge (Schacter, 1987).

The second major human memory system--short-term memory--has also been the subject of extensive research, developed mainly in the 1960s and early 1970s, as pointed out by Engle and Oransky (1999), Logie (1996), and Richardson (1996a). Among a number of short-term store models that emerged to explain the nature of forgetting, Atkinson and Shiffrin's (1968) came to be the most influential one, according to many researchers (e.g., Baddeley, 1986, 1990 and 1999; Cowan, 1988; Engle & Oransky, 1999; Harrington, 1992; Richardson, 1996a; Shah & Miyake, 1999, among many others).

This was, in fact, a general model of memory and included a sensory memory, a short-term store, and a long-term store. Atkinson and Shiffrin's three-stage model assumed that incoming information first entered sensory memory, then proceeded to the short-term store, and finally reached long-term memory. The most important component of the model was the short-term store, which was conceptualized as a unitary system of limited capacity and a necessary step in both the acquisition and use of information. In their proposal, short-term memory could maintain information by a control process consisting of rehearsing the last few items presented. In Atkinson and Shiffrin's view, the capacity of the short-term store had to be shared between a variety of other control processes besides rehearsing, and these included coding procedures and search strategies.

Thus, for Atkinson and Shiffrin (1971), short-term memory, equated with consciousness, was the arena of control processes:

> Because consciousness is equated with the short-term store and because control processes are centred in and act through it, the short-term store is considered a working memory: A system in which decisions are made, problems are solved and information flow is directed. (Atkinson & Shiffrin, 1971, p. 83)

Among the control processes claimed to take place within the short-term store, rehearsal was the most focused on. In their view, the probability of an item being transferred to long-term memory was greater if this item was kept longer in the short-term store by means of rehearsal. Although aware of the existence and relevance of the semantic aspects of the material to be maintained, Atkinson and Shiffrin (1968) were concerned mainly with rote rehearsal.

Although able to account for a variety of data, the Atkinson and Shiffrin model (1968) was particularly deficient in dealing with the evidence from patients with short-term memory problems who had long-term store unimpaired and could thus lead a normal life. In addition, the assumption made that the longer an item is maintained in short-term memory the greater its possibility of being transferred to long-term memory seemed to be falsified by a number of studies showing that previous rote repetitions did not necessarily bring about learning (e.g., Tulving, 1966; Craik & Watkins, 1973).

Noticing that current multi-store models did not account for neuropsychological evidence from patients with short-term impairments who could nevertheless perform complex cognitive activities, Baddeley and Hitch (1974) decided to investigate the issue by using the concurrent task, a dual-task technique in which participants were asked to remember a digit string of up to six items while performing a cognitively demanding task. According to Atkinson and Shiffrin (1968), performing the digit task would occupy most of the participants' system capacity, leading to great impairment in the performance of the cognitively demanding task. Baddeley and Hitch found, in fact, that the concurrent digit span task did interfere in the cognitive task. However, as Baddeley (1992) puts it, "the degree of disruption was far from being catastrophic" (p. 284). By means of the concurrent task, Baddeley and Hitch hypothesized that if performance on one task did not interfere with the other, then the tasks relied on different components of the cognitive system. In contrast, if one task did interfere with the other, then the same cognitive pool was being used. They thus decided to propose a multicomponent model of short-term store, which they termed working memory, abandoning then the idea of a unitary short-term memory system.

2.2 Working memory

Any discussion of the concept of working memory must start with Baddeley and Hitch's (1974) model. Although, as Richardson (1993, 1996a) and Shah and Miyake (1999) note, the term can be found in early work developed by Miller, Galanter, and Pribram (1960), the connotation with which the phrase "working memory" is used nowadays was first introduced by the multicomponent model proposed by Baddeley and Hitch (1974).

Baddeley and Hitch's (1974) original model consists of a central executive, which functions as a limited capacity attentional controller, and two subcomponents--also originally called "slave systems" (Baddeley, 1990, 1992a, 1992b, 1992c, 1996, 1999; Baddeley & Logie, 1999)--the phonological loop, responsible for storing and manipulating speech-based information, and the visuospatial sketchpad, which controls visual and/or spatial material (Baddeley & Logie, 1999:29).

The central executive is responsible for (1) controlling and coordinating the two subsystems, (2) focusing attention, (3) switching attention, (4) activating representations in long-term memory, and (5) controlling and integrating actions and activities (Baddeley, 1996; Baddeley & Logie, 1999). Baddeley (1990 and elsewhere) and Baddeley and Logie (1999) sustain that the central executive is the least studied of the three components, with research in the model concentrating on the subsidiary systems, which, in their view, deal with more tractable problems.

In discussing the central executive, Baddeley (1990) suggested that this construct is similar to the model proposed by Norman and Shallice (1986), the supervisory attentional system (SAS), which offers an account of the control of action. Briefly, Norman and Shallice (1986) assume that actions are carried out through the activation of schemas. In the case of well-learned actions, schema activation will take place automatically and will initiate a sequence of actions. Several schemas can be activated simultaneously (e.g., talking and driving at the same time), but there may be occasions in which two concurrent activities will conflict with one another and priority will have to be given to one over the other. To control the activation of schemas and resolve conflicts, Norman and Shallice propose that action goals can enhance activation of some schemas and inhibit activation of others. This process is relatively automatic. The SAS is a limited capacity attentional system that aids in the selection of the most appropriate schemas for action, thus interrupting behavior when activated schemas are incompatible with current goals. The activities performed by the SAS--planning, initiating activities, monitoring the selection of schemas, controlling conflicts between action goals, and controlling the activation of well-learned schemas--are also taken to be performed by the central executive (Baddeley, 1990; Smyth, Collins, Morris, & Levy, 1994).

Since its original formulation, the central executive--the most important component of the model--has undergone several changes. In their first proposal, Baddeley and Hitch (1974) assumed that some of the limited processing capacity of the central executive could be shared with temporary storage of information (Baddeley, 1981). In a recent publication, however, Baddeley and Logie (1999) state that any storage beyond that performed by the phonological loop and the visuospatial sketchpad is possible only via accessing long-term memory or other (unknown) subsystems.

The phonological loop is assumed to comprise two components--a phonological store, which holds speech-based information that decays with time, and an active rehearsal process (Baddeley & Logie, 1999), which recycles the decaying information in the phonological store. Baddeley (1990, 1992a, 1992b, 1992c) argues that auditory spoken information has automatic and obligatory access to the phonological store.

The phonological loop has been able to account for a variety of laboratory findings. Thus, for instance, the phonological similarity effect, in which memory span is affected by the similarity in articulatory or sound features of the material being presented, is interpreted as strong evidence that the phonological loop is based on a phonological, rather than semantic, code (Baddeley, 1990). Similarly, as shown by Salamé & Baddeley (1982), the unattended speech effect seems to disrupt immediate recall. This effect is obtained by presenting participants with material to be immediately recalled accompanied by irrelevant speech, which gains access to the phonological store, thus affecting performance.

Although there is sound evidence for the existence of a system within working memory that deals specifically with speech-based material, it is not clear what use the phonological loop might have in the acquisition, comprehension, or production of language. Some theorists (e.g., Ashcraft, 1994) tend to equate the phonological loop to the view of the short-term store depicted in classical memory span tasks--a passive system for the storage of information. Baddeley and colleagues, however, have attributed greater importance to this component since it seems to play a principal role in vocabulary learning in first language (Baddeley, Papagno, & Vallar, 1988; Gathercole & Baddeley, 1989, 1990), in a foreign language (Papagno, Valentine, & Baddeley, 1991), and in the comprehension of discourse (Baddeley & Wilson, 1988). With regard to native and foreign language vocabulary learning at the practical level, however, Baddeley, Papagno, and Vallar (1988) and Papagno, Valentine, and Baddeley (1991) are careful to say that the phonological loop seems to have no relevant participation in the establishment of semantic values, and that individuals are likely to compensate for short-term phonological constraints by making use of the semantic aspects of the material to be learned.

The other subsidiary system, the visuospatial sketchpad, is in charge of maintaining and manipulating visual and/or spatial information. The amount of evidence for this particular system has been far less than for the phonological loop. The visuospatial sketchpad has recently been subdivided by Logie (1995) into a passive visual cache, which retains visual patterns, and a spatially based system, the inner scribe, which retains sequences of movements. Evidence for this fractionation comes from studies demonstrating a disruptive effect of concurrent movements on the storage of spatial patterns (e.g. Logie, Zucco, & Baddeley, 1990) and of concurrent irrelevant visual material on the storage of visual information (e.g. Logie, 1986).

The Baddeley and Hitch (1974) model is clearly a more complex and elaborate proposal than that of earlier unitary models. Contrasting it to Atkinson and Shiffrin's (1968) influential model, the crucial difference is that Baddeley and Hitch emphasize a multicomponent system consisting of at least three components, each of these being further subdivided. In addition, and most importantly, they also emphasize the efficiency of the system, which is able to deal with a multitude of different types of information from the environment, thus adding a dynamic feature to the model.

After Baddeley and Hitch's proposal, a new paradigm of research on human memory was established. Short-term memory, as it has been described in current cognitive psychology, is now assumed to be a fragment of a more powerful system responsible for on-line cognition (Baddeley & Logie, 1999), for the coordination of moment-to-moment processing, for our mental work: working memory. In one of their latest publications, Baddeley and Logie (1999, pp. 28-29) offer the following definition of working memory:

> Our own definition of working memory is that it comprises those functional components of cognition that allow humans to comprehend and mentally represent their immediate environment, to retain information about their immediate past experience, to support the acquisition of new knowledge, to solve problems, and to formulate, relate, and act on current goals.

Two decades after the seminal work of Baddeley and Hitch (1974), working memory has become, as Miyake and Shah state (1999, p. xiii), "one of the hottest topics in cognitive psychology and cognitive neuroscience". Two decades of research resulted in a number of models and interpretations of the construct "working memory". For instance, one can find in Miyake and Shah (1999) a discussion of 10 different models of working memory, representing diverse approaches to the study of the system and thus varying in theoretical scope and emphasis. The 10 models presented in Miyake and Shah also differ in their conceptualization of the nature, structure, and functions of working memory.

These diverse approaches to working memory, although positive to the growth of research, many times reflect conflicting views. For instance, although researchers generally agree that working memory refers to the system--or mechanism--responsible for the temporary storage and processing of information required for the performance of everyday cognitive tasks such as learning, language comprehension, language production, reasoning, thinking, problem solving, and decision making (e.g., Baddeley, 1900 and elsewhere; Cantor & Engle, 1993; Carpenter & Just, 1999; Carpenter, Miyake, & Just, 1994; Cowan, 1999; Daneman, 1991; Daneman & Carpenter, 1980, 1983; Engle, Cantor & Carullo, 1992; Just & Carpenter, 1992, 1996; Masson & Miller, 1983; Miyake & Shah, 1999; Richardson, 1996a, 1996b, 1993; Tirre & Peña, 1992; Towse & Hitch, 1995; Turner & Engle, 1989, among many others), there seems to be no agreement as to whether working memory consists of various components or is a unitary system and whether the capacity of the system varies as a function of the task being performed or is a stable factor. In reading the existing literature, a researcher interested in working memory will encounter a number of metaphors trying to represent the concept, some of which being "the interface between memory and cognition" (Baddeley 1992a), "a mental workspace" (Stoltzfus, Hasher, & Zacks, 1996), "an arena of computation" (Baddeley & Hitch, 1974; Just, Carpenter, & Hemphill, 1996), "the hub of cognition" (Haberlandt, 1994), "the blackboard of the mind" (Just & Carpenter in Goldman-Rakic, 1992). In all these, it seems to be clear that working memory is taken as the center where cognitive action takes place.

In spite of the various senses in which the notion of working memory has been used, it is possible to distinguish between two different but complementary approaches to the construct (Baddeley 1992a, 1992c). Both approaches make use of the dual-task methodology[2]. The first, uses the dual-task methodology and neuropsychological evidence from brain-damaged patients to verify the structure of working memory. This is the approach taken by Baddeley. The other approach, the psychometric correlational, emphasizes the correlation between individuals' working memory capacity and their performance on complex cognitive skills. The methodology of the latter approach consists of correlating performance on dual tasks with performance on demanding cognitive tasks such as reading comprehension. Results of studies of this type have consistently shown that individual differences in working memory capacity predict performance on complex cognitive tasks.

While Baddeley and those following the dual-task and neuropsychological approach focus on the structure of the working memory system, emphasizing its

---

[2] The dual-task methodology consists, of asking participants to hold sequences of numbers while also performing a reasoning, learning, or comprehension task (Baddeley, 1990, p.95).

subcomponents, researchers working within the psychometric correlational approach are mainly concerned with the central executive--the system in charge of the coordination of mental resources applied to the processing and integration of information--rarely making any explicit reference to subsidiary systems (see, for instance, Just & Carpenter, 1992). The present study has been carried out under the psychometric correlational approach. The remainder of this chapter, thus, focuses on the literature on working memory from an individual differences perspective.

2.3 Individual differences in Working Memory

As already pointed out, there is a consensus among theoreticians that working memory is involved in the performance of a number of cognitive tasks (e.g., Carpenter & Just, 1999; Engle, 1996; Engle & Conway, 1998; Gilhooly, 1998; Just & Carpenter, 1992; Jurden, 1995, among many others). Recall from section 2.2 that working memory is a limited capacity system in nature. However, the psychometric correlational approach postulates that this limited capacity is different among individuals and that these differences are good predictors of performance on important cognitive tasks, individuals with larger working memory capacity performing better on these tasks than individuals with smaller capacity.

The research on individual differences in working memory capacity has been most extensively developed in the area of first language reading comprehension, according to Baddeley (1999). In the area of first language, Daneman (1991), reviewed below in this chapter, is to my knowledge the only study investigating the relationship between working memory capacity and production.

As Engle and Oransky (1999) claim, research on individual differences in working memory capacity originates from studies showing that, differently from traditional short-term memory measures such as the digit span or the word span, measures of working memory span correlate with performance on complex cognitive tasks. As a rule, tasks assessing working memory span require an individual to hold a small amount of information in mind while concurrently carrying out other operations that generally require, in turn, manipulation of information (Swanson, 1993). In contrast, tasks assessing short-term memory span require a person to hold passively a small amount of information and then reproduce it without further manipulation or transformation.

In the field of first language comprehension, many researchers (e.g., Kintsch & van Dijk, 1978; Perfetti & Lesgold, 1977, cited in Daneman & Merikle, 1996) believed that the source of individual variation in reading performance was short-term memory capacity (measured by means of digit or word span tests). However, they successively failed to find a correlation between these measures of short-term memory and performance in reading comprehension. Daneman and Carpenter (1980) showed that the problem was in the measure used, which reflected passive storage. In their view, reading required both storage and processing of information, two functions Baddeley and Hitch (1974) had attributed to the central executive. They then showed that a test that could assess these two functions during reading comprehension was a better predictor of reading performance than a test that assessed only passive storage. This test was the reading span test.

In its original design, the reading span test (Daneman & Carpenter, 1980) requires a person to use both functions of working memory during reading comprehension: the processing component is sentence comprehension while the storage component is maintaining and retrieving the final word of each sentence of a presented set. A person's reading span is the maximum number of sentence-final words recalled in the order they were presented and is taken as index of his/her working memory capacity. In Daneman and Carpenter's view, the processing and storage functions of working memory compete for its limited capacity. The more an individual uses his/her available resources for processing, the less he/she will have available for storage. Thus, the hypothesis underlying the reading span test is that an individual's higher reading span is due to his/her more efficient reading comprehension processes, which enable him/her to leave a greater amount of their resources free for storage of sentence-ending words. In fact, they were able to find a significant correlation between the reading span test and performance on a global test of language comprehension, the Verbal Scholastic Aptitude Test (VSAT), as well as two components of reading comprehension--fact retrieval and pronominal reference--which made the reading span test a good predictor of reading and led researchers to abandon the digit, letter, and word span tests as predictors of individual variation in reading performance.

Reading is an activity that draws heavily on working memory. Roughly speaking, readers will need both functions of the system to compute semantic and syntactic information from successive words, phrases, and sentences while also holding the intermediate products of this computation for subsequent integration and construction of a coherent representation of a passage (Just & Carpenter, 1992; Daneman & Carpenter, 1980, 193; Daneman & Merikle, 1996). During reading, the limited capacity of the system needs to be shared between the work and the memory (Daneman & Merikle, 1996) and the reading span test is assumed to capture how individuals coordinate these two activities.

The reading span test has been the basis for most of the research on individual differences in working memory capacity and reading comprehension, being extensively used as a predictor of performance in various reading skills, including (a) the ability to detect inconsistencies in sentences with homonyms (Daneman & Carpenter, 1983), (b) the ability to make inferences of ideas not explicitly mentioned in the text (Masson & Miller, 1983), (c) the ability to make use of contextual cues to infer the meaning of new words in the text (Daneman & Green, 1986), (d) the resolution of lexical ambiguity in reading (Miyake, Just, & Carpenter, 1994), and (e) the perception of text structure (Tomitch, 1995). Various researchers (e.g., Turner & Engle, 1989; Masson & Miller, 1983) have also found strong correlations between the reading span test and standardized measures of reading ability such as the Verbal Scholastic Aptitude Test (VSAT) and the Nelson-Denny reading test.

However, the correlation between working memory capacity and higher-level cognitive tasks has also been consistently shown in other skills. As pointed out by Engle (1996), individual differences in working memory capacity have been shown to be significantly related to learning to spell (Ormrod & Cochran, 1988), following directions (Engle, Carullo, & Collins, 1991), vocabulary learning (Daneman & Green, 1986), notetaking (Kiewra & Benton, 1988), writing (Benton, Kraft, Glover, & Plake, 1984), and complex learning (Shute, 1991). These studies have used either the reading span test or adaptations following the principle of a good working memory span test: simultaneous processing and storage of information.

Of particular relevance to the present study are two studies carried out by Daneman & Green (1986) and Daneman (1991), which investigated whether individual differences in working memory could account for variation in speech production. Daneman (1991) was based on an earlier study, Daneman and Green (1986), which showed that individuals vary in the speed with which they can retrieve appropriate words to convey their ideas. In that study, Daneman and Green (1986) argued that speaking is a cognitively demanding task involving complex coordination of processing and storage of information. Among the processes required in speech production, selection of context-appropriate words from our mental lexicon is seen by the authors as a fundamental one.

Based on the work of Goldman-Eisler (1968), Clark and Clark (1977), and Tannembaum, Williams, and Hillier, (1965), Daneman and Green (1986) argued that pauses during speech production occur more frequently before the first content word of a phrase or sentence and indicate that speakers start the execution of speech before finishing planning. In Daneman and Green's (1986) view, it is during the planning stage that words are accessed and selected. They, thus, suggest that accessing words in real time during speech production is related to a speaker's ability to coordinate planning and execution of speech and that this coordination taxes working memory. To measure working memory during speech production, Daneman and Green (1986) proposed the speaking span test, which, like the reading span test, involved simultaneous storage and processing of material.

The speaking span test focuses on sentence level speech production processes. Daneman and Green (1986) presented 34 participants with increasingly longer sets of unrelated words, which they had to read silently. Each word was seven letters in length and was individually presented for 1 second on a computer video screen. Words were organized in five sets each of two, three, four, and five words, with intervals of 10 milliseconds between the words of each set. At the end of the set, when words were no longer visible on the screen, participants were required to produce aloud a sentence for each individual word presented. There were no restrictions on the length of the sentences or the position of the word within the sentences, but they had to be grammatical as regards syntax and semantics. In addition, there were no restrictions on the time participants took to generate the sentences. Speaking span, which is assumed to measure working memory capacity, was operationalized in terms of “maximum set size” and “total performance”. The maximum set size measure consisted of the highest set size in which a participant was correct on three out of the five sets. The total capacity measure consisted of the total number of sentences participants were able to make. In addition to the Speaking Span Test, Daneman and Green (1986) also applied the reading span test, following the same procedures of previous studies by Daneman and Carpenter (1980. 1983).

Daneman and Green (1986) correlated the results of the speaking span test to the participants' performance on a contextual vocabulary task, consisting of completing a sentence fragment with an appropriate word under time pressure. The authors report that small speaking span participants were less fluent (slower) in generating (accessing and retrieving) a context-appropriate word. They also found a significant correlation between the speaking span test and the reading span test, which led them to suggest that the two tests tap similar systems, with the speaking span test being a better predictor of word production.

Following Daneman and Green (1986), Daneman (1991) investigated the relationship between the speaking span test and first language speech production at the discourse level and articulatory level. Claiming that speaking is a complex cognitive task that requires coordination of storage and processing of information in the various stages of the speech production process, Daneman hypothesized that individuals with a larger working memory capacity--the capacity to store and process information in real time--would perform better on three tasks measuring what Daneman calls fluency in speech production, viz. a speech generation task, an oral reading task, and an oral slip task. In addition, Daneman hypothesized that the reading span test would correlate with the oral reading task, but not as strongly as the speaking span test. The study was carried out with 29 English L1 university students. Participants were given a set of experiments assessing both working memory capacity--the speaking span test and the reading span test--and language production.

The speaking span test was constructed like the one used in the 1986 study, but this time it consisted of 100 words organized in increasingly longer sets of two to six words. Speaking span was operationalized in terms of total capacity, that is, the total number of words for which a sentence was produced. This total capacity was expressed in two speaking span scores: speaking span strict, counting only those sentences with the exact form of the word presented, and speaking span lenient, counting also sentences containing the word in a different form. The reading span test was also constructed with 100 sentences and organized in sets of two to six sentences. A participant's reading span score was his/her total performance on the test.

Speech production was assessed in terms of a speech generation task, an oral reading task, and an oral slip task. The speech generation task was used as a general measure of fluency in first language production and consisted of a picture description. Participants were instructed to speak about the picture for 1 minute, as fluently and continuously as possible. The measure of fluency in this task was number of words completely articulated. Participants' speech samples were also evaluated in terms of richness and originality of content by independent raters.

The oral reading task required participants to read aloud a 320-word passage taken from *The Great Gatsby*. Participants were explicitly instructed to read the passage as fast as they could, their reading times measured in seconds with a stopwatch. In addition to reading time, the main measure, emphasis was also given to accuracy in the articulation of words.

Finally, the oral slip task consisted of eliciting spoonerisms[3] in the laboratory. Based on Motley and Baars' (1976) SLIP technique, Daneman devised a list of 309 pairs of words, which were presented one by one on a computer video screen, for 900 milliseconds each pair. The interval between one pair and the next was 100 milliseconds. From the 309 word pairs, 69 were cued via a beep. Participants were instructed to attend to all pairs and to say each cued pair as soon as they heard the beep, which was sounded 500 milliseconds after the removal of the pair, thus entering 400 milliseconds into the presentation of the next pair. The other non-cued pairs were read silently. From the 69 cued word pairs, 39 were filler pairs aimed at disguising 30 target pairs aimed at eliciting spoonerisms. Each of the 30 target word pairs was immediately preceded by three phonological interference word pairs, aimed at inducing the spoonerism error. The interference word pairs were phonologically similar to the spoonerism error participants were expected to make. The 39 filler pairs were preceded by a different number of word pairs without the phonological interference feature.

Daneman (1991) reports finding similar means between scores of the speaking span test and the reading span test, which she interprets as evidence that the two tasks tap a common limited system. However, she does not report testing for correlation between the two measures. She also reports that the two span tasks differed in their power of predicting performance in language production: The speaking span test correlated significantly with the three speech production tasks, whereas the reading span task correlated significantly only with the oral reading task. Furthermore, the strict and the

---

[3] A type of speech error in which the position of sounds, words, or sentences is reversed: e.g., *queer dean* for *dear queen* (Richards, Platt & Platt, 1992; Daneman, 1991).

lenient scores of the speaking span test differed in the aspects of fluency they could predict. Speaking span strict correlated better with the oral reading task and the oral slip task, whereas speaking span lenient correlated significantly with the two measures of the speech generation task--words per minute and subjective ratings. The findings Daneman obtained are consistent with Daneman and Green's (1986) previous finding that working memory capacity is involved in speech production and stands as an important source of individual differences in the ease of access and storage of the information necessary in each of the stages of the process.

After Daneman and Green (1986) and Daneman (1991), the speaking span test is claimed to be a complex measure of working memory span for language production, which taxes both the storage and processing functions of this limited capacity system during speech production. While the storage component of the test is to recall the words presented, the processing component consists of generating grammatical sentences containing these words. Both functions compete for the limited capacity of the system. Daneman (1991) argues that the ability with which an individual coordinates storage and processing in this task is related to his/her ability to produce fluent speech, which also requires efficient coordination of storage and processing of information.

The speaking span test and the reading span test are recall tests devised to measure working memory capacity under language production or comprehension processing demands. As claimed by Daneman and Carpenter (1980, 1983), good readers have a larger working memory capacity for storing products of the reading comprehension process--such as facts, pronoun referents, and propositions (Turner & Engle, 1989)--because their processing during reading comprehension is efficient, thus leaving more of their working memory capacity available for storage. On the other hand, less skilled readers have a smaller working memory capacity, as measured by the reading span test, because they devote more capacity to processing, which results in less capacity for storage of data. By the same token, Daneman (1991) argues that more fluent L1 speakers have a larger working memory capacity as measured by the speaking span test because their speech production processes are more efficient, leaving greater resources available for the storage of intermediate products of this processing. These claims reflect the processing efficiency view of working memory capacity.

At this point, two observations need to be made. The first one is that, compared to reading comprehension, it seems more complex to determine what is processed and what is stored during speech production. Thus, Daneman (1991) does not specify what the intermediate products referred to above might be. The second observation is that the speaking span and reading span tests do not measure processing efficiency per se; rather, they are assumed to reflect the storage capacity an individual has left as a result of his/her processing efficiency while producing or comprehending language. As claimed by Daneman and Green (1986) and Daneman (1991), working memory capacity is task specific and varies as a function of how efficient a person is at the task in which working memory is involved (Daneman & Green, 1986).

It is a controversial issue, in studies on working memory, whether the relationship between the limited capacity of the system and performance on complex cognitive tasks is a task-specific factor that varies as a function of the individual's processing efficiency or a general, domain free factor stable across tasks. The next section reviews the recent literature on current theorizing on the nature of the relationship between working memory capacity and higher-order cognition.

2.4 The relationship between working memory capacity and higher-order cognition: task-specific or domain free?

As indicated above, the task-specific view of working memory capacity stems from the work of Daneman and Carpenter (1980), who first showed that individual differences in this system's capacity is related to reading comprehension. They argued that the limited resources of working memory have to be shared between processing and storage demands during reading comprehension. In their view, individuals differ in their ability to coordinate processing and storage demands during reading comprehension. Efficient readers have a higher span, as measured by the reading span test, because they can allocate more of their available resources to storing the to-be remembered items and less to the processes themselves (Daneman & Merikle, 1996). Daneman and Carpenter also argued that the capacity of working memory is functional and varies according to the individual's processing efficiency in a given task. The implication of the task-specific view is that the processing component of the span test must require the same processes present in the cognitive task whose performance is being predicted.

Further elaboration of the task-specific view led Daneman and Tardiff (1987) to argue that individual differences in working memory capacity can be measured through processing efficiency alone, without including a simultaneous storage component in the task. They examined the relationship between three span tasks (verbal span, math span, and spatial span) and comprehension. The span tasks had both a processing and a storage component. The verbal and math span tasks correlated with verbal abilities, whereas the spatial span task did not. This finding was interpreted as evidence that there are at least two major systems--one for representing and processing verbal and symbolic information and one for representing and processing spatial information. However, to show that the crucial variable in individual differences in working memory is processing efficiency (and not storage), Daneman and Tardiff added three storage-free span tasks in which only processing was tested. They also found a correlation between these tasks and comprehension, which led them to reinforce the claim that it is individual differences in processing efficiency that explain differences in language abilities.

Nevertheless, Turner and Engle (1989) disputed the view that working memory capacity was functional and offered an alternative hypothesis for the relationship between working memory capacity and reading comprehension. For them, good readers have a larger working memory capacity in general, independent of the task to which working memory is being applied. To test this hypothesis, Turner and Engle devised a working memory span task, which they termed operation-word span task. In this task, the processing component was a math operation such as "(4/2)-1=1?" (Engle, Cantor, & Carullo, 1992, p. 975) and the storage component was a word (e.g., "snow"), which was written just beside the operation. They measured participants' working memory capacity by presenting them with increasingly longer sets of operation-word strings, asking them to verify whether the result of the operation was correct, and then asking them to read the word following the operation. Participants' memory span was the number of words recalled.

Turner and Engle (1989) found that the number of words recalled correlated significantly with measures of reading comprehension. That is, the processing component of the memory span task does not need to be reading-related, as Daneman and Carpenter (1980, 1983) had argued, to correlate with language comprehension. In addition, Turner and Engle showed that the operation-word span task and the reading span task explained similar amounts of variance in reading comprehension. They interpreted these findings as evidence for the general capacity hypothesis, which sees working memory capacity as a single, unitary resource independent of the nature of the task. According to the general capacity hypothesis, limitations in working memory capacity reflect a common factor and are domain free (Engle & Oransky, 1999; Engle, Kane, & Tuholski, 1999).

Engle and his group have been pursuing the question of whether working memory capacity is task-specific or domain free over the past decade and have consistently replicated the finding that the operation-word span test is a predictor of reading comprehension (e.g., Cantor & Engle, 1993; Engle, Cantor, & Carullo, 1992; Conway & Engle, 1996; Engle & Conway, 1998; Engle, Nations, & Cantor, 1990; La Pointe & Engle, 1990), thus accumulating evidence for the general capacity view of working memory. Likewise, the reading span test, which reflects the task-specific view, has been validated through numerous studies replicating Daneman and Carpenter's (1980, 1983) findings (e.g., Daneman, 1991; Daneman & Tardiff, 1987; Masson & Miller, 1983; Miyake, Just, & Carpenter, 1994; Tomitch, 1995). The present research project hopes to contribute to this debate by examining whether there is a significant relationship between working memory capacity and L2 speech performance and whether this relationship is task-specific or domain-free. The next section reviews the literature on working memory and L2 skills.

2.5 Working memory and L2 skills

As mentioned earlier, research on individual differences in working memory has focused heavily on first language reading comprehension. Recent studies, however, have addressed the relationship between working memory capacity and L2 skills. These studies have focused on reading comprehension, syntactic acquisition and comprehension, and speech production.

Harrington (1991) investigated the extent to which vocabulary and grammatical knowledge affects the relationship between L2 working memory capacity and L2 reading measures. He obtained measures of vocabulary knowledge, grammatical knowledge, L2 reading comprehension, and working memory span (measured by means of reading span tests in the L2) from 55 Japanese learners of English as a foreign language. He found a significant correlation between working memory capacity and scores on L2 vocabulary, grammar, and L2 reading measures. However, and most importantly, the correlation between working memory capacity and L2 reading comprehension measures was maintained even after the contributions of vocabulary and grammatical knowledge were partialled out. Harrington interprets these results as evidence that the L2 reading span test assesses more than lexical and grammatical knowledge, thus being an important index of reading comprehension in the L2. In addition, the results also replicate the findings of studies in L1 reading comprehension showing that the reading span test taps processes other than vocabulary knowledge (e.g. Dixon, LeFevre, & Twilley, 1988; Engle, Nations, & Cantor, 1990).

Harrington and Sawyer (1992) investigated the relationship between working memory capacity and L2 reading comprehension. They applied working memory span tests in both L1 and L2 to 32 advanced Japanese students of English as a foreign language, and compared performance in the span tests to performance in the Grammar and Reading sections of the Test of English as a Foreign Language (TOEFL). L2 proficiency was controlled by means of participants' overall scores on the TOEFL, which had to be at least 500. The memory span tests consisted of a reading span test based on Daneman and Carpenter (1980, 1983), a digit span test and a word span test. To avoid floor effects, Harrington and Sawyer devised their reading span test with 42 sentences smaller and syntactically simpler than those of Daneman and Carpenter's original tests.

The results showed that the mean scores for the digit and word span tests in the participants' first language were significantly higher than in the L2, English. However, the researchers report that there was no significant difference between L1 and L2 reading span, between which a moderate correlation was found. No significant correlation was found between L2 digit or word span and L2 reading comprehension measures, as in the L1 research. The most important finding of the study was that there was a significant correlation between the L2 reading span--which reflects L2 working memory capacity--and the results on the TOEFL Grammar and Reading sections, thus replicating in the L2 the findings of L1 reading comprehension studies.

Berquist (1998) investigated the relationship between L1 working memory, L2 working memory, and L2 proficiency. Participants were 60 French learners of English as a foreign language. Berquist administered two memory tests: a word span test and a reading span test, each one in French and in English. L2 proficiency was measured by means of the listening and reading sections of the Test of English for International Communication (TOEIC). To control for processing of L2 sentences in the reading span test, Berquist introduced a cloze in which participants had to fill in missing words of randomly chosen sentences from the test. Results showed that memory span in the participants' L1, French, was greater than in the L2, for both the simple word span and the reading span tests, a finding similar to that obtained by Harrington and Sawyer (1992). Berquist found significant correlations between L1 and L2 word spans, L1 and L2 reading spans, and between all memory spans and L2 proficiency, as measured by the TOEIC, with the best correlations being those between L2 memory spans and L2 proficiency. Interestingly, he reports that L2 word span correlated more strongly with TOEIC than did the L2 reading span. However, the most important finding of the study was that the highest correlation obtained was that between the L2 cloze test, a submeasure of the L2 reading span test, and L2 proficiency. Berquist interprets these results as evidence that L2 working memory is a good predictor of L2 proficiency and is a function of processing efficiency in the L2, rather than a fixed or reduced capacity. In addition, in his view the results also show that L2 working memory does not seem to be superior to L2 short-term memory in predicting L2 level, contrary to the view in L1 studies that working memory, and not short-term memory, predicts performance in complex cognitive tasks.

Miyake and Friedman (1998) investigated the relationship between working memory capacity and the analysis and comprehension of complex sentence structures in L2. Participants were 59 Japanese advanced learners of English as a foreign language. Miyake and Friedman applied a listening span test in the participants' L1 and L2, an L2 syntactic comprehension task, and an L2 agent identification task aimed at determining to what extent participants' linguistic cue preferences (e.g. word order, animacy, agreement, or case marking) for understanding syntax were similar to those of native speakers of English, who rely heavily on word order.

Using a sophisticated statistical technique to go beyond correlational and regression techniques--path analysis--Miyake and Friedman were able to show that L1 working memory capacity determines L2 working memory, L2 working memory contributes to syntactic comprehension, and L2 working memory determines linguistic cue preferences. In the researchers' view, these results indicate that, at least for advanced learners, L1 and L2 working memory share the same pool of resources and that those with a larger working memory span are more efficient at comprehending language. Specifically, those with a larger working memory span are more efficient at selecting linguistic cues that aid in the comprehension of L2 complex sentences.

Still in the area of L2 skills, Mota (1995, published as Fortkamp, 1999), based on Daneman (1991), examined whether working memory capacity correlated with L2 speech rate and articulation. Adapting Daneman's methodology, she administered a set of seven experiments to 16 advanced speakers of English as a foreign language. Working memory was assessed by means of the speaking span test (Daneman & Green, 1986; Daneman, 1991) and the reading span test (Daneman & Carpenter, 1980 and 1983; Harrington & Sawyer, 1992) in the participants' L1, Portuguese, and in the L2, English. Speech production was assessed by means of a picture description task, an oral reading task, and the Oral Slip Task (Motley & Baars, 1976). The variables of speech production used were speech rate--total number of words per minute--oral reading rate, and number of articulation errors.

Results showed that L1 and L2 working memory correlate significantly when measured by the reading span tests--thus confirming Harrington & Sawyer (1992) and Berquist (1998), but not when measured by the speaking span tests. In addition, L2 working memory for speaking does not correlate with L2 working memory for reading. L2 working memory capacity, as measured by the speaking span test, correlates significantly with speech rate and articulation errors, two direct measures of speech production. That is, those with a larger working memory capacity spoke faster and were less prone to articulation errors in the L2. However, working memory capacity, as measured by the L2 speaking span test, did not correlate with oral reading rate, an indirect measure of speech production. Working memory capacity, as measured by the reading span tests in English and in Portuguese, correlated significantly with oral reading rate. These results were interpreted as evidence that working memory capacity is specific to the task to which it is applied.

As already pointed out, research on the human cognitive system in charge of the temporary storage and processing of information--working memory--has consistently shown that the system is involved in fundamental ways in the performance of everyday cognitively complex tasks, such as understanding language, producing language, solving problems, and learning complex tasks. Working memory has been conceptualized in many different ways, but all lines of research agree on the dynamic nature of the system as well as on its relevance to the understanding of human performance. Whereas most research on working memory capacity has been developed in the area of language processing, comprehension is the aspect of processing that has achieved most attention. Although there has been some initiative towards investigating the role of working memory capacity in language production, much more systematic research is needed so that we can understand better how this central cognitive system relates to speech production. The next Chapter reviews the literature on L2 speech production found relevant to the present study.

CHAPTER 3

Review of the Literature on L2 Speech Production

In a recent paper, Levelt (1995) remarks that despite playing a major role in the survival and development of human society, speaking has been a neglected subject of research in psychology. Although production is considered one of the three core topics in the psychology of language (Bock, 1996; Dell, 1996), it was not until the early 1960s that the systematic study of the skill of speaking was undertaken and given a psychological perspective (Levelt, 1995). Nevertheless, the general emphasis psychology researchers have given to language comprehension and the lack of studies on language production have both had their effects on L2 acquisition and use research.

Studies on L2 speech production are fairly limited in number and the field still lacks consensus on the constructs underlying L2 speech performance as well as on the most effective ways to approach oral performance, either from a theoretical or from a pedagogical perspective. Most studies of L2 speech production are scattered around several areas, including psycholinguistics, sociolinguistics, pragmatics, instruction/pedagogy, or any combination of these. Using a variety of methods, these studies have been principally concerned with the concept of fluency in L2 speech production from different perspectives (e.g. Butler-Wall, 1986; Ejzenberg, 1992; Riggenbach, 1989, 1990; Freed, 1995, among others). A few studies have focused on other aspects of L2 speech production, besides fluency (e.g., Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999). This chapter presents an overview of research on L2 speech production that was judged relevant for the present study. It is divided into three main sections. In the first section, two influential monolingual models of speech production are reviewed--Dell (1986) and Levelt (1989). The reason these two models were chosen to be reviewed is that they are the basis of bilingual models of speech production, specifically of those proposed by De Bot (1992), Green (1986), and Poulisse & Bongaerts (1994), which are reviewed in section two. Finally, section three reviews empirical studies on L2 speech production.

3.1 Speech production in L1

The first speech production model to be reviewed here is the one proposed by Dell (1986). Dell's model is proposed within the activation theoretical paradigm in cognitive psychology and is based on assumptions from linguistic theory--dealing with units and structures of language--and the mechanism of spreading activation. Because of the relevance of the mechanism of spreading activation, Dell's model can be classified as a spreading-activation theory of speech production.

In Dell's view, the idea that behavior is ordered can be equally applied to speech production. Thus, he suggests that speaking requires the construction of internal representations of planned behavior prior to overt behavior. In the case of speech production, representations and behavior entail both linguistic competence and cognitive processing.

Among the linguistic assumptions underlying the theory, one is of central importance, viz. that a speaker's knowledge of his/her language is separated into at least four levels: the semantic, syntactic, morphological, and phonological. Of these, Dell's theory deals with the three latter levels. Each of these levels--syntactic, morphological, and phonological--is associated with a set of generative rules that establish the possible combinations of units at that level in the language. In other words, the set of generative rules determines whether a given combination of units is a well-formed sequence in the language. Units can be syntactic (e.g., noun and verbs), morphological (stem, prefix, suffix) and phonological (vowel, consonants, initial stops, etc.). The set of rules determining the combinatorial possibilities of units at a given level interacts with the lexicon.

In Dells' theory, the lexicon is a network of stored knowledge that contains nodes for the linguistic units. Thus, there are nodes for concepts, words, morphemes, phonemes, and phonemic features. These nodes are connected in a somewhat hierarchical way: conceptual nodes connect to words, words to morphemes, morphemes to phonemes, and these to phonemic features. In order to describe the interaction between the set of generative rules at a given level and the lexicon, Dell proposes that the rules generate frames with categorized slots.

Thus, for instance, in the sentence "This cow eats grass" (1986, p. 286), the syntactic rules would produce a frame with slots for Determiner, Noun, present tense Verb, and Noun. The morphological rules, at the morphological level, would generate a frame with slots for word stems and affixes. The phonological rules would, according to the theory, produce a frame with slots for consonants and vowels, for example. Thus, the linguistic aspect of the model proposes that (a) sets of rules produce frames with slots, (b) the slots are filled in with units, (c) units are stored in the lexicon, represented as a network with connections between nodes. In addition, linguistic units are labeled as to the category they belong to. Dell terms the process of specifying units with category information "insertion rules" (1986, pp. 286-287). The insertion rules make it possible for the appropriate units to be inserted into the slots.

The most important cognitive processing assumption of Dell's model is that each level of encoding--syntactic, morphological, and phonological--requires the construction of a mental representation of the utterance. Thus, according to Dell, a planned utterance will have various simultaneous representations--one at each level. The syntactic, morphological, and phonological representations consist of "ordered sets of various words, morphemes, and phonemes" (1986, p. 287). Dell suggests these representations could be thought of as a set of order tags which, attached to nodes in the lexicon, determine the contents of the representation as well as their order. The construction of representations takes place in a hierarchical fashion. The tagged nodes of a higher representation--say, syntactic--activate the nodes of the next lower representation--say, morphological. The mechanism that allows information from one level to trigger the tagged nodes of the next level is the retrieval mechanism of spreading activation.

Spreading activation takes place in two ways at the same time: from the higher level to the lower level and back. In addition, various nodes get activated at the same time to fulfill one slot in the frame built at a particular level. Level of activation will determine which node is selected for the slot, the most activated node being the one selected. After being selected, its level of activation drops towards zero. The set of nodes selected to fill in the slots at a given level constitutes the representation for that level.

The evidence for the two principal mechanisms of Dell's model, viz. the frame-and-slot mechanism and the spreading activation mechanism, comes mainly from speech errors. The model is powerful in that it accounts for a great deal of current speech error data (Dell, 1986; Poulisse, 1999), in both explaining and predicting the several types of slips of the tongue. However, the model is based on a number of assumptions that still need to be empirically studied.

A much more comprehensive and ambitious model of monolingual speech production is the one advanced by Levelt (1989), who labels it "a blueprint for the speaker" (1989, p. 8). Arguing that "the speaker is a highly complex information processor who can, in some still rather mysterious way, transform intentions, thoughts, and feelings into fluently articulated speech (1989. p. 1), Levelt proposes that both conceptual and linguistic processes take place in the construction of a message.

According to Levelt (1989), in order to speak, the individual has to go through a number of steps. First, s/he has to apply conceptual processes such as making an initial choice of purpose or conceiving an intention to speak. Conceptual processes, Levelt argues, depend on the speaker's state of motivation, the amount of knowledge shared with the interlocutor, and the discourse being created by the speaker and other participants in a given context. In addition to the conceptual processes, there are specifically linguistic processes the speaker needs to carry out to convey a message. After having in mind the concepts to be expressed, the speaker has to access words that map onto these concepts, retrieve syntactic forms that correspond to these words, and build a grammatical structure. The grammatical structure has to be given a phonetic plan that will, in turn, activate the processes of the articulatory system and finally be executed in terms of words, phrases, and sentences.

The processes involved in the conception of the message and in its development into a phonetic plan are assumed to constitute a general planning stage. The delivery of the message in actual sounds, as overt speech, is the general execution stage. Planning and execution processes do not take place in a linear way. Levelt (1989) argues that most sub-processes of planning and execution occur in parallel and execution may begin at any moment of the planning stage In other words, when speaking, the individual starts executing one piece of the message while still planning the following piece, carrying out planning and execution concurrently.

Levelt (1989) distinguishes four major components in his model of speech production: a conceptualizer, a formulator, an articulator, and a speech-comprehension system. Each of these components is autonomous and is responsible for specific processes. The notion of autonomy is fundamental to the functioning of the model. According to Levelt, each component operates independently and does not interact with the other components. In other words, there is no information sharing between the components and each works as if 'unaware' of what is taking place in the other components at a given moment during speech production. In order to produce fluent L1 speech, the components work in parallel. This is possible because each component contains a number of procedures, which are part of the speaker's procedural knowledge. These procedures act on the speaker's declarative knowledge, which is stored in his/her long-term memory as factual knowledge.

It is in the conceptualizer that message generation takes place. Generating a message involves selecting and ordering relevant information in terms of concepts. Message generation takes place through macroplanning and microplanning. Macroplanning consists of planning the content of the message by the elaboration of communicative goals and the retrieval of the information necessary to realize these goals. Microplanning, on the other hand, consists of planning the form of the message and involves making decisions about the appropriate speech act--for instance, a question, an assertion, a suggestion--determining what is "given" and what is "new", and assigning topic and focus. Levelt assumes that it is not necessary to finish macroplanning for microplanning to begin. The output of microplanning, and thus of the conceptualizer, is the preverbal message, which is also the input for the next component, the formulator.

It is in the formulator that the preverbal message is developed into a phonetic plan through the selection of the appropriate lexical units and the application of grammatical and phonological rules. The translation of the preverbal message into a linguistic structure takes place in two stages: grammatical encoding and phonological encoding. For grammatical and phonological encoding to take place, the formulator first has to access the mental lexicon, where lexical units are stored in the form of declarative knowledge. By accessing the mental lexicon, the formulator makes it possible for the lexical units that match the concepts of the preverbal message to be selected.

The lexical units of the mental lexicon contain information about their meanings as well as their syntactic, morphological, and phonological specifications. According to Levelt (1989), lexical units consist of two parts: the lemma and the lexeme. The lemma is that part of the lexical unit that carries semantic and syntactic information. The lexeme is the part that carries morphological and phonological information. Grammatical encoding consists of accessing lemmas and of building syntactic constructions such as noun phrases or verb phrases. The building of syntactic constructions takes place by means of procedural knowledge. Thus, the concepts of the preverbal message trigger the appropriate lexical units in the formulator which, in turn, trigger the syntactic information about these units, which trigger the syntactic procedures necessary to build a surface structure of the message. The surface structure is the output of grammatical encoding and is necessary for phonological encoding to take place.

Phonological encoding consists of building a phonetic plan for each lemma and for the utterance. The construction of this plan is carried out through the information contained in the lexeme: the information about the word's morphology and phonology. The output of the formulator, the phonetic plan or internal speech, is the input for the articulator, the next processing component. In the articulator, internal speech is converted into actual overt speech. According to Levelt (1989), articulation requires the coordination of a set of muscles and is highly procedural. The product of the articulator is what Levelt calls overt speech.

Finally, Levelt proposes that the speech-comprehension system is the component that allows the speaker to monitor both overt and internal speech. Monitoring may take place at all phases of the speech production process, and it allows the speaker to compare what was intended to what was linguistically planned.

As Poulisse (1999) points out, Dell's (1986) and Levelt's (1989) models of speech production have undergone several changes since they were first proposed. For instance, Dell, Juliano, & Govindjee (1993) suggest that the frame-and-slot mechanism proposed in Dell (1986) can be replaced by mechanisms of parallel distributed processing (PDP), which does not assume linguistic structures, frames and rules and claims, instead, that rule-like behavior results from the strengthening of the connections through practice. However, Dell et al. are careful to say that so far the possibility of replacing the frame-and-slot mechanism by PDP mechanisms applies only to a limited number of phonological phenomena.

As for the changes in Levelt's (1989) original model, these resulted from the construction of a computational model called WEAVER ++ (Levelt 1995; Levelt, Roelofs, Ardi, & Meyer, 1999) and several experimental studies investigating reaction times on lexical decision tasks and word-naming tasks. For instance, Levelt (1995) and Levelt et al. (1999) have adopted the network view of the mental lexicon as in Dell (1986), with spreading activation as the major mechanism of retrieval of information from the lexicon, and have thus made the process of grammatical encoding in the formulator more explicit. As already mentioned, Dell's (1986) and Levelt's (1989) original models are the basis for several bilingual models of speech production, three of which reviewed in the next section.

3.2 Bilingual models of speech production

Although researchers in the L2 area have been increasingly concerned about the psycholinguistic aspects of acquiring and using an L2 since the 1960s (Poulisse, 1997, 1999), it was not until very recently that explicit attempts were made at describing the processes involved in L2 speech production. Poulisse (1999) claims that one additional factor influencing L2 speech production model building is the need to account for how L2 speakers manage to mix and separate two (or more) languages, either willfully or accidentally. There are currently three main models aiming at accounting for L2 speech production: Green (1986), De Bot (1992), and Poulisse & Bongaerts (1994).

Green (1986) proposed a model of bilingual speech production for explaining how normal as well as brain-damaged patients perform in an L2. Based on evidence showing that bilingual brain-damaged patients may lose command of one language but not the other, Green claims that a speaker's L1 and L2 are organized in separate subsystems. In order to account for the fact that L2 speakers can choose in which language to speak, Green proposes the mechanism of activation. The language selected to be used will be the most activated. The nonselected language, however, can still be active, but its level of activation falls as a result of the degree of activation of the selected language, as well as of mechanisms of inhibition of the nonselected language.

Green proposes that the words of a particular language possess "tags" (1986, p. 216), which allow the speaker to identify the words and thus activate them. He also proposes a device, the specifier, to account for language switches and translation during L2 speech production. The specifier determines how the levels of activation must be controlled in order for the output to be in L1 and L2. In the case of code-switching, the specifier determines that the output is free to vary and the word used--in L1 or L2--will be the one that reaches the required threshold level of activation first. In the case of translation, the regulation of the system is more complex. More specifically, Green suggests that both languages will be active, but the output of one of them will have to be suppressed.

Green is particularly concerned with the resources consumed in controlling activation levels during speaking. He describes these resources as energy (1986, p. 211) and claims that, because they are limited, production may be impaired and errors will result if there is no time for the resources to get replenished (Poulisse, 1997) and inhibit activation of the nonselected language.

The advantage of Green's (1986) model is that it also postulates the mechanism of activation as part of L2 speech production processes. In addition, as Poulisse (1997, 1999) points out, he addresses the issue of limited resources as a possible cause of disruptions in L2 speech production. However, the model fails to give detailed accounts of message generation and syntactic, morphological, and phonological processes in L2 speech production.

A more comprehensive model of L2 speech production is that proposed by De Bot (1992), who based his proposal on Levelt (1989). Due to the enormous explanatory power of Levelt's model, De Bot made only those changes that were necessary to account for L2 speech behavior.

The first aspect De Bot (1992) deals with is related to the decision the speaker has to make as to which language to use. This takes place in the conceptualizer, since it is this component that deals with information about the participants and the situation in which speaking will take place. Deciding which language to use would take place during the macroplanning stage of the preverbal message. During microplanning, the speaker would thus use language-specific information to trigger the appropriate lexical units in the formulator.

With regard to the formulator, De Bot's (1992) proposal is that it is language-specific, thus suggesting that grammatical and phonological encoding of L1 and L2 takes place through different procedures. De Bot also suggests that L2 speakers produce two speech plans concurrently. Following Green (1986), he hypothesizes that L2 speakers produce speech for the selected language--the one being used to speak at a given moment--and for the active language, which is not being used but is regularly used. Simultaneous encoding of two languages allows code switching to happen.

In order to explain how the mental lexicon of an L2 speaker is organized, De Bot (1992) follows Paradis (1985, 1987). Based on neurolinguistic research, Paradis (e.g.

1985) claims that bilinguals have one conceptual store and two distinct semantic stores that are differentially connected to the conceptual store. This store, which corresponds to our experiential and conceptual information, contains mental representations of things, events, properties and qualities of objects, and our knowledge of the world. The lexical items representing these concepts are stored separately for each language. In his Subset Hypothesis (1987), Paradis proposes that the conceptual store is language-independent and that L1 and L2 lexical items are organized in different subsets. Having separate lexical stores implies applying different procedures to carry out grammatical and phonological encoding. De Bot suggests that, during L2 speech production, the L2 lexical subset gets more activated than the L1 subset.

In De Bot's (1992) proposal, the articulator is assumed to be one for both languages and to store a large set of sounds and pitch patterns, encompassing the two languages. By proposing one articulator where sounds and pitch patterns are stored together, De Bot is able to explain L1 phonological interference in the L2.

As Poulisse (1999, p. 61) suggests, De Bot's (1992) model of L2 speech production explains the following aspects of L2 oral performance: (a) lexicalization, (b) grammatical encoding, (c) code-switches, and (d) phonological interference. However, in Poulisse and Bongaerts' (1994) view, the model is particularly problematic in explaining how the speaker is able to carry out two speech plans concurrently. Thus, Poulisse and Bongaerts (1994) propose a model of L2 speech production, also based on Levelt (1989), that is aimed at overcoming this problem in Levelt's model.

Poulisse & Bongaerts (1994) follow De Bot (1992) in assuming that choice of the language to be used is made in the conceptualizer, during the construction of the preverbal message. However, Poulisse and Bongaerts propose that L1 and L2 lexical units are stored in a single network, and for this reason they need to contain information specifying to which language they belong. In Poulisse and Bongaerts' view, then, lexical units are tagged with a language label, an idea they adopted from Green (1986). In addition, they also propose that concepts have a distributed form. That is, concepts contain different meaning elements in the form of conceptual primitives (also features).

Thus, according to Poulisse & Bongaerts (1994), lexical selection occurs via the mechanism of spreading activation. During L2 speech production, L2 lexical units--lemmas and lexemes--receive more activation than those of the L1. This process gets started in the preverbal message, which contains the feature [+ L2]. For instance, the selection of the word "boy" will first require access to information, at the conceptual level, about the features [+ human, + male, - adult]. At this point, several lemmas can be activated to different extents, including L1 lemmas, but it is the degree of activation of the lemma "boy" that determines its selection. Selection of the lemma allows access to the lexeme, thus allowing grammatical encoding and subsequent phonological encoding. In the whole, grammatical and phonological encoding takes place the same way as proposed by De Bot (1992).

According to Poulisse (1997, 1999) the advantage of the Poulisse & Bongaerts (1994) model is that it does not postulate the elaboration of two concurrent speech plans, since the lexicon is accessed through activation of individual lexical items and not through the activation of entire subsets as proposed by De Bot (1992).

The models of monolingual speech production (Dell, 1986; Levelt, 1989) and bilingual speech production (Green, 1986; De Bot, 1992; Poulisse & Bongaerts, 1994) reviewed here are useful in that they make explicit the main processes involved in the production of speech. These models will be mentioned again in the discussion of the results obtained in the present study (Chapter 6). The next section reviews empirical studies on L2 speech performance.

3.3 Empirical studies on L2 speech production

As noted by Crookes (1989), a great deal of the work developed on L2 speech production has been in the psycholinguistic area and is put together in Dechert & Raupach (1980) and Dechert, Möhle, & Raupach (1984). Early studies of L2 speech production carried out by Dechert and his group at Kassel University in the early 1980s focused on the study of temporal variables and speech errors and adopted a research strategy that they termed "contrastive psycholinguistics" (Dechert, 1987; Dechert & Raupach, 1980). In this type of research, participants were asked to perform the same oral tasks both in their L1 and L2 and their speech samples were analyzed in terms of the similarities and differences in the speech production processes of the two languages.

The studies published by The Kassel Research Group utilized data from the Kassel Corpus, collected from 1979 to 1982, from L2 learners of English, French, or German telling stories or describing pictures both in their L1 and L2 (Dechert, Möhle, & Raupach, 1984). However, the collection of papers published in Dechert & Raupach (1980) and Dechert, Möhle, & Raupach (1984) do not all report empirical studies--some of them are theoretical discussions of specific aspects of L2 speech production. Thus, only three papers reporting results of empirical studies were found of interest for the present study, and these are reviewed below.

Möhle (1984) set out to examine differences and similarities in the foreign language planning processes of advanced learners of different foreign languages--French and German--who had different native languages--also French and German. The study was carried out with participants--3 French speakers studying German as an L2 and 6 German speakers studying French as their L2. L2 speech production was elicited by means of two tasks. The French participants were asked to describe a series of cartoons, first in their native language, then in their L2. They were also asked to answer two questions related to their experiences living in Germany on a study-abroad program. In this "free discourse task", as Möhle (1984, p. 29) calls it, the questions were first answered in their L1 and then in their L2. The procedures with the German students were slightly different. Of the 6 German students learning French, 3 were asked to describe the cartoon series and 3 were asked to answer the questions. Both tasks were performed first in the speakers' first language and then in their L2.

Möhle analyzed participants' speech samples in terms of (a) temporal variables, which included speech rate, articulation rate, length of speech units, number of pauses per 100 syllables, and length of pauses in seconds, (b) linguistic variables, which consisted of complexity and variation in syntactic structure, morpho-syntactic grammaticality and acceptability, lexical refinement, idiomatic acceptability, communicative effectiveness, and degree of L1 influence, and (c) the strategies participants used to solve L2 problems.

Möhle found that the cartoon description task yielded slower speech rates than the question-answer task for the two groups of learners, but that the learners of French as an L2 tended to have a faster speech rate than their counterparts in the same task. In terms of the linguistic variables investigated, Möhle reports that, although more disfluent, the learners of L2 German produced more complex language than the learners of L2 French. The lack of fluency demonstrated by the learners of German is probably related to the fact that German, compared to French, in Möhle's view, imposes heavier processing demands that interfere in the planning and execution processes of speech. Möhle also relates this finding to the type of L2 language education the participants received in their country, claiming that French speakers of L2 German tend to prioritize grammatical accuracy whereas German speakers of L2 French prioritized content. This, in her view, resulted in "an acoustically smoother flow of speech" by the learners of L2 French (Möhle, 1984, p. 37). Finally, Möhle found no significant differences between the two groups in the strategies used to solve L2 speech production problems.

Lennon (1984) examined temporal variables and hesitation phenomena in the retelling of a story by 12 advanced learners of English as an L2. Lennon compared the argumentative and temporal structures of the learners' speech samples to a recording of the story made by a native speaker. In terms of temporal variables and hesitation phenomena, Lennon focused on the time taken to retell the story, speech rate, speech-pause time ratio, distribution of pauses, use of final intonation falls, repetitions, and self-corrections. Lennon's hypothesis was that there should be harmony between the syntactic, narrative, and temporal structures and that this harmony should be close to that produced by the native speaker.

Lennon reports that participants tended to keep very close to the narrative structure found in the native speaker's recording. In addition, participants tended to perform the task in about the same time as the native speaker, but they presented a slower speech rate, higher pausing time, more repetitions, and more self-corrections than the native speaker. The greatest difference between participants' speech samples and the native speaker's was in the distribution of pauses. Lennon explains this variation in terms of the narrators' individual styles.

In Lennon's view, learners' speech samples sound less fluent than the native speaker's because they tend to segment speech in shorter units. He suggests that, whereas the native speaker paused after final intonation falls, which mark the end of a segment, and used this pausing time to plan the next segment, nonnative speakers probably found the pausing time after intonation falls insufficient for planning, thus finding extra time to plan by building shorter units of speech.

Raupach (1984) investigated the role of formulae--prefabricated routines—in the speech production of 2 German learners of French as an L2. His hypothesis was that formulae work as planning units in L2 speech production. By investigating changes in temporal variables, such as silent pauses, and in hesitation phenomena, such as repeats, self-corrections, false starts, and drawls (lengthening of syllables) in the participants' speech sample after one semester on a study-abroad program, Raupach found that the main changes in the participants' speech production were in the distribution of hesitation phenomena and not in the type of prefabricated routines used. For instance, before their semester abroad, participants tended to segment prefabricated routines into smaller units and to present a greater number of silent pauses. In Raupach's view, speech planning processes would take place during these unfilled pauses. After a semester abroad, participants tended to produce longer prefabricated routines, unimpeded by silent pauses, and to use drawls followed by shorter silent pauses as their strategy to carry out their planning processes.

Dechert (1983) examined temporal variables and hesitation phenomena as indicators of speech planning processes in the narrative of an advanced L2 speaker of English. Dechert noticed that the speech produced could be classified into two types--one marked by hesitations, repetitions, corrections, false starts, and silent pauses, and the other by formulaic, ready-made expressions of varying length and syntactic complexity, which Dechert termed "islands of reliability" (p.183). In Dechert's view, the production of fluent L2 speech is largely based on these islands, which function as anchoring points for planning and execution processes, and thus, as time-buying devices.

In their various publications, the researchers of the Kassel Group dealt with L2 speech production by comparing native to nonnative oral production in description and narrative tasks and by comparing nonnative oral production of the same speakers before and after their experience living in the country where the L2 was spoken as a native language. Their general assessment of L2 speech production was quantitative, through the measurement of temporal variables--speech rate, articulation rate, silent pauses, and mean length of run--and hesitation phenomena--filled pauses, repetitions, self-corrections, and drawls (Wiese, 1984). These researchers' general assumption was that temporal variables and hesitation phenomena indicate planning and execution activities, being thus important in the investigation of L2 speech production. In evaluating L2 speech production, the Kassel Group usually referred to the notion of 'fluency', this being used in a broad sense to express the degree to which learners' speech samples approached native oral production. In spite of the lack of additional studies on L2 oral production and the lack of a powerful model of speech production such as Levelt's

(1989) on which they could base their findings, the Kassel researchers helped establish a tradition in the examination of L2 speech production.

Recently, Towell, Hawkins, & Bazergui (1996) investigated the development of fluency in the L2 speech production of 12 advanced learners of French. Towell et al. asked participants to retell the story of a film both in their first language English, and in the L2. Data were collected before participants went on a study program in France and, again, a year later, when they returned. Towell et al. assessed L2 speech production by means of four temporal variables--speech rate, phonation time ratio, articulation rate, and mean length of runs. Their hypothesis was that increase in oral fluency would be mainly due to the proceduralization of knowledge of syntax and lexical phrases in the L2. This increase in fluency would be reflected in the learners' use of temporal variables.

Their analysis consisted of comparing each learner's speech sample before and after residence in France, in addition to comparing each learner's performance in L1 and L2. Towell et al. found that there was an increase in speech rate, articulation rate, and mean length of run from the first to the second recording. In addition, participants' phonation time ratio, "the percentage of time spent speaking as a percentage proportion of time taken to produce the speech sample" (Towell et al., 1996, p. 91), did not fall and their average length of pause did not increase. They also found, however, that the increase in speech rate, articulation rate, and mean length of run did not bring their speech to levels comparable to those of their speech in their native language. Towell et al. interpret these results as evidence that after a period of residence abroad, increase in oral fluency is mainly due to changes in the formulator, the component in Levelt's (1989) model of speech production responsible for the processing of lexicogrammatical information, rather than to changes in the way participants conceptualized the message in the L2 or the way they articulated speech in the L2. In the qualitative analysis of the speech samples of 2 participants, Towell et al. were able to show that there was an increase in the length and complexity of linguistic units between pauses, which shows that access to lexicogrammatical knowledge was faster and probably proceduralized.

It is important to note that Towell et al. were primarily concerned with fluency in L2 speech production. Although not offering a working definition of the term, they seemed to interpret fluency as a primarily temporal phenomenon. Taking the same view of fluency, Temple (1997) examined a number of temporal variables in the speech of learners of French and compared these to the recordings of twenty native speakers.

Temple analyzed speech rate, pause time ratio, mean length of run, hesitation rate, and repair rate. Eleven learners of French as an L2 at the intermediate and advanced levels were asked to perform an interview task at the beginning and end of a semester. Comparing the two sets of speech samples, Temple reports that there were no statistically significant gains in oral fluency during the semester. Temple also compared learners' speech samples to the recordings of native speakers performing an interview task, taken from a commercial cassette. The comparison between nonnative and native speakers' speech samples revealed statistically significant differences in speech rate, hesitation rate, and repair rate. Temple interprets these results as evidence that while native speakers' speech production processes occur, to a large extent, in parallel, nonnative speakers' processes take place serially, since a great deal of their knowledge of the L2 is declarative and stored as explicit memory. Following Towell et. al (1996), Temple relates her findings to Levelt's (1989) model of speech production and suggests that nonnative speakers' processing difficulties take place mainly in the formulator, where grammatical and lexical information is dealt with. In her view, lack of fluency in L2 speech production is due to insufficient knowledge of the L2 or, alternatively, to the use of explicit knowledge through controlled processes, which can only take place in serial fashion, thus overloading working memory.

The studies reviewed so far have examined the 'fluency' dimension of L2 speech production through the use of temporal variables and hesitation phenomena, from a psycholinguistic perspective. Several other studies have also used temporal variables and hesitation phenomena to study L2 speech production, but aiming at specifying the notion of fluency and describing the profile of fluent and nonfluent L2 speakers.

For instance, Lennon (1990) attempted to quantify the components of fluency. His method consisted of analyzing speech samples of a group of four adult German learners of English as a foreign language recorded before and after a period of residence abroad. Lennon devised a wide range of measures of fluency, which encompassed temporal variables and what he calls 'dysfluency markers'. He assessed 8 temporal variables (unpruned and pruned words per minute, total unfilled and filled pause time, mean length of runs, T-units followed by pauses, total and mean pause time at T-unit boundaries)[1], and 4 dysfluency markers (repetitions, self-corrections, filled pauses, percentage of repeated and self-corrected words). Participants' speech samples were also submitted to the evaluation of native-speaker teachers of the L2, who were asked to judge learners' fluency on a subjective and holistic basis.

[1] According to Lennon (1990, p. 406), the category "unpruned words per minute" includes all words produced in the speech sample, whereas " pruned words per minute" consist of all words except self-corrected words, repeated words (except when for rhetorical reasons), words related to comments on the task, and words addressed to the experimenter.

Comparing each learner's first and second narratives, Lennon found that the improvements in their oral fluency were mainly due to an increase in speech rate and a decrease in the number of filled pauses. The comparison also showed that learners tended to have fewer repetitions and filled pauses per T-Unit in the second sample, but their number of self-corrections did not decrease. In addition, their unfilled pause time decreased and their mean length of runs increased. Lennon's general interpretation of findings is that there are at least two key components of the dimension 'fluency' in L2 speech production: a temporal component, revealed in terms of speech/pause relationships, and a dysfluency marker component, revealed in terms of repetitions, self-corrections, and filled pauses.

Furthermore, the three variables that were found to increase significantly were pruned words per minute, filled pauses per T-Unit, and T-Units followed by a pause, which, Lennon argues, should be considered as constituting the core of any set of measures of fluency.

Whereas Lennon attempted to describe a set of variables that function as indicators of the fluency dimension of L2 speech production and identified changes in learners' speech after exposure to the L2, Riggenbach (1989, 1990) analyzed the speech of 6 Chinese learners of English as a second language, three rated as very fluent, and three as very nonfluent. Her primary goal was to identify which features of the speech of highly fluent nonnative speakers differed from those of highly nonfluent speakers.

Following Hunt (1974) and Vorster (1980), Lennon defines a T-unit as "one main clause and all its subordinate clauses and nonclausal units" (1990, p. 406).

Riggenbach asked participants to record a dialogue with a native speaker on any topic. Participants' speech samples were first rated for fluency by native speaker teachers of English as an L2. Riggenbach selected 5-minute excerpts of each participant's dialogue and analyzed them in what she terms fluency-related features--hesitation phenomena such as filled and unfilled pauses, repair phenomena, such as repetitions and restarts, and rate and amount of speech. Riggenbach's analyses of participants' speech samples also included an analysis of interactive phenomena--such as backchannels, echo, questions, and laughter particles--and interactive features such as latches, overlaps, gaps, and collaborative completion.

Riggenbach reports that the quantitative analysis showed few statistically significant differences in features between fluent and nonfluent subjects. However, she was able to verify that fluent and nonfluent nonnative speakers differed in terms of speech rate and number of filled pauses, which supports Lennon's (1990) findings. Riggenbach also reports that type of hesitation phenomena, in addition to the speech rate and repair phenomena, analyzed quantitatively, were the most salient features distinguishing highly fluent from highly nonfluent nonnative speakers. The qualitative analysis showed that the fluent L2 speakers tended to produce lexical filled pauses (e.g. "y'know"), whereas the nonfluent speakers tended to produce nonlexical filled pauses ('uhm') and silent pauses. Also, fluent L2 speakers tended to produce pauses that were considered nativelike because they occurred at junctures and in isolation, in contrast to the nonfluent participants, whose pauses occurred at places other than junctures and in proximity with other disfluencies, which contributed to the effect of choppiness of speech. As for the interactive phenomena and interactive features, Riggenbach suggests that fluent L2 speakers show the ability to initiate topic changes, to carry on the conversation through demonstrations of comprehension, to show anticipation of ends-of-turn, and to balance their amount of speech in relation to the interlocutor's.

Riggenbach was one of the first studies to use conversational data and to include interactive features in the examination of L2 speech production, more specifically, fluency. Her study is basically descriptive in that it attempts to identify components of fluency, like Lennon's (1990) reviewed above. Two other studies have used conversational data to examine L2 speech production.

Olynak, d'Anglejan, and Sankoff (1990) investigated several hesitation phenomena in L1 and L2 speech production. Participants were 10 French-speaking learners of English judged by the researchers to be high-fluency and low-fluency speakers. Participants were required to tape record a conversation in three situations: English planned, French unplanned, and English unplanned. In the first situation, participants were required to teach a skill or present new information to a peer in the L2, English. In the second and third situations, they were interviewed by a native speaker in their first and second languages.

Olynak et al. assessed first and second language oral fluency in the recorded conversations by means of hesitation phenomena, which they termed speech markers: cut-offs (a glottal or other stop), repeats, repairs, 'uhs' (filled nonlexical pauses), and transitions (turn changes). Their hypothesis was that frequency of occurrence of these speech markers would be an important component of oral fluency. Quantitative analyses of the speech samples--which lasted from 25 seconds to 2 minutes--revealed that, contrary to what was expected, the high-fluency speakers used more speech markers than the low-fluency speakers in general, but this difference was not statistically significant. The markers produced by the high-fluency speakers tended to occur at the end of a speech unit, a transitionally relevant place, which led the researchers to conclude that it is not the frequency of speech markers that is important, but their function in speech production.

Olynak et al. also found that the situation, or context, in which L2 speech was produced affected the production of speech markers by both high- and low-fluency speakers. The researchers report that the lowest number of speech markers occurred in the French unplanned situation, whereas the highest number occurred in the English unplanned situation, with English planned lying in between. Finally, comparing participants' L1 and L2 speech samples, Olynak et al. found that the types of speech markers used were the same for each individual in his/her L1 and L2, which shows, in the researchers' interpretation, that L2 speakers transfer their pattern of hesitation phenomena from their L1 to their L2.

Ejzenberg (1992, 1995) also focused on the fluency dimension of L2 speech production and investigated the effects of the nature of the task--dialogues versus monologues--on oral fluency in the L2. Ejzenberg assessed fluency through subjective ratings, as well as four quantitative variables which she calls fluency markers: amount of speech, speech rate, talk unit length, and fluent unit length. In addition, she verified whether there were quantitative and qualitative differences in the speech produced by high-fluency and low-fluency L2 speakers relative to the task being performed. Participants were 50 Brazilian learners of English as a foreign language at the intermediate and advanced levels.

By manipulating the structure of the tasks--two monologues versus two dialogues--and the amount of external support to the tasks (i.e., external information), Ejzenberg was able to show that interactivity and external support are important variables affecting nonnative speakers' display of fluency (1995, p. 17). With regard to interactivity, participants were more fluent in dialogues than in monologues. With regard to amount of external support, participants' fluency decreased when they had to perform tasks without external support. Their fluency increased in tasks with a high amount of external support and interactivity, since in one of the dialogue tasks in which they were required to deliver a presentation following a set of rules, their fluency increased.

Ejzenberg reports that the four fluency markers discriminate well between high-fluency and low-fluency nonnative speakers, with rate of speech as the most salient indicator in the four tasks. In addition, high-fluency speakers, as subjectively judged by four raters and as shown by the fluency markers, also made use of strategies that helped them maintain "an air of fluency" (Ejzenberg, 1995, p. 38). For instance, Ejzenberg reports that high-fluency speakers optimized their use of time by trying to say more, competing for the floor with the interlocutor, in the case of the dialogue tasks, and showing willingness and eagerness to accomplish the tasks. In contrast, low-fluency speakers were reluctant to speak, acknowledged difficulty to accomplish the tasks, and showed concern with making themselves understood, which affected their fluency ratings and measures.

Both Ejzenberg (1992, 1995) and Olynak et al. (1990) made use of a number of different variables and tasks to address the notion of fluency in L2 speech production elicited by means of conversational data. Whereas Dechert (1983), Lennon (1984), Möhle

(1984), Towell et al. (1997), and Raupach (1984) were interested in L2 speech production from a psycholinguistic perspective and claimed that temporal variables and hesitation phenomena reflected mental processes during speech production, Ejzenberg, (1992, 1995) and Olynak et al. (1990), examined fluency from a sociolinguistic perspective and introduced the idea that oral fluency in nonnative speech is context-dependent and varies across tasks.

Freed (1995) has recently addressed the fluency dimension of L2 speech production. More specifically, she investigated whether global perceptions of fluency by native speaker judges would distinguish between two groups of L2 learners--one with experience in studying for one semester in the country of the target language, France--and the other with formal classroom instruction only. Freed also attempted to describe features of fluency that would distinguish the two groups. The speech samples of 30 participants were first subjectively analyzed by a group of 6 native speaker judges on a 7-point scale.

Speech was elicited by means of the Oral Proficiency Interview (OPI), a widely used speaking test, in the beginning and end of the semester. Linguistic analyses of a subset of the speech samples were also performed in order to identify attributes of fluency that would help determine those learners who had been abroad from those who had not. For this linguistic analysis, Freed chose mainly temporal variables and what she calls "dysfluency markers" (1995, p. 129), including amount of speech (total number of non-repeated words), rate of speech (number of non-repeated words per minute), total number of filled and unfilled pauses, mean length of runs and several variables of repair phenomena (repetitions, reformulation, false starts, grammatical repairs, and clusters of dysfluencies).

Freed's (1995) results showed that the subjective ratings of fluency by the native speakers revealed a small difference in the perceived global fluency between the two groups, with slightly higher ratings, unexpectedly, for the less advanced learners. Freed notes that the judges were influenced by factors other than the temporal variables she assessed, including notions such as accuracy of grammar, richness of vocabulary, accent, rhythm, tone of voice, and confidence in speech. She concludes, however, that living abroad did have an effect on the learners' speech performance, as shown by the linguistic analyses: subjects who had lived abroad tended to speak more and faster, with fewer silent and nonlexical filled pauses, longer speech runs, and a greater number of reformulations and false starts.

There have been only a few attempts, to the best of my knowledge, to address instructional issues related to L2 speech production development. Two studies were located in the literature on L2 speech production--one theoretical (Gatbonton & Segalowitz, 1988) and one empirical (Arevart & Nation, 1991).

Assuming that formulaic speech is an important characteristic of fluent speech, Gatbonton & Segalowitz (1988) proposed a theory of 'creative automatization' to be applied in L2 classroom settings. In this theory, fluency in speech is related to two broad aspects of language. One aspect reflects speakers' general communicative ability to convey their message, which involves knowing what to say, when, where, and how, and attending to complex sociolinguistic and intercultural factors. The other aspect reflects the ease with which the mental processes involved in the actual production of utterances take place, thus emphasizing the cognitive and linguistic aspect of speaking.

According to Gatbonton & Segalowitz, in the L2 it might be the lack of linguistic knowledge--that is, lexical, syntactic, and phonological knowledge--that causes nonfluent speech. The search for this kind of information, in these researchers' view, demands a great deal of the speaker's attentional resources, resulting in discontinuous speech, full of hesitations. They suggest, then, that one way of developing fluency in L2 learners' speech is to provide them with formulaic speech, "speech forms produced as unanalyzed wholes, prepatterned expressions, or routinized utterances" (Gatbonton & Segalowitz, 1988, p. 473), within the communicative framework.

The creative automatization theory proposes that the automatization of specific utterances or sentence frames reflecting various language functions (e.g. requesting, directing, describing past activities) can be developed in the L2 classroom by means of activities in which the learner is given extensive practice in using these items while conveying messages.

Arevart & Nation (1991) examined the effect of a pedagogical technique aimed at developing L2 learners' rapidity in speaking. The 4/3/2 technique consists of requiring learners to deliver a 4-minute talk on a familiar topic to a partner, then the same talk in 3 minutes to a different partner, and finally in 2 minutes to a new partner. The idea of the technique is to make learners deliver the same talk with less and less time available, each time to a new interlocutor. Arevart & Nation wanted to test the idea that repetition, a changing audience, and decreasing time would contribute to improvement in L2 oral fluency. Participants were 20 learners of English as a second language. Oral fluency was assessed by means of speech rate in words per minute and hesitation phenomena--filled nonlexical pauses, repairs, sentence incompletion, repetition, self-correction, and throat clearings and sighs.

Arevart & Nation report that there were significant improvements in speech rate and number of hesitations from the first to the third delivery (Gatbonton & Segalowitz, 1991, pp. 88-89). In their view, repetition allowed learners to focus on the message and free their attentional resources from having to cope with lexical and syntactical problems. The researchers suggest that oral fluency is a trainable skill and improvement achieved through training will carry over to other tasks sharing the same language resources.

The studies reviewed so far all have attempted to investigate L2 speech production by using primarily the term "fluency" in their assessment of L2 speech. In most of these studies, fluency is used as an umbrella term, in its broad sense, as Lennon (1990) puts it, to refer to the speakers' general L2 oral competence or proficiency. In other studies researchers, implicitly or more explicitly, relate the idea of fluency to speech rapidity, to the flow of speech without this being unimpeded by hesitations.

Several conclusions can be drawn from this review of empirical and theoretical studies. In terms of L2 speech production, these studies show that:

a) Nonnative speakers tend to transfer their individual pause profiles from their native language to their L2 (Chafe, 1980; Olynak et al., 1990).

b) Compared to L2 speech, nonnative speech shows a greater number of pauses, greater pause time, increased hesitation phenomena, and decreased speech rate (Deschamps, 1980; Raupach, 1980).

c) Nonnative speech production is sensitive to context (Ejzenberg, 1992, 1995; Olynak et al., 1990; Riggenbach, 1989, 1990) and task structure (Ejzenberg, 1992, 1995).

d) Individual differences in fluency play an important role: nonnative fluent speakers share a great number of fluency features, whereas nonfluent speakers will be nonfluent in idiosyncratic ways. That is, what characterizes one speaker as nonfluent may be present in the speech of other nonfluent speakers, but may not be as relevant (Ejzenberg, 1992, 1995; Freed, 1995; Olynak et al., 1990; Riggenbach, 1989, 1990).

e) Frequency of hesitation phenomena is related to the production of new utterances and to the level of cognitive difficulty of the task, whereas a higher speech rate is observed when the speaker is being repetitive (Goldman-Eisler, 1968).

In terms of methods and concepts, the review shows that:

(a)Fluency is an ill-defined concept in the L2 speech production literature, probably encompassing a multitude of interactive factors--linguistic, psycholinguistic, and sociolinguistic in nature (Freed, 1995).

b) L2 speech production has been elicited by means of monologues (descriptions and narratives) and dialogues (conversations and interviews).

c) L2 speech production has been measured by means of temporal variables and hesitation phenomena, both categories being composed of several sub-measures, many times interrelated and referred to by different labels (e.g. disfluency markers, speech markers, fluency markers, repair phenomena).

In spite of a certain lack of cross-referencing among researchers focusing on L2 speech production, the field has made important methodological progress with the publication of recent groundbreaking studies examining the effect of planning time on the production of L2 speech in the context of the task-based approach as put forward by Skehan (1996a, 1998). Although not directly related to the present research project, three of these studies will be reviewed here because they provided a sound methodological apparatus on which to base my choices as to how to best assess L2 speech performance. Before going on to review these studies, though, the background to Skehan's proposal for a task-based approach to L2 instruction will be presented below.

In a series of recent publications, Skehan (1992, 1996a, 1996b, 1998) has advanced a framework for the implementation of task-based L2 instruction based on current developments in cognitive psychology, more specifically, on information-processing theories. The task-based approach movement emerged in the mid-1980s as an alternative to the traditional approach to L2 instruction known as PPP--presentation, practice, and production (Skehan, 1996a). Criticism to the effectiveness of the PPP model was raised on the basis of research showing that levels of attainment in L2 learning, provided through the use of this model, were poor and that most learners did not reach satisfactory levels of proficiency (e.g. Carroll, 1975; Skehan, 1989; Stern, 1983). Further research (e.g. Brumfit & Johnson, 1979; Ellis, 1985) also challenged the view, inherent in the PPP model, that focus on specific language forms leads to learning and subsequent use.

Thus, convinced that instruction is a relevant component of L2 learning, but unsure of how best to deal with it in the classroom for communicative purposes, researchers started to examine classroom interaction and the ways it affected learning (e.g. Doughty & Pica, 1986; Duff, 1986; Gass & Varonis, 1985). The new research in classroom interaction coincided with L2 teachers' attempts to design and apply communicative activities in the classroom, focusing on the "meaning" aspect of these activities. Investigations into the validity of these activities from the classroom interaction perspective led researchers to propose a task-based approach to L2 instruction (e.g. Crookes & Gass, 1993a, 1993b; Long & Crookes, 1991; Nunan, 1989; Prabhu, 1987).

Skehan (1998, p. 95) defines a task, in the context of task-based instruction, as an activity in which (1) meaning is the primary focus, (2) there is some type of communication problem to solve, (3) there is a relationship to real-world activities, (4) completion is important, and (5) assessment of outcome is important. However, given its broad scope, its concern with focus on meaning to the detriment of form, and growing evidence from research on L2 acquisition that some focus on form is necessary to foster acquisitional processes, Skehan (1996a, 1996b, 1998) claims that the task-based approach should work along with some principles aimed at guiding the use of tasks, if these are to benefit L2 learning and use. Thus, he proposes that one way to implement the task-based approach is to set goals as to what is expected from performance in the L2. One general goal of L2 learning, in his view, is that of becoming more native-like in performance. Within this general goal, it is possible to organize learners' goals, from a pedagogic perspective, into three main categories of L2 performance: fluency, accuracy, and complexity (Skehan, 1996b).

This three-way distinction makes it possible to assess L2 performance within realistic parameters by allowing the combination of feasible pedagogical goals with a well-defined set of dimensions of L2 performance on which researchers can draw. Thus, by fluency, Skehan (1996b, 1998) means "the capacity to mobilize one's linguistic resources in the service of real-time communication, i.e. to produce (and comprehend) speech at relatively normal rates, approaching (but not necessarily identical to) one's own native-language speech rates" (1996b, p. 48). Skehan states that to examine fluency, according to this definition, would require determining features such as speech rate, pausing, reformulation, hesitation, and redundancy. Accuracy and complexity both concern form, with a difference in degree. Accuracy is related to "freedom from error, based on whatever language is used" (Foster & Skehan, 1996, p. 304), being thus also related to the norms of language and to learners' beliefs in these norms (Skehan, 1996b). Complexity concerns the organization and internal elaboration and structuring of what is said, reflected in the variety of syntactic patterning the learner uses (Foster & Skehan, 1996; Skehan 1996b).

One line of research within task-based approaches to L2 learning and use includes studies examining the effects of planning time on L2 speech production, three of which are of particular relevance to the present study because of the way researchers assessed L2 speech production. The major focus of the first study, Foster & Skehan (1996), was to examine the effects of different types of task design on the variables of fluency, complexity, and accuracy on L2 speech production under three conditions for each task--unplanned, planned without detail, detailed planning. Participants were 32 preintermediate-level learners of English as a foreign language, from a wide variety of L1 backgrounds. Fluency was operationalized as continued performance and as repair avoidance. The measures used to reflect this view were reformulations (phrases or clauses repeated with some modification), replacements, false starts, repetitions, hesitations (initial phoneme or syllables uttered more than once before complete pronunciation of the word), pauses (silences of 1 second or longer), and silence total (the sum of pauses). Complexity was measured by number of clauses and syntactic variety. Accuracy was assessed by means of number of error-free clauses.

Participants were divided into three groups and all groups were required to perform a personal information exchange task, a narration, and a decision-making task. Group one had no planning time. Group two was given planning time but no guidance as to how to use the time, and group three was given time as well as guidance as to how to use the available to time to plan what to say in each task. Results showed that there was a straightforward effect for planning on the fluency and complexity aspects of L2 speech performance on the three tasks. However, effects of planning on accuracy were only significant for group two--undetailed planners.

Foster and Skehan (1996) explain these results by claiming that attention is limited in capacity and learners have to choose what to prioritize when performing orally: allocating attentional resources to fluency and complexity penalized accuracy for group three (detailed planners), which also had to follow the suggestions given as to what to focus on when performing the tasks.

Similar results were obtained by Mehnert (1998), who investigated the effects of planning time on the speech performance of 31 intermediate learners of German as a foreign language. Participants were divided into four groups and all were required to perform two narrative tasks. Group one was given no planning time, whereas groups two, three, and four were given 1, 5, and 10 minutes planning time, respectively.

Following Skehan (1996b), Mehnert (1998) assessed L2 speech production in terms of fluency, accuracy, and complexity. Fluency was assessed by number of pauses, total pausing time, mean length of run, and speech rate. Accuracy was measured by the percentage of error-free clauses and number of errors per 100 words. Complexity was measured by number of dependent clauses per communication unit (a unit used in the analysis of spoken interaction that takes into account ellipsis as well as utterances such as *hello* and *till tomorrow*). In addition, she included a further dimension of L2 speech production, lexical density, measured in terms of weighted lexical density (see Chapter 4, section 4.2.4.4 for details).

Results showed that planning time had an effect on all measures of L2 speech production, with 10-minute planners producing more fluent, more accurate, more complex and denser speech than nonplanners. However, Mehnert (1998) points out that effects of planning time were not the same across the four dimensions investigated. For instance, fluency increased across time, but the degree of improvement diminished with time. Accuracy improved only in the 1-minute condition. Complexity increased only in the 10-minute condition. Density seems to have been the only measure to improve regularly across the four conditions. Mehnert explains these results by claiming that there is a competition for attentional resources between the goals of fluency, accuracy and complexity, with gains in the three measures not being achieved simultaneously.

The third study investigating L2 speech performance using the framework proposed by Skehan (1996a, 1996b, and 1998) was conducted by Ortega (1999) with 32 dyads of advanced learners of Spanish as a foreign language. Each member of the pairs was assigned the role of speaker or listener.

Also focusing on the effects of planning time on L2 speech performance, Ortega asked participants to perform a story-retelling task under two conditions: planned (with

10 minutes for planning) and unplanned. L2 speech performance was assessed in terms of fluency, accuracy, lexical range, and syntactic complexity. Fluency was measured by pruned speech rate. Accuracy was measured by analysis of native-like use of noun-modifier agreement and use of the Spanish article system. Lexical range was measured by dividing the number of different words by the total number of words produced. Finally, syntactic complexity was measured by number of words per utterance, defined as the stream of speech produced under one intonation contour between pauses (Ortega, 1999: 124).

Results showed greater fluency and complexity in the planned condition, but there was no difference in lexical range. In addition, effects of planning time on accuracy were significant in terms of noun-modifier agreement but not in terms of use of the Spanish article system. Ortega interprets her results as corroborating previous studies on the effects of planning on L2 oral production, in that planning time favors fluency and complexity in speech. She also argues that limited attentional resources do not allow L2 speakers to focus simultaneously on all dimensions of L2 speech production, which penalizes accuracy. Ortega also claims that the lack of significant effects of planning time on lexical range may be due to methodological problems in the way lexical range was assessed.

At this point, the conclusion we can draw from the three studies reviewed above is that fluency, accuracy, complexity, and lexical density in L2 speech production interact in nonlinear ways when planning time prior to performance is controlled. These studies are relevant to the area of task-based instruction, in general, and to the construct of planning, in particular. Since the main variable under investigation is planning time, the central results of these studies--that is, the effects of planning time--are of no direct relevance to the present research project. However, more important than the effects of planning time on L2 oral performance is the way L2 speech was assessed in these studies. Fluency, accuracy, complexity, and lexical density seem to be reasonable dimensions of L2 speech production that can be operationalized by means of relatively straightforward variables, thus making it possible to approach coherently an otherwise overly complex skill (Bygate, 1998).

The review of the literature presented in Chapter 2 and in this Chapter makes it apparent that speech production, in general, is a neglected area in the research agenda of the working memory community. It also indicates that L2 skills are not an issue in this agenda, despite the importance of these skills in the real world. Similarly, L2 researchers have shown only slight interest in the working memory construct and have, as a rule, overlooked the importance of investigating more systematically the nature of L2 speech performance. The present study aims at bringing these two areas together. The next chapter outlines the method used for examining the relationship between working memory capacity and L2 speech performance.

CHAPTER 4

METHOD

In order to address the research questions and hypotheses presented in the Introduction, an experiment assessing individual's working memory capacity and their L2 speech production was carried out. The experiment was carried out at the University of Minnesota, where this researcher was completing her third year of doctoral studies as a visiting graduate student on a grant from CAPES.

In order to maximize the stay at the University of Minnesota and gather as much data as possible to implement future research projects on the relationship between working memory capacity and L2 speech production from a developmental perspective, data were collected from the same set of participants on two occasions--in the beginning and again at the end of the winter term of 1999. However, to examine the questions and hypotheses motivating the present study, only the data collected in the beginning of the term are relevant. This chapter describes the methods for applying the experiment and for analyzing the data collected for the present study. First, it describes the participants, instruments, design, and data collection procedures. Then, it presents the method for establishing reliability of the measures used. Finally, it outlines data transcription procedures and the data analysis method.

4.1 Participants

Participants for this study were 13 advanced learners of English as a second language (ESL) at the Minnesota English Center (MEC), at the University of Minnesota. All participants were enrolled in ESL classes at the MEC and all had 20 hours of ESL classes a week covering speaking, listening, reading, and writing, in addition to grammar and pronunciation. All participants were from the same "Advanced Speaking Skills" class. This class had a total of 14 students, 13 of which were selected to participate in this study on a volunteer basis. The fourteenth student started the course the week in which the experiment was carried out--the third week after the beginning of the winter term--and thus was not selected to participate in the study due to the difference in the number of hours of formal instruction at the MEC in comparison to the rest of the group.

Participants' ages ranged from 18 to 41 years, with a mean of 28.2, thus a predominantly adult population. Except for one participant, who would resume college in the fall quarter of the same year, all participants held a university degree from their countries, and their areas of work included electrical engineering, chemistry, business, journalism, psychology, and education. At the time of data collection, all participants were preparing to pursue a graduate degree at the University of Minnesota. Participants' length of residence in the United States, prior to data collection, varied from one to three weeks, but three of them had visited the country before.

Of the 13 participants, 8 were female and 5 were male. There were 4 Brazilians, 1 German, 2 Koreans, 2 Japanese, 1 Chinese, 1 Israeli, 1 Indonesian, and 1 Turk. All participants reported having studied English in their countries during their school years. Number of years of formal instruction in English varied from 8 to 13 years, with a mean of 10.46 years. However, participants' level of proficiency in English could not be rigorously controlled in the present study. Apart from their reported number of years studying English in formal settings, the only other information about their proficiency level was an in-house placement test subjects had to take, to which this researcher was denied access.

Assuming that the 13 participants had similar scores on the placement test, and based on their educational background as well as on their reported number of years studying English, it was thought that they would form a relatively homogeneous group both in terms of educational history and level of proficiency in English, and would thus be suitable participants for the present study.

4.2 Instruments

4.2.1 Materials and equipment

The experiment consisted of four tasks: two tasks aimed at measuring working memory capacity and two tasks aimed at eliciting speech production in the L2. The working memory tasks were conducted using an IBM ThinkPad Laptop Computer with a VGA monitor. Participants' responses in all four tasks were recorded on magnetic tape using a SONY Voice Operated Recording tape recorder. A separate tape was used for each participant. Participants' performance on the first speech production task--the description task--was timed through the use of a SPORTLINE Model 220 stopwatch to signal the beginning and end of the time allotted for the task (2 minutes).

4.2.2 Assessment and measures of working memory capacity

Participants' working memory capacity was assessed by means of the speaking span test (Daneman, 1991) and the operation-word span test (Turner & Engle, 1989). The versions used in the present study were those used in Fortkamp (1998, 1999), designed to measure working memory capacity in English as a foreign language.

4.2.2.1 The speaking span test

Based on Daneman (1991), the speaking span test was constructed with 60 unrelated one-syllable words, organized in three sets each of two, three, four, five, and six words (see Appendix A, item I). Each word was presented on the middle line of the computer video screen for 1 second and was accompanied by a beep. Participants were required to read the word silently. Ten milliseconds after the word had been removed, the next word in the set would appear beside the place the previous word had been presented, on the same line. This procedure was followed, each word slightly to the right of the previous word, until a blank screen signaled that a set had ended. Participants were then required to produce orally a sentence for each word in the set, in the order they had appeared and in the exact form they were presented. For instance, after being presented with the following set of three words:

club spring knife

a subject produced the sentences:

"I went to the club yesterday"

"The spring is a beautiful season"

"That knife is very dangerous"

Participants were told that there were no restrictions as to the length of the sentences, but they were required to be syntactically and semantically acceptable. After each subject finished generating the sentences for a given set, the next set would be presented, and this procedure was followed until all sets had been presented. A two-word set was presented first, followed by a three-word set, and so on, ending the sequence with a six-word set. The same sequence was repeated two more times until the 60 words had been presented. Participants were given practice trials and the actual span test would begin only when the subject reported being comfortable with the test.

Following Daneman (1991) and Daneman & Green (1986), the measure applied to a subject's speaking span in English as an L2 was his/her total performance on the test, that is, the total number of words for which a syntactically and semantically acceptable sentence was produced--in this case, the maximum being 60. To be scored as acceptable, the sentence should contain the word in its original order within the set, be in its original form of presentation, and should not present deviations from grammatical English. In addition, the sentence should be semantically viable, in the sense that it could be understood without much contextual information.

4.2.2.2 The operation-word span test:

Following Turner & Engle (1989), the operation-word span test was constructed with 60 operation strings and 60 English words. Each operation string was accompanied by a one-syllable word to the right of it and presented one at a time on the middle line of the computer video screen. For each subject the pool of 60 operations was randomly paired with the pool of 60 words, to minimize test difficulty effects. The 60 combinations of operation plus word in English were arranged in three sets each of two, three, four, five, and six. Participants were required to pace themselves through the operation string, evaluate whether the result presented for the operation was true or false, then press the letter T or F on the keyboard, while retaining the word accompanying the operation for subsequent recall. After solving the operation, checking if the result presented was true or false, and retaining the word for subsequent recall, participants would press "enter" and the next operation-word pair would appear, as shown in the example of a two-pair set below:

(9 X 3) + 2 = 4	DANCE
(7/1) + 6 = 12	FILE

This procedure was followed until a blank screen followed by a beep signaled that a set had ended. Participants were then required to recall the words accompanying each operation in the order they had appeared and in the exact form presented. Contrary to the speaking span test, in this test the number of operation-word pairs presented before recall was also randomized to prevent participants from being able to predict the number of words they would have to recall. The operation strings (see Appendix B, item I) were taken from Cantor & Engle (1993) and the words (see Appendix B, item I) from Cantor & Engle (1993) and La Pointe & Engle (1990). As in the speaking span test, practice trials were given to each subject and the actual span task would begin only when the subject felt comfortable enough with the test. As in Engle, Cantor, & Carullo (1992), a participant's' operation-word span was his/her total performance on the test, that is, the total number of words correctly recalled--in this case, the maximum being 60.

As already said, the two different working memory span tests used in the present study reflect two different views of working memory capacity (Cantor & Engle, 1993; Conway & Engle, 1996; Engle, Cantor, & Carullo, 1992; Daneman & Merikle, 1996; Turner & Engle, 1989). The operation-word span test was first devised by Turner & Engle (1989) to test the hypothesis that working memory capacity is general in nature and not specific to the cognitively complex language-related task being performed. Thus, the general capacity hypothesis (Cantor, Engle, & Hamilton, 1991; Conway & Engle, 1996; Engle, Cantor, & Carullo, 1992) would predict significant correlations not only between the speaking span test and measures of L2 speech production, but also between the operation-word span test and the several measures of L2 speech production used in the present study. The speaking span test, however, reflects the task specific hypothesis (Conway & Engle, 1996; Cantor & Engle, 1993; Daneman, 1991; Daneman & Carpenter, 1980 and 1983; Engle, Cantor, & Carullo, 1992), which posits that working memory capacity is functional and dependent on the complex task being performed. The task specific hypothesis would predict that only the speaking span test, but not the operation-word span test, correlates significantly with the measures of L2 speech production.

4.2.3 Assessment of L2 speech production

Participants' L2 speech production was elicited by means of a description task and a narrative task. The picture description task and the narrative task were chosen for this study because both tasks are traditionally used in the elicitation of L1 and L2 speech production and because they can be manipulated so as to be monologic rather than interactive tasks (Ortega, 1999). Several researchers have suggested that L2 speech elicited through interactive tasks (e.g. dialogues and interviews) may be a learner's least fluent variety of interlanguage (Freed, 1995, p. 143; Lennon, 1990, p. 397; Olynyk, d'Anglejan, & Sankoff, 1990, p. 153), which would make them probably inadequate tasks to obtain speech for detailed analysis of temporal variables, as is the case in the present study.

4.2.3.1 Picture description task

Participants were presented with a colorful picture taken from a popular magazine (see Appendix H). The picture, an ad for a TV channel, portrays two different moments of the history of the human race. On the left-hand side of the picture, an Egyptian scenario is presented with half of a statue of an Egyptian sphinx. On the right-hand side of the picture, completing the other half of the statue, an astronaut in space is portrayed. At the top of the page, the name of the TV channel is shown, followed by the phrase "The official network of every millennium".

Participants were required to describe the picture and express their opinion about the message, if any, conveyed by it. They were given 2 minutes to perform this speaking task, measured with a stopwatch. Participants were explicitly instructed to give as much information as possible. There were given time to analyze the picture and were told that, if they thought it necessary, they could plan what they would say (see Appendix D, Item III for instructions). In planning, they were allowed to make notes of words and sentences they wanted to use in their description. They were also free to check any vocabulary difficulties they had before the beginning of the task and to use their notes, while speaking, for the specific words and sentences planned. However, they were not allowed to speak as if they were reading their notes.

It was thought that allowing participants to analyze the picture, plan what to say, and check vocabulary difficulties prior to performance would optimize their use of time and encourage them to produce speech that was more fluent, accurate, complex, and lexically dense by prior activation of the linguistic resources necessary to complete the task. The task would begin when participants signaled they were ready.

4.2.3.2 Narrative task

This task was intended to elicit speech in a more natural situation. There were no time constraints on the performance of this task. Participants were required to retell a movie they had seen that they had liked or disliked. Again, participants were explicitly instructed to give as much information as possible. As in the descriptive task, participants were free to plan what they would say and to check any vocabulary difficulties they had. They could also check their notes while speaking, but were not allowed to speak as if reading them (see Appendix D, item IV for instructions). Participants were also instructed to talk as much as possible. As in the description task, it was thought that allowing participants to plan what to tell and check vocabulary difficulties would contribute to their producing speech that was more fluent, accurate, complex, and lexically dense by prior activation of information.

Although there were no time constraints on the performance of the narrative task, only the first two minutes of talk of all participants were selected for analysis in this study. This cut-off point was chosen for two reasons: First, the principal reason for not setting time constraints on the performance of this task was to counterbalance the description task and minimize the emergence of a faster speech rate due to time pressure and anxiety. Second, of the 13 participants, only 5 participants actually spoke more than 2 minutes, thus showing that the general tendency of the group was to accomplish the narrative task within the first two minutes.

Thus, the 13 participants of this study provided 2 speech samples each, a picture description and a narrative, resulting in a total of 26 samples. All participants' speech samples were tape-recorded and later transcribed (see section 4.4 for transcription procedures).

4.2.4 Measures of L2 speech production

Participants' speech production in English as an L2 was measured in terms of fluency, accuracy, complexity, and weighted lexical density in the picture description task and the narrative task (described in sections 4.2.3.1 and 4.2.3.2, respectively) . These measures were adapted from the framework proposed by Skehan (1996; 1998) in the area of task-based instruction and have been extensively used in research on the effects of planning time on L2 speech production (Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999). Examining L2 speech production through these four aspects seems to give us a global view of L2 speech performance since they are intended to capture complementary aspects of this multidimensional process.

4.2.4.1 Fluency

Following Skehan (1996; 1998) and Foster & Skehan (1996), in this study the notion of fluency was operationalized so as to reflect continued performance in real time. Four temporal variables were assumed to reflect this notion of fluency in L2 speech production. The selection of these variables was based on the findings of previous studies showing that they are salient features in the display of continuous L2 speech (Ejzenberg, 1992; Freed, 1995; Lennon, 1990; Riggenbach, 1989, and Temple, 1992):

(1) Speech rate in two versions: Unpruned and pruned (Lennon, 1990; Ortega, 1999). Speech rate unpruned was calculated by dividing the total number of semantic units produced, including repetitions, by the total time--including pause time and expressed in seconds--the subject took to complete the speech production task. The resulting figure was then multiplied by 60 to express number of semantic units per minute. Semantic units consisted of complete words and partial words (Ejzenberg, 1992; Riggenbach, 1989; Freed, 1995). Partial words consisted of at least a consonant and a vowel that could be recognizable as a syllable. Speech rate pruned was calculated the same way as speech rate unpruned, but all semantic units that were repeated (excluding repetitions for rhetorical effect and including only immediate repetitions) or that were abandoned before completion were excluded from the count. In both measures, contractions were counted as one word. Speech rate unpruned is a more general measure that is assumed to reflect the relationship of articulation to silence. Speech rate pruned is a more specific measure that reflects a more straightforward expression of ideas and unimpeded articulation of words.

(2) Number of silent pauses per minute: As already discussed, there seems to be no agreement on the cut-off point to be used in determining silent pauses in L2 speech production (Towell, Hawkins, & Bazergui, 1996). Foster & Skehan (1996) and Mehnert (1998) used 1 second as a cut-off point for silent pauses; Freed (1995) used .4 seconds or longer; Grosjean & Deschamps (1972, 1973, 1975) and Raupach (1987) used .25 seconds; Lennon (1990) used .2 seconds and Riggenbach (1989) established a threefold distinction between micropauses (silence of .2 seconds or less), hesitations (.3 to .4 seconds), and unfilled pauses (.5 seconds to 3 seconds). In the present study, a break in the speech flow equal to or larger than .5 seconds was considered a silent pause. This cut-off point was chosen because, as Deese (1980), Fillmore (1979) and Riggenbach (1989) suggest, silent pauses of .4 seconds or less are frequent in nonnative speech production and may be in the range of normal speech.

All unfilled pauses in the speech samples were first located and measured with a stopwatch during transcription. As Crookes (1991) and Griffith (1991) point out, reliability in the measurement of pauses is a problematic issue in the research on speech production. Thus, in order to establish the length of unfilled pauses in a more precise way, participants' speech samples were also copied onto a CD-ROM so that the location and length of silent pauses could be verified by a software system especially designed for speech analysis. The software used was SpeechStation2 (Sensimetrics, 1998). Through location and measurement of all unfilled pauses on the spectrogram and waveform of each speech sample displayed by the software, it was possible to determine in a more reliable way the length, in milliseconds, of every silent pause in each speech sample. The total number of silent pauses equal to or longer than .5 seconds in each subject's picture description and narrative was determined and divided by the total time taken to speak, in seconds. The resulting figure was then multiplied by 60 to express number of silent pauses per minute.

(3) Number of hesitations per minute: In the present study, unfilled pauses of .49 seconds or less, filled nonlexical pauses (e.g. "uh" and "uhm"), immediate repetitions and partial words were considered hesitations. The location and length of unfilled pauses of .49 seconds or less were determined the same way as described in (2) above. In each subject's picture description and narrative, all unfilled pauses, filled nonlexical pauses, immediate repetitions and partial words were counted and summed up. The total number of hesitations was then divided by the total time taken to speak, in seconds, and the resulting figure multiplied by 60 to express number of silent pauses per minute.

(4) Mean length of run: In this study, mean length of run was calculated as the mean number of semantic units--words and partial words, including repetitions, immediate or not--between pauses--unfilled, of any length, and filled (e.g., *uhm* and *uh*). Each subject's mean length of run in the description and narrative task was determined by dividing the total number of semantic units produced by a selected number of pauses. A criterion was established that when there were chunks of filled and unfilled pauses, the whole chunk was counted as one pause, as shown in the following example (see section 4.4 for transcription conventions):

(Subject 9: Picture description task)

> "but still the two different sides give a very different impression and a different mood **(.) uh** and **uh** give the the: **uhm (1.8)** the impression of really two very different pictures that are just **(.)** thrown together..."

In this excerpt, there are 6 pauses: 3 unfilled pauses--2 shorter than .5 seconds and one longer than .5 seconds--and 3 filled nonlexical pauses. Of these, only 4 pauses were taken into the count of mean length of run. The short silent pause **(.)** followed by **uh** and the filled nonlexical pause **uhm** followed by a long silent pause **(1.8)** were counted as one pause each. This procedure was necessary because taking all pauses into the count of the mean length of run would yield a misleading figure since there are no semantic units between pauses in the chunks.

4.2.4.2 Accuracy

As Foster & Skehan (1996, p. 304) and Mehnert (1998, p. 91) argue, when the study is exploratory in nature, as is the case of the present study, and no hypotheses can be made as to what specific forms of language will occur in a given speech sample, using general measures of accuracy seems an appropriate procedure. In the present study, the number of errors per 100 words was used as a general measure of accuracy. The analysis to determine number of errors counted all errors in syntax, morphology, and lexical choice, including repetitions. Errors in pronunciation and intonation were not included in the analysis. Errors that were immediately corrected were not counted. The total number of errors in each subject's speech sample was divided by the number of semantic units produced and the resulting figure multiplied by 100 to express number of errors per 100 words.

4.2.4.3 Complexity

Foster & Skehan (1996) and Skehan (1998) propose that subordination is an index of internal complexity of speech. According to Quirk & Greenbaum (1973), "subordination is a non-symmetrical relation, holding between two clauses in such a way that one is a constituent part of the other" (p. 309). As such, subordination is realized through the dependent clause. In the present study, complexity of speech was measured in terms of number of dependent clauses per minute. Following Mehnert (1998, p. 90), analysis of number of dependent clauses included finite and nonfinite subordinate clauses, coordinate clauses with subject deletion, coordinate clauses with subject and finite verb deletion, and infinitive constructions, excluding infinitives with modal verbs. The total number of dependent clauses in each speech sample was divided by the time taken to accomplish the task--in seconds--and the resulting figure was then multiplied by 60 to express number of dependent clauses per minute.

4.2.4.4 Weighted lexical density

Following Mehnert (1998) and O'Loughlin (1995), lexical density of the speech data collected for this study was measured by weighted lexical density. In order to establish the weighted lexical density of each speech sample, it was first necessary to classify all linguistic items as grammatical or lexical items. In his framework of analysis, O'Loughlin (1995) points out that, because there is no direct correspondence between linguistic items and words in English, the notion of item may be more revealing in the analysis of lexical density in spoken data than the traditional concept of "word". This distinction is important because multiword verbs, phrasal verbs, and idioms, although consisting of more than one word, are counted as one linguistic item. In addition, the notion of "item" encompasses particles such as "oh" and "wow" that are frequently present in spoken data. Thus, in the present study, it was the notion of "linguistic item" which was used in determining weighted lexical density. The following criteria were used to assign items to the grammatical or lexical category (based on O'Loughlin, 1995:228):

<u>(A) Grammatical items:</u>

1- All modals and auxiliaries.

2-All determiners, including articles, demonstratives and possessive adjectives, quantifiers (*some*, *any*), and numerals (cardinal and ordinal).

3-All pronouns (*she*, *it*, *they*, *someone*, *something*) and *this* and *that* when used to replace clauses.

4-Interrogative adverbs (*what*, *when*, *how*) and negative adverbs (*not*, *never*).

5-All contractions of pronouns and auxiliary verbs. Contractions were counted as one item.

6-All prepositions and conjunctions.

7-All discourse markers including conjunctions (*but, so, and*), sequencers (*next, finally*), particles (*oh, well*), lexicalized clauses (*you know, I mean*), and quantifier phrases (*anyway, somehow, whatever*).

8-All lexical filled pauses (*so, well*)

9-All interjections (*gosh, really, oh*)

10-All reactive tokens (*OK, no!*)

(B) Lexical items:

1- Nouns, adjectives, verbs, adverbs of time, manner, and place. Multiword verbs, phrasal verbs, and idioms were counted as one item. Contractions of pronouns and main verbs were counted as one lexical item.

In addition, lexical and grammatical items were divided into high-frequency lexical and grammatical items and low--frequency lexical and grammatical items. A high-frequency lexical or grammatical item was one appearing more than once in the same speech sample. Different word forms of the same lexical or grammatical item formed by inflection or derivation (e.g., *go/went, study/student, this/these*) were considered repetitions and thus counted as a high-frequency lexical or grammatical item. The numbers of high- and low-frequency lexical and grammatical items in each speech sample were first tallied as frequency counts. Following Mehnert (1998), high-frequency lexical and grammatical items were given half the weight of the low-frequency lexical and grammatical items. The total number of weighted lexical items was thus determined, then divided by the total number of weighted linguistic items and multiplied by 100 so as to obtain the percentage of weighted lexical items over the total number of weighted linguistic items in the speech sample.

In summary, a set of 10 variables in two tasks was used to examine the relationship between working memory capacity and L2 speech production. Table 4.1 shows how the information was organized for each subject's speech sample:

Table 4.1
Model of summary of subject's scores

Part.	Task	SST	OWST	SRU	SRP	SPpm	Hpm	MLR	Acc.	Comp.	WLD

Note. Part.: subject identification
Task: description/narrative
SST: Speaking span test
OWST: Operation-word span test
SRU: Speech rate unpruned
SRP: Speech rate pruned
SPpm: Pauses per minute
Hpm: Hesitations per minute
MLR: Mean Length of run
Acc: errors per 100 words
Comp.: number of dependent clauses per minute
WLD: weighted lexical density

4.3 Procedures

All necessary procedures to obtain permission to carry out this study in an American university were followed as required by the Institutional Review Board of the Research Subject's Protection Program at the University of Minnesota. This included obtaining prior permission from the Dean of the College of Liberal Arts and the Director of the Minnesota English Center and submitting to the Board copies of all the instruments used as well as the recruitment script and the consent form participants would be required to sign. The permission to carry out the study was given at the end of the fall quarter of 1998.

Prior to data collection, in the first week of the winter quarter of 1999, the researcher contacted the teacher responsible for the "Advanced Speaking Skills" class (Speaking 235-1) to ask permission to invite students to participate in the study. In order not to use class time and also to leave students totally free to choose whether they wanted to participate in the study or not, students were first informed by their teacher of what the study was about and that I would be inviting them to participate. After that, I contacted participants by email (see Appendix E for the letter sent) and arrangements were made either by phone or email.

The data for this study were collected individually with each subject in a room at the Institute of Linguistics at the University of Minnesota in the third week of the winter quarter of 1999. Each subject was given the set of 4 tasks--two assessing working memory capacity and two assessing L2 speech production-- in one session that lasted about one hour and a half. Prior to performing the tasks, participants were given a consent form to read and sign (see Appendix F) and filled in a biographical information/personal data sheet (see Appendix G). All participants were given first the memory span tests--the speaking span test followed by the operation-word span test--and then the L2 speech production tasks--the picture description task followed by the narrative task[1]. Participants were explicitly told that the span tests were memory tests and that it was necessary to focus their attention on the stimuli. They were also told that they should try to speak as much as possible in the speech production tasks. As described in 4.2.2.1 and 4.2.2.2 all participants were given a training session before taking the span tests. For the sake of uniformity, each subject first received written instructions and then an oral explanation followed (see Appendix D, items I and II for instructions). Instructions were given in English to all participants, including Brazilians, by this researcher, who also collected the data.

---

[1] Participants were also required to answer 3 strategy use questionnaires aimed at assessing their use of strategies for improving speaking skills (see Appendix I). These are not analyzed in the present study.

4.4 Data transcription procedures

Participants' responses to all tasks were tape-recorded for transcription. All participants' responses were transcribed (see Appendix A, item II, for the speaking span test, Appendix B, item II, for the operation-word span test, and Appendix C, for speech data). Participants' speech data were duplicated to ensure against damage to the original recordings. The transcription system used was adapted from Jefferson (1979), which has been used in recent studies on L2 speech production (Ejzenberg, 1992; Riggenbach, 1989). Below is an explanation of the conventions used in the transcription of the speech data.

As explained in 4.2.4.1, unfilled pauses were first located and timed with a stopwatch. The length of each unfilled pause was again measured by means of a software designed to analyze speech data (SpeechStation2). The software is more precise than a stopwatch in determining the length of unfilled pauses in that the analyst can use both the waveform and the spectrogram of the segment being analyzed as additional tools to locate and measure the pause. Furthermore, the visual display makes it easier to always select the same portion to be measured. Each subject's description and narrative was recorded on a CD-ROM as an individual file. Thus, there were 26 files. The software displays each file in segments of 2 seconds. The display always presents the frequency scale, spectrogram, pitch track, and the waveform of the 2-second segment. The unfilled pauses--located and measured first with the stopwatch--could then be double-checked by selecting the portion of the segment equivalent to the pause and verifying on the screen the beginning, end, and duration of the pause in milliseconds.

Unfilled pauses equal to or longer than .5 seconds are indicated by the time period in parentheses--for instance, (1.8) means a silent pause of one second and eight hundred milliseconds. Unfilled pauses shorter than .5 seconds are indicated by a period inside a parenthesis--for instance, (.). Filled nonlexical pauses are indicated by "uh", "uhm", and "uh-uh".

Sound stretches are indicated by colons (:), as in "the:". Laughter is indicated by the word "laugh" in parenthesis. A period indicates falling intonation. A question mark indicates rising intonation. An exclamation mark indicates enthusiasm on the part of the speaker. Underlined words indicate greater stress than the other words in the same environment. Finally, although the speech sample of each subject was transcribed in its entirety, only the first two minutes of each sample was coded following the procedures described above.

4.5 Interrater reliability

After determining the score for each variable of L2 speech production in all participants' speech samples, the samples were submitted to different raters for computing interrater reliability. Three raters reanalyzed different portions of the data, following the criteria the researcher had used. All three raters are teachers of English as a foreign language and have been in the profession for more than 10 years. Rater 1, who is pursuing a Ph.D. in education, reanalyzed 53.8% of the data--7 descriptions and 7 narratives--for the fluency variables and agreement reached 92%. Rater 2, a native speaker of English who holds a Ph.D. in Applied Linguistics, reanalyzed all the data for accuracy and agreement reached 87,76%. Rater 3, who holds an MA in Applied Linguistics, reanalyzed 69.23 % of the data for complexity and agreement reached 100%. The same rater reanalyzed 53.8% of the data for lexical density, and agreement reached

98.37% after lengthy discussion on the criteria used. All discrepancies were resolved by discussion.[2]

4.6 Reliability of span tests

Reliability estimates were computed for the speaking span test and the operation-word span test using Cronbach's alpha formula of internal consistency[2], also used by Turner and Engle (1989) and Engle et al. (1992). Internal consistency, according to Klein and Fiss (1999), measures the degree to which different test items measure the same variable. Following Turner and Engle (1989), the reliability estimates were based on three composite scores computed for every subject in the two span tests. As already explained, the span tests were organized in three trials each of two, three, four, five, and six sequences of to-be remembered items. In the case of the speaking span test, the total number of correct sentences in the first trials of all set sizes was calculated as one span. A second span was calculated from the total number of correct sentences in the second trials of all set sizes. Finally, the total number of correct sentences in the third trials of all set sizes was calculated as the third span. Thus, each subject had 3 individual spans in the speaking span test and intercorrelations were computed among the three spans of each subject. The same procedure was followed to establish intercorrelations among the operation word-span test composite scores. Reliability estimates of the speaking span test and the operation-word span test were .88 and .77, respectively. There are no reports, in the literature, on reliability scores for the speaking span test. However, given that the closer to 1 the alpha coefficient is, the higher the internal consistency of the test, the estimate obtained in the present study for the speaking span test is at an acceptable level.

---

[2] Several researchers have used interrater analyses of portions of the data, including Mehnert (1998) and O'Loughlin (1995).

[2] See Appendix L for formula.

For the operation-word span test, reliability scores reported in the literature range between .74 and .81 (Engle et al. 1992; Engle & Turner, 1989; Klein & Fiss, 1999). Thus, based on the alpha coefficients reported on the literature, the internal consistency statistic for the operation-word span test is also acceptable.

4.7 Data analysis

The goal of this research was twofold: (1) to examine the relationship between working memory capacity, as measured by complex span tasks, and L2 speech production, as measured by several variables covering fluency, accuracy, complexity, and lexical density of speech, and (2) to examine the nature of this relationship, that is, whether it is task-dependent or a general capacity, independent of the task being performed. The approach adopted to assess the relationship between working memory capacity and L2 speech production was that traditionally used in most studies on working memory capacity and complex cognitive behavior, the correlational one (Atkins & Baddeley, 1998; Daneman, 1991; Daneman & Carpenter, 1980, 1983; Daneman & Green, 1986; Daneman & Merikle, 1996; LaPointe & Engle, 1990; Miyake & Friedman, 1998; Turner & Engle, 1989; Roberts & Gibson, 1999; Woltz, 1988, among many others). In the present study, adopting the correlational approach involved determining the degree of association between working memory capacity and measures of L2 speech production and determining whether working memory capacity is a significant predictor of L2 oral performance. The main analytic technique used to measure the amount and significance of the relationship between working memory capacity and measures of fluency, accuracy, complexity, and weighted lexical density was the Pearson Product

Moment Coefficient of Correlation (*r*). The predictive power of working memory capacity was determined through simple linear regressions.

The present study pursued two research questions, now formulated in specific terms:

Research Question 1: Is there a statistically significant relationship between task-dependent working memory capacity, as measured by the speaking span test, and measures of fluency, accuracy, complexity, and weighted lexical density in L2 speech production?

Research Question 2: Is there also a statistically significant relationship between independent working memory capacity, as measured by the operation-word span test, and fluency, accuracy, complexity, and weighted lexical density in L2 speech production?

From these two research questions, 5 hypotheses follow. The first research question generates the following set of hypotheses:

Hypothesis 1: There will be a relationship between working memory capacity, as measured by the speaking span test, and fluency in L2 speech production, as measured by speech rate unpruned, speech rate pruned, pauses per minute, hesitations per minute, and mean length of run. Working memory capacity will correlate positively with speech rate unpruned, speech rate pruned, and mean length of run, and negatively with pauses per minute and hesitations per minute.

Hypothesis 2: There will be a negative correlation between working memory capacity, as measured by the speaking span test, and accuracy in L2 speech production, as measured by the number of errors in syntax, morphology, and lexical choice per hundred words.

Hypothesis 3: There will be a positive correlation between working memory capacity, as measured by the speaking span test, and complexity in L2 speech production, as measured by the total number of dependent clauses.

Hypothesis 4: There will be a positive correlation between working memory capacity, as measured by the speaking span test, and weighted lexical density in L2 speech production.

The second research question generates the following hypothesis:

Hypothesis 5: In addition to the relationships predicted in Hypotheses 1-4, there will be a statistically significant relationship between working memory capacity, as measured by the operation-word span test, and fluency, accuracy, complexity, and weighted lexical density in L2 speech production, in the directions predicted in Hypotheses 1-4.

Correlational analyses have some inherent limitations. Several researchers (e.g. Carpenter, Miyake, & Just, 1994; Harrington, 1992; McClave, Dietrich, & Sincich, 1997; Miyake & Friedman, 1998; Walsh, 1990) have acknowledged that correlational data are ambiguous in nature and open to alternative interpretations. Moreover, correlational analyses are particularly problematic when applied to small sample sizes, in which the participants' range of scores in one of the variables is generally restricted, thus lowering the correlation (Goodwin, 1995). Small sample sizes in correlational studies are also sensitive to outliers, individuals with atypical behavior whose impact on the data can lead to serious misinterpretation of the relationship between two variables.

In order to overcome these limitations, many researchers (e.g., Cantor & Engle, 1993; Just & Carpenter, 1992; McDonald, Just, & Carpenter, 1992; Miyake, Carpenter, & Just, 1994, among many others) adopt an analysis of variance (ANOVA) oriented approach in which they identify participants with high and low spans and, treating them as two independent and homogeneous groups, compare their performance on complex tasks (Engle, Cantor, & Carullo, 1992). This approach, however, also has its drawbacks since it ignores data from participants in the middle group, often overestimates the relationship between variables, and generally requires an arbitrary cut-off point to determine what is a high span and a low span when the scatterplot does not indicate any natural grouping of the data (Roberts & Gibson, 1999).

Although some researchers suggest that combining the correlational and the ANOVA approaches strengthens both techniques (e.g., Engle et al, 1992), other researchers (e.g. Roberts & Gibson, 1999) claim that the correlational approach is more appropriate if no natural grouping of the data is shown by the scatterplot and if the objective is to study the common variance between sets of variables. Thus, for this reason the present study adopted the correlational approach.

It is often the case that, in examining linear relationships between pairs of variables, correlation and linear regression analyses are used together (Goodwin, 1995; McClave, Dietrich, & Sincich, 1997; Walsh, 1990). Whereas correlation techniques determine whether there is a relationship between any two variables and indicate the strength of this relationship, regression techniques allow us to predict the value of one variable based on knowledge of another. Together, these two analytic techniques allow the researcher to go beyond global correlations and investigate further the linear relationship between any two variables. Thus, to complement the correlational analyses and gain further insights on the relationship between working memory capacity and L2 speech production, simple linear regressions were also performed in order to determine the predictive power of working memory capacity on measures of L2 speech production.

Statistically speaking, the number of participants in the present study--13--is not high enough to make the results obtained generalizable to the population that they represent. To minimize this limitation, every effort was made to ensure that the findings of the present study be a possible indication of trends in the relationship between working memory capacity and L2 speech production. It was decided that if outliers were found in the data, two additional analytic techniques would be used to investigate further the degree to which the association between working memory capacity and the measures of L2 speech production was affected by the presence of outliers: the Pearson Product Moment Coefficient of Correlation computed without the outliers and Spearman's Rank Order Correlation (r_s) [3], including the scores of all participants.

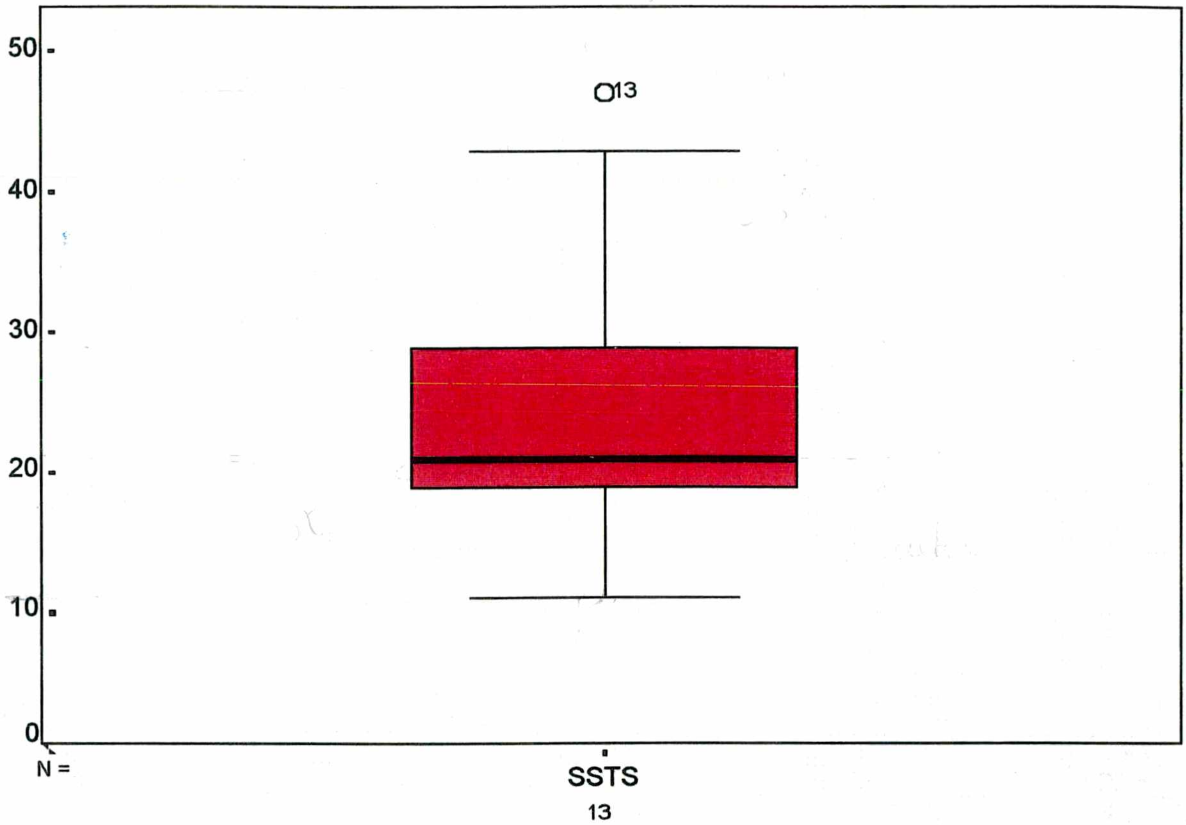
In Spearman's Rank Order Correlation, the values of a set of variables are ordered into ranks and the coefficient provides a measure of association between ranks (Walsh, 1990; McClave et al., 1997). As with the Pearson Product Moment Correlation (r), the magnitude of r_s indicates the direction of the association as well as its strength. Spearman's Rank Order Correlation is, however, a non-parametric test whereas the Pearson Product Moment Coefficient of Correlation is a parametric test[4].

[3] As suggested by Rocha (December 1999, personal communication), Spearman's Rank Order Correlation (r_s) is generally less affected by the presence of outliers.
[4] It should be noted that the correlation coefficient r_s can not be squared to be interpreted as the amount of variance accounted for by the variables. For this purpose, it is necessary to compute Pearson correlation (r).

While the objective of the two tests is the same, that is, to determine the degree of association between two variables, the method of calculating the probability of rejecting the null hypothesis is different in the two tests. As Green & D'Oliveira (1982) note, ordering participants' scores into rank order measures variability in their scores only indirectly. In the Pearson Correlation, there is more information about differences in scores because the exact proportion of the total variability in participants' scores is taken into account, making the test more powerful, so to speak. However, there is a lot of controversy on whether one type of test is really more powerful than the other. In any event, the objective of running Spearman's Rank Order Correlation would be to look at the data from another angle, to verify whether the same trend would be found in the relationship between working memory capacity and measures of L2 speech production.

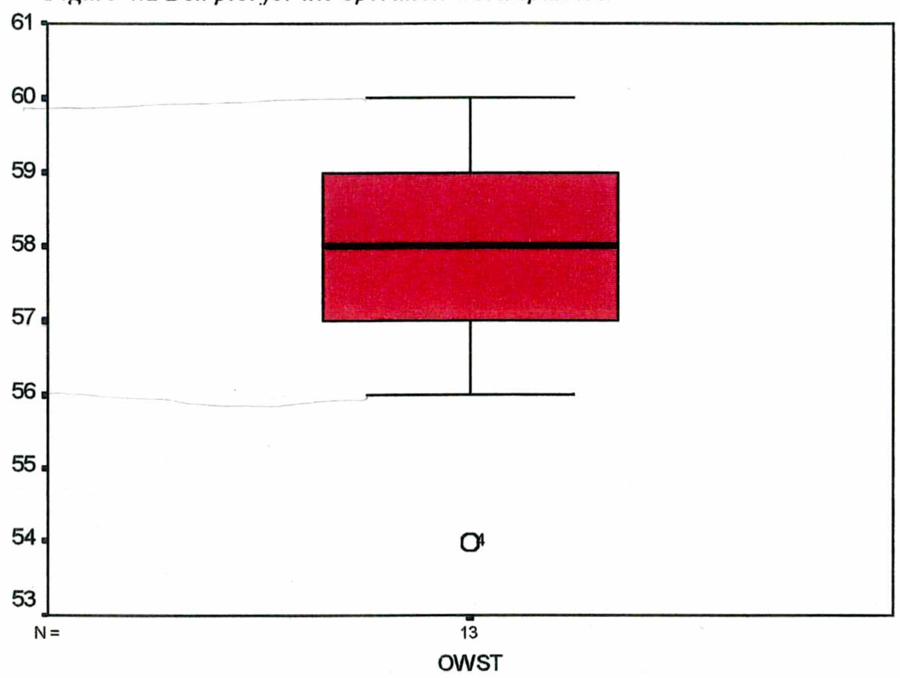
Since scores on the working memory span tests are the main variables of the present study, the data related to these tests were checked, through a box plot, to see if they were evenly distributed, a prerequisite of correlational analysis. Briefly, the box plot is a descriptive measure based on the quartiles of a data set. Quartiles divide the data into four sets, each representing 25% of the measurements. Thus, the lower quartile corresponds to the 25th percentile of a data set, the middle quartile to the 50th percentile and the median *M*, and the upper quartile to the 75th percentile. The box plot indicates the distance between the lower and upper quartiles, showing the sample variability. The line across the box shows the median of the sample. The vertical lines emanating from the box, called the upper and lower whiskers, indicate the highest and lowest values in the sample, excluding outliers. Measurements that fall beyond these lines are outliers. Figure 4.1 shows the box plot for the speaking span test:

Fig. 4.1 Box plot for the speaking span test



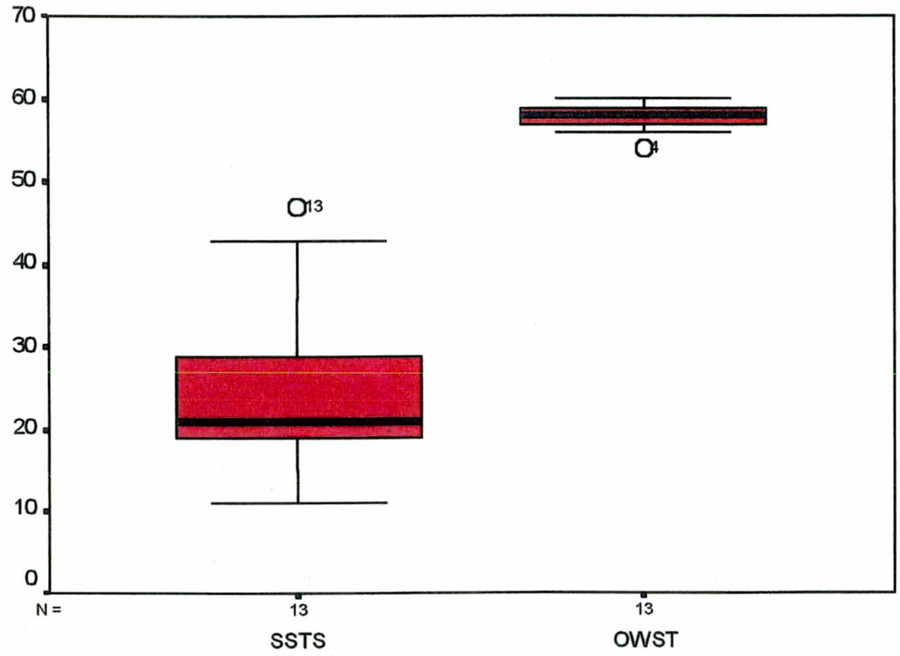
As can be seen in Figure 4.1, the median is close to the lower quartile, indicating that there is lower variability among those participants below the median and higher variability among those above the median. In addition, the upper whisker is longer than the lower one, indicating that the scores are positively--but weakly--skewed and that there are participants performing in the high end of the distribution. Finally, the box plot indicates the presence of one outlier (participant 13), who is outside the upper whisker. On the whole, however, this initial inspection shows that the data for the speaking span test are reasonably well distributed and thus adequate for undergoing correlational analyses. Figure 4.2 shows the box plot for the operation-word span test:

Figure 4.2 Box plot for the operation-word span test



As can be seen from Figure 4.2, the box plot for the operation-word span test shows that, although the data for this test are symmetrically distributed, there is an outlier performing in the lower end of the distribution. More importantly, the box plot indicates that there is almost no variability in the participants' scores on the test. This, added to the fact that the scores all hover close to 60--the maximum possible score--implies a potential ceiling effect. Analysis of the two box plots in the same numerical scale, as shown in Figure 4.3, makes it clear that there is much more variability of scores in the speaking span test (SST). Another point to notice is that the entire range of scores for the operation-word span test (OWST) is above even the outlier for the SST, which makes the ceiling effect more apparent.

Figure 4.3 Box plots for the SST and the OWST



The following chapter presents the results of the data analysis. For all analyses, a probability level of $p < .05$ was used to determine statistical significance. The computation of Cronbach's alpha (see section 4.6) was performed using Excel 8.0. All other statistical analyses were performed using the SPSS 8.0 Student Version package.

CHAPTER 5

RESULTS

This chapter presents the results of the statistical analyses performed to address the hypotheses and research questions of this study. As stated in the Introduction and Chapter III (Method), the aim of the present study was to examine the relationship between working memory capacity and fluency, accuracy, complexity and lexical density in L2 speech production.

This Chapter is divided into 4 main sections. Section 5.1 presents the descriptive statistical analysis performed for the working memory span tests. Section 5.2 presents the descriptive statistical analysis performed for the measures of L2 speech production in the picture description task and the narrative task. Section 5.3 presents the results of the correlational analyses: The Pearson Product Moment Coefficient of Correlation, with and without outliers, and Spearman's Rank Order Correlation. Finally, Section 5.4 reports the results from simple linear regressions performed in order to determine whether working memory capacity, as measured by the speaking span test, is a significant predictor of L2 speech production measures.

5.1 Working memory measures: The speaking span test and operation-word span test

This section presents the descriptive statistics of the speaking span test and the operation word span test. As discussed in Chapters 2 and 3, the speaking span test reflects the task-specific view of working memory capacity, whereas the operation-word span test was designed to reflect the claim that working memory has a domain-free capacity independent of the task being performed. Table 5.1 reports the descriptive statistics--the mean (M), standard deviation (SD), and the minimum and maximum scores for the speaking span test and the operation span test:

Table 5.1

Descriptive Statistics for the Speaking Span Test (SST) and Operation Word Span Test (OWST)

	SST	OWST
Mean	24.5	58.07
SD	10.82	1.75
Minimum	11	54
Maximum	47	60

N=13

For both span measures, the highest possible score was 60. However, as can be seen from Table 5.1, while the speaking span test scores varied over a 36-point range with a large standard deviation, the scores on the operation-word span test covered a mere 6-point range, clustering around the mean and close to the maximum or ceiling.

As Goodwin (1995) points out, "when doing a correlational study, it is important to include individuals with a wide range of scores" (p. 243). While this condition was met for the speaking span test, it was clearly not met for the operation-word span test, for which a strong ceiling effect is apparent. This ceiling effect, thus, had disruptive effects on the validity of the operation-word span test as an adequate variable on which to perform inferential statistical analyses. The lack of variability left almost no possibility for significant correlations with any of the other variables, as can be seen in Tables 5.2 and 5.3:

Table 5.2
Correlations between the Speaking Span Test (SST) and the Operation Word Span Test (OWST)

	SST	OWST
SST	1.00	.23
OWST	.23	1.00

Table 5.3

Correlations between the Operation-Word Span Test (OWST) and Fluency, Accuracy, Complexity, and Weighted Lexical Density in the Picture Description and Narrative Tasks

	Fluency Variables					Other		
	SRU	SRP	MLR	SPpm	Hpm	Acc.	Comp.	WLD
Description	.12	.15	.05	-.24	.02	-.07	.09	-.19
Narrative	.09	.08	.02	-.45	.21	-.29	.25	-.18

Note. SRU = speech rate unpruned
SRP = speech rate pruned
MLR = mean length of run
SPpm = silent pauses per minute
Hpm = hesitations per minute
Acc. = accuracy
Comp. = complexity
WLD = weighted lexical density

Table 5.2 shows no significant correlations between the two span tests. Table 5.3, in turn, shows no significant correlations between the operation-word span test and any of the measures of L2 speech production in either of the two tasks[1].

These results appear to indicate a lack of support for Hypothesis 5, which predicted that there would be a relationship between the operation-word span test and measures of fluency, accuracy, complexity, and weighted lexical density in the directions predicted in Hypotheses 1-4, both in the picture description task and in the narrative task.

[1] In the present study, one-tailed correlations were used to determine the correlation coefficients between memory spans and measures of speech production because the direction of the correlations was specified in Hypotheses 1-5.

However, since the operation-word span test has been shown to be an inadequate variable for inferential statistical analyses, these results can not be taken to indicate anything at all about Hypothesis 5. Thus, the second research question addressed in the present study--whether working memory capacity is specific to the task being performed or is a general task-independent capacity--cannot be answered from these data.

One possible explanation for the ceiling and the lack of variability in the operation-word span test is the fact that, being participant-paced, this test allowed extensive rehearsal of to-be remembered items and was, as a result, performed as a traditional word span test in which only the participants' storage capacity was actually assessed. Traditional word or digit span tests are considerably easier than working memory span tests and do not seem to exceed capacity, a prerequisite for individual differences to emerge (e.g., Just & Carpenter, 1992). The performance of the operation-word span test as a short-term memory span test and its implications for the second research question and Hypothesis 5 will be discussed more fully in Chapter 6, section 6.7. In the next section, the descriptive statistics for the measures of L2 speech production in the picture description and narrative tasks are presented.

5.2 Measures of L2 speech production: Fluency, accuracy, complexity, and weighted lexical density

In the present study, fluency in L2 speech production reflects continuous performance in real time (Foster & Skehan, 1996; Mehnert, 1998) and was assessed by 5 variables: speech rate unpruned (SRU), speech rate pruned (SRP), mean length of run (MLR), number of silent pauses per minute (SPpm), and number of hesitations per minute (Hpm). Accuracy (Acc.) was measured in terms of the number of errors in syntax, morphology, and lexical choice--including repetitions--per 100 words. Complexity (Comp.) was measured in terms of number of dependent clauses per minute. Finally, weighted lexical density (WLD) indicates the percentage of weighted lexical items over all linguistic items in each participant's speech sample. Table 5.4 shows the means (M) and standard deviations (SD) as well as the minimum (Min) and maximum (Max) values for each L2 speech production variable in the picture description and narrative tasks:

Table 5.4
Descriptive Statistics for L2 Speech Production Scores

		Fluency					Other		
		SRU	SRP	MLR	SPpm	Hpm	Acc.	Comp.	WLD
Descrip.	M	82.76	79.09	3.23	18.05	18.24	9.65	3.81	51.60
	SD	19.06	18.60	.69	4.51	8.42	5.75	2.61	4.88
	Min	54.91	52.46	2.45	7.99	4.15	1.55	.59	44.66
	Max	117.32	111.42	5.09	22.12	30.49	19.44	8.57	61..25
Narrat.	M	95.09	93.81	3.52	15.59	21.79	6.62	5.72	51.19
	SD	22.71	25.79	.80	4.08	7.60	3.84	2.51	5.58
	Min	42.99	38.49	1.70	6.99	8.25	1.68	1.76	42.03
	Max	122.49	144.99	4.76	22.05	36.99	15.11	9.02	63.29

As can be seen from Table 5.4, the means of both measures of speech rate were higher for the narrative task. The narrative task yielded a lower mean number of silent pauses per minute--those of .5 seconds or longer--and a higher number of hesitations per minute than the picture description task. This counter-balance might be the explanation for the similar mean length of run in the two tasks. Table 5.4 also shows that the narrative task allowed participants to produce more accurate and more complex speech in comparison to the picture description task. Finally, there was little difference in the means of weighted lexical density between the description and narrative tasks, which might indicate that, in both tasks, participants tended to draw on approximately the same number of different lexical items.

Tables 5.5 and 5.6 display participants' scores on all variables of L2 speech production, in the picture description and narrative tasks. Participants are organized from lower-to higher-working memory spans (see Appendixes A and B for individual scores on the span tests):

Table 5.5
Participants' Scores on Fluency Variables in the Picture Description Task and the Narrative Task

	Picture Description Task					Narrative Task				
Part.	SRU	SRP	MLR	SPpm	Hpm	SRU	SRP	MLR	SPpm	Hpm
1	54.91	52.46	2.45	20.84	16.27	42.99	38.49	1.70	19.50	15.99
2	57.03	53.46	2.59	21..38	14.25	86.04	82.76	3.18	12.00	19.66
3	78.78	74.24	2.71	21. 90	20.69	76.76	75.88	2.93	22.05	17.64
4	74.50	66.00	3.04	19.99	24.00	95.10	87.58	3.20	18.62	26.20
5	82.39	81.62	3.27	22.12	12.96	86.78	85.44	2.88	20.02	16.02
6	56.10	55.06	2.51	21.81	4.15	69.96	70.45	3.42	16.51	8.25
7	86.11	83.44	3.58	20.69	10.01	109.62	107.78	4.76	16.58	14.73
8	103.31	99.05	3.35	14.37	26.62	112.99	109.99	4.18	12.99	21.99
9	92.00	90.00	3.60	7.99	30.49	105.49	102.99	3.98	6.99	30.49
10	102.61	99.13	3.37	12.17	28.69	122.49	144.99	3.60	11.49	36.99
11	83.75	78.21	3.59	17.74	6.10	94.50	90.00	3.31	14.49	27.99
12	87.07	84.17	2.88	19.73	22.05	112.99	108.49	4.18	15.00	22.99
13	117.32	111.42	5.09	13.92	20.89	120.53	114.69	4.45	16.46	24.42

Note. Part. = participant
SRP = speech rate pruned
SRU = speech rate unpruned
MLR = mean length of run.
$N = 13$
SPpm = number of silent pauses per minute
Hpm = number of hesitations per minute

Table 5.6
Participants' (Part.) Scores on Accuracy (Acc.),Complexity (Comp.), and Weighted Lexical Density (WLD) in the Picture Description and Narrative Tasks

	Picture Description Task			Narrative Task		
Part.	Acc.	Comp.	WLD	Acc.	Comp.	WLD
1	19.44	2.03	58.91	15.11	3.00	52.83
2	18.75	.59	54.61	6.66	4.09	63.29
3	13.07	2.43	61.25	8.13	1.76	57.26
4	17.44	1.50	55.49	10.86	2.06	52.07
5	4.62	1.52	44.66	8.46	6.00	49.32
6	8.33	1.55	51.35	10.00	7.15	51.51
7	1.55	6.67	49.01	1.68	7.36	44.17
8	9.27	3.72	48.57	3.98	7.50	50.18
9	7.60	3.49	53.42	3.31	4.99	54.19
10	4.23	7.82	48.99	3.67	7.99	42.03
11	6.62	3.88	48.14	3.17	9.00	46.78
12	7.33	5.80	50.00	7.07	4.50	54.20
13	7.30	8.57	46.16	3.96	9.02	47.70

N=13

Participants' individual scores, presented in Tables 5.5, demonstrate that the range of variation in mean length of run in the picture description task is barely more than 1 semantic unit--from 2.45 to 3.60, except for participant 13. In the narrative task, the range of scores on mean length of run is much wider, varying from 1.70 to 4.76, a range of 3 semantic units, and participant 13 is within the normal range.

Table 5.5 also indicates that there appears to be an interaction between number of silent pauses and hesitations in both tasks. For instance, in the picture description task the first three participants display a higher number of silent pauses and a lower number of hesitations per minute, whereas the last four participants exhibit the opposite pattern: fewer silent pauses but a higher number of hesitations. In the narrative task, there seems to be more instability in the scores of the first participants, but the pattern for the last 5 participants is the same: the number of hesitations is always higher than the number of long silent pauses. In other words, it seems that those individuals with a higher speech rate tend to use fewer long silent pauses and rely more extensively on devices such as repetitions, filled nonlexical pauses, and shorter silent pauses than those with a slower speech rate.

The individual scores on accuracy and complexity, presented in Table 5.6, also indicate an apparent tendency for accuracy and complexity to increase (that is, for a lower number of errors and a higher number of dependent clauses to be produced) as speech rate increases in both tasks. Finally, it is interesting to notice that, for both tasks, the percentage of lexical items over the total number of items tends to decrease as speech rate, accuracy, and complexity increase.

Taken together, the results from the descriptive statistical analyses performed for the measures of L2 speech production show that the participants maintained roughly the same pattern in both tasks, exhibiting slightly better performance in the narrative task, since this elicited a faster speech rate, fewer silent pauses, fewer errors in syntax, morphology, and lexical choice, and a higher number of dependent clauses. The participants also maintained approximately the same mean length of run in the two tasks, but produced more hesitations in the narrative task. In general, a smaller number of long silent pauses was counter-balanced by a higher number of hesitations per minute, in both tasks. Furthermore, the percentage of weighted lexical items over the total number of linguistic items produced was approximately the same both in the picture description task and narrative task. Finally, an increase in speech rate seems to be accompanied by increases in accuracy and complexity, and by a decrease in weighted lexical density. The next section presents the results from the inferential statistical analyses.

5.3 Inferential statistics: Correlations and simple linear regressions

This section presents the results of the inferential statistical analyses--the Pearson Product Moment Coefficient of Correlation (one-tailed) and simple linear regressions--on the speaking span test and L2 speech production measures. It is further subdivided into 4 sections, each dealing with one of the four Hypotheses.

A preliminary inspection of the joint distribution of scores on the speaking span test and each L2 speech production measure (see Appendix L for scattergrams) indicated that some of the correlations between the speaking span test and L2 speech production measures were affected by the behavior of participants 12 and 13, who scored high in the speaking span test but tended to score low in some of the measures of speech production. Participant 13 had already been spotted as an outlier in the speaking span test. Participant 12 emerges as an outlier when participants' scores are plotted according to their joint position on both the speaking span test and L2 speech production measures.

There seems to be no consensus on the best way to deal with outliers in a given sample. Walsh (1990), for instance, suggests that, once identified, problems with outliers are best dealt with by eliminating them from the data set. McClave et al. (1995), on the other hand, maintain that outliers frequently reveal useful information on the variables and play an important role in the data analytical process. Hatch & Farhady (1982) state, in turn, that outliers should be removed from the analyses for the true relationship to be seen, but that their behavior should be further dealt with through case studies. Since the sample size of the present study is extremely small and it is difficult, in general, to justify eliminating extreme observations from the data, it was decided that the outliers would be kept, and that, as noted in Chapter 4, complementary statistical computations would be performed for all pairs of variables. These consisted of (a) the Pearson Product Moment

Coefficient of Correlation *without* the outliers' scores and (b) Spearman's Rank Order Correlation with *all* participants' scores, both correlations being one-tailed.

The rationale for computing these additional correlational tests is that, given the small sample size, it is not possible to determine whether participants 12 and 13 are really atypical observations or whether they emerge as extreme values as a result of the restricted range of scores. As will be seen shortly, Pearson Product Moment Correlation without the outliers and Spearman's Rank Order Correlation with all participants' scores strengthens the degree of association between the speaking span test and L2 speech production. Thus, the most cautious approach, in the case of the present study, is to consider the Pearson Product Moment Correlation with *all* participants' scores as the basis for the statistical association between the speaking span test and measures of L2 speech production and further inspect this association through Pearson Product Moment Correlation without the outliers' scores as well as through Spearman's Rank Order Correlation with all participants' scores included.

5.3.1 Working memory capacity, as measured by the speaking span test, and measures of L2 fluency

Hypothesis 1 predicted that there would be a relationship between working memory capacity, as measured by the speaking span test, and fluency in L2 speech production, as measured by speech rate unpruned, speech rate pruned, mean length of run, number of silent pauses per minute, and number of hesitations per minute, both in the picture description and narrative tasks. Working memory was predicted to correlate positively with speech rate unpruned, speech rate pruned, and mean length of run, and negatively with number of silent pauses per minute and number of hesitations per minute.

The results related to this hypothesis are presented in two sections. Section 5.3.1.1 presents the results from the three correlational performed for the relationship between the speaking span test and speech rate unpruned, speech rate pruned, and mean length of run. Section 5.3.1.2 presents the results from the same three tests performed for the relationship between the speaking span test and number of silent pauses and hesitations per minute.

5.3.1.1 The relationship between the speaking span test and L2 speech rate unpruned, speech rate pruned, and mean length of run

Table 5.7 reports the correlation coefficients calculated between the speaking span test and speech rate unpruned, speech rate pruned, and mean length of run, in the picture description and narrative tasks. Recall that Pearson Product Moment Correlation with all participants' scores included assumes that participants 12 and 13 are not atypical observations, that Pearson Product Moment Correlation without these participants' scores included assumes they are atypical observations obscuring the relationship, and that Spearman's Rank Order Correlation checks for a linear association between two variables when the scores are ranked, thus being less sensitive to extreme values since variation in scores is not taken into account.

Table 5.7
Pearson Product Moment Correlations and Spearman's Rank Order Correlation Between the Speaking Span Test (SST) and Speech Rate Unpruned (SRU), Speech Rate Pruned (SRP), and Mean Length of Run (MLR), in the Picture Description and Narrative Tasks:

	Picture Description Task			Narrative Task		
	SRU	SRP	MLR	SRU	SRP	MLR
SST[a]	.73**	.72**	.70**	.69**	.68**	.62*
SST[b]	.73**	.72**	.80**	.69**	.68*	.51
SST[c]	.83**	.83**	.75**	.79**	.82**	.74**

[a] Pearson Product Moment Correlation (*r*), N=13
[b] Pearson Product Moment Correlation (*r*), N = 11
[c] Spearman's Rank Order Correlation (r_s), N = 13
* $p < 0.05$
** $p < 0.01$

As can be seen in Table 5.7, results from The Pearson Product Moment Coefficient of Correlation with **all** participants' scores included show that there is a statistically significant correlation between working memory capacity, as measured by the speaking span test (SST), and speech rate unpruned, r (13) = .73, $p < 0.01$, speech rate pruned, r (13) = .72, $p < 0.01$, and mean length of run, r (13) = .70, $p < 0.01$, in the picture description task. Similarly, there is a statistically significant correlation between the speaking span test and speech rate unpruned, r (13) = .69, $p < 0.01$, speech rate pruned, r (13) = .68, $p < 0.01$, and mean length of run, r (13) = .62, $p < 0.05$, in the narrative task. These significant correlations suggest that working memory capacity, as measured by means of the speaking span test in the participants' L2, may be related to continuous performance in real time in the L2 oral production of a picture description and a narrative, as measured by rate of speech and length of runs between pauses and hesitations. These results, thus, lend substantial support to Hypothesis 1.

Results from the Pearson Product Moment Coefficient of Correlation **without** participants 12 and 13 tend to reproduce those of the Pearson correlation with all participants' scores included. The outliers do not seem to affect the association between the L2 speaking span test and the two versions of speech rate in the picture description task: r (11) = .73, $p < 0.01$, for speech rate unpruned, and r (11) = .72, $p < 0.01$, for speech rate pruned. However, without the outliers, the association between the speaking span test and mean length of run in this task increases: r (11) = .80, $p < 0.01$.

In the narrative task, the association between the speaking span test and speech rate unpruned does not change without the outliers: r (11) = .69, $p < 0.01$, but statistical level of significance of the association with speech rate pruned decreases without them: r (11) = .69, $p < 0.05$.

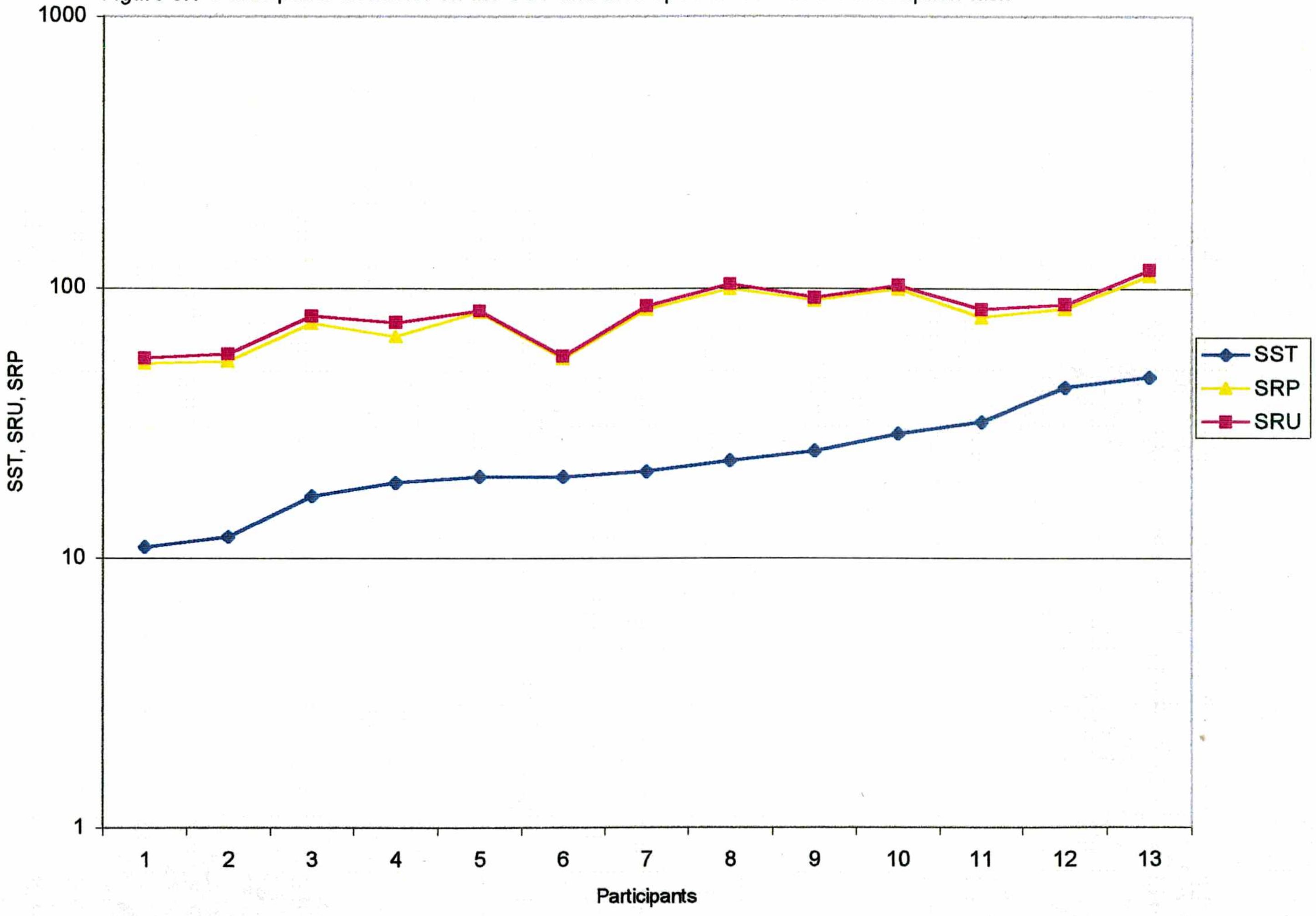
Finally, the association between the speaking span test and mean length of run in the narrative task is severely affected without the outliers, since it loses significance when they are left out of computations: r (11) = .51, p < .052. In this case, participant 13 seems to be the one affecting the relationship and when only this participant's results are left out of the computation, the association between the L2 speaking span test and mean length of run reaches statistical significance, as the following results show: r (12) = .55, $p < 0.05$.

Together, the results from the Pearson Product Moment Coefficient of Correlation without the outliers might be an indication that working memory capacity, as measured by the speaking span test in the participants' L2, is related to continuous performance in real time in the L2 oral production of a picture description and a narrative, as measured by rate of speech and length of runs between pauses and hesitations.

The results from the third statistical technique, Spearman's Rank Order Correlation, with **all** participants' scores included, lend further support to Hypothesis 1. As can be seen in Table 5.7, there is a statistically significant correlation between the speaking span test and the two versions of speech rate, r_s (13) = .83, $p < 0.01$, for both, and between the speaking span test and mean length of run, r_s (13) = .75, $p < 0.01$, in the picture description task. Similarly, in the narrative task, there is a statistically significant correlation between the speaking span test and speech rate unpruned, r_s (13) = .79, $p <$ 0.01, speech rate pruned, r_s (13) = .82, $p < 0.01$, and mean length of run, r_s (13) = .74, $p <$ 0.01. These coefficients of correlation also show that, when participants' scores are rank-ordered, the association between working memory capacity, as measured by the L2 speaking span test, and L2 continuous oral performance in real time, as measured by speech rate and mean length of run, appears stronger than when tested by the Pearson Correlation.

Thus, the results from The Pearson Product Moment Coefficient of Correlation and Spearman's Rank Order Correlation--both with all participants' scores included--provide support for Hypothesis 1, which predicted that the speaking span test would correlate positively with speech rate unpruned, speech rate pruned, and mean length of run. Assuming that participants 12 and 13 are, in fact, atypical observations, the results from the Pearson correlations without their scores also support Hypothesis 1, except that the relationship between the speaking span test and mean length of run in the narrative task loses statistical significance. Taken **together**, these results might indicate that working memory capacity, when measured in the participants' L2 during speech production, is related to L2 oral fluency, more specifically to the speed with which participants can speak the L2 and to the number of semantic units they can produce between silent pauses of various lengths and filled nonlexical pauses. In other words, individuals with a higher working memory capacity tend to speak faster and produce longer stretches of talk between pauses and hesitations. It can be more easily perceived in Figures 5.1 and 5.2 that in both the description and narrative tasks, as participants' scores in the speaking span test increase their speech rates--both unpruned and pruned--increase as well. The same trend is observed for the mean length of run (MLR), as shown in Figures 5.3 and 5.4.[2]

Figure 5.1 Participants' behavior on the SST and their speech rates in the description task



[2] Logarithmic graphs were thought to be the best solution to deal with the different numerical scales of the variables.

Figure 5.2 Participants' behavior on the SST and their speech rates in the narrative task

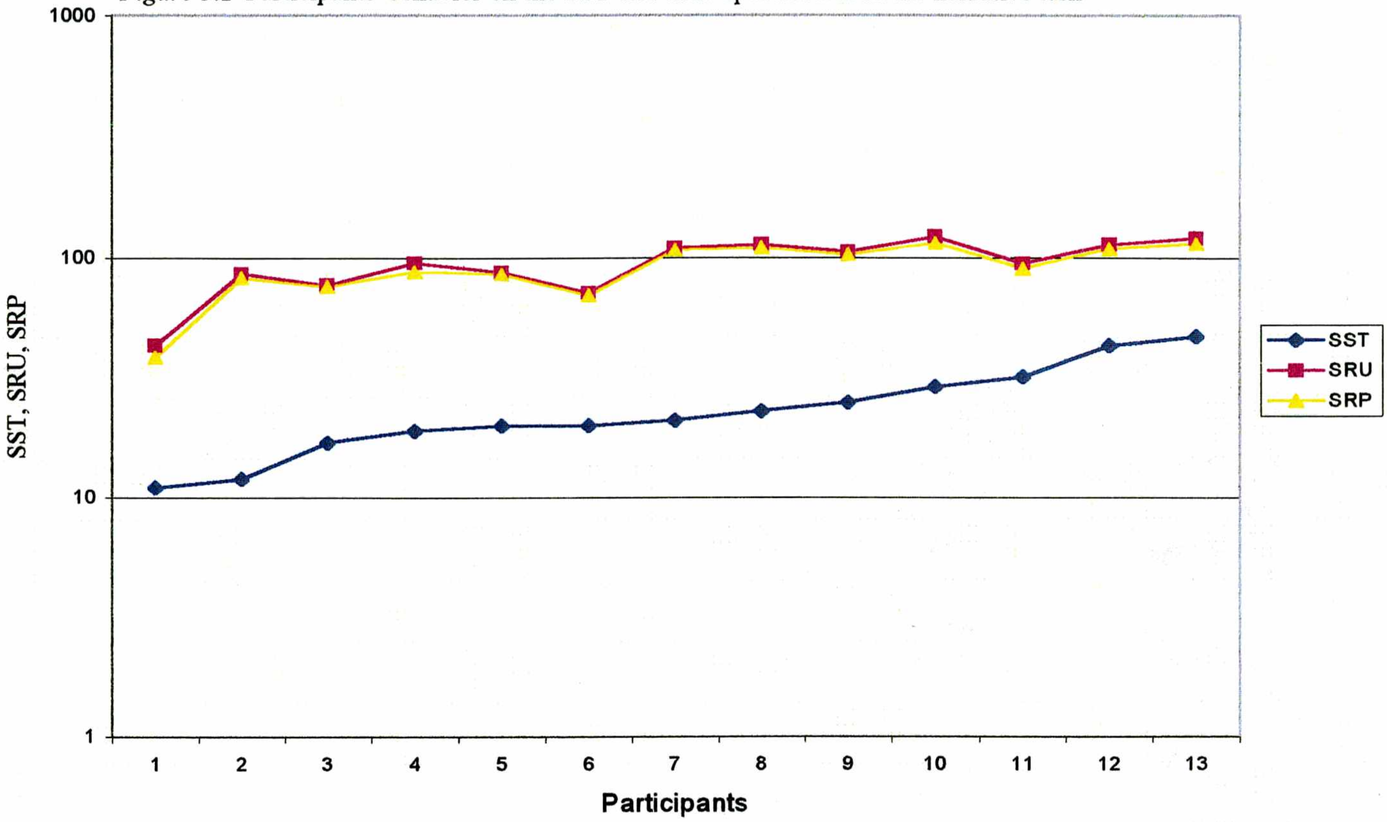


Figure 5.3 Participants' behavior on the speaking span test and their mean length of run in the picture description task

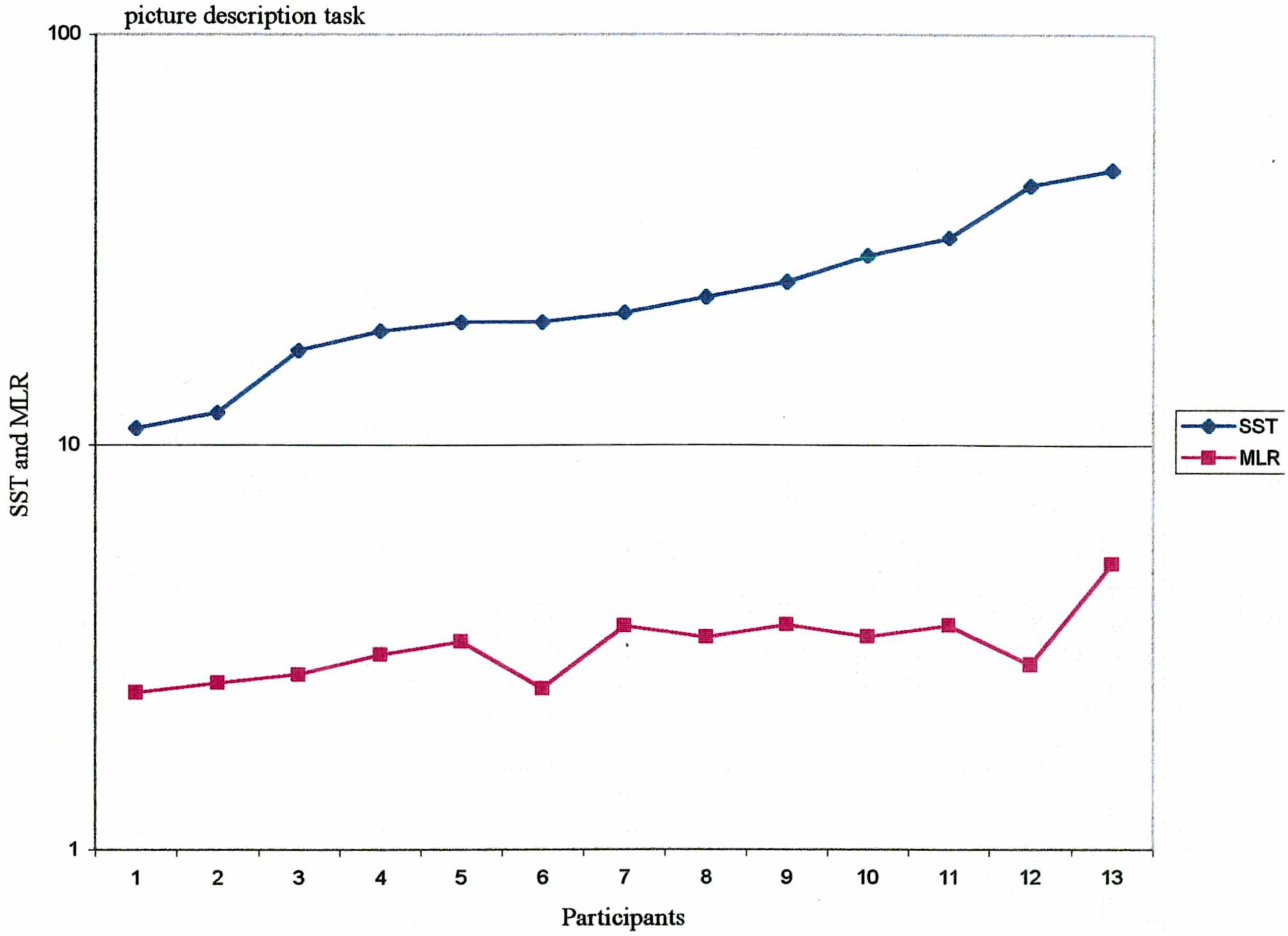
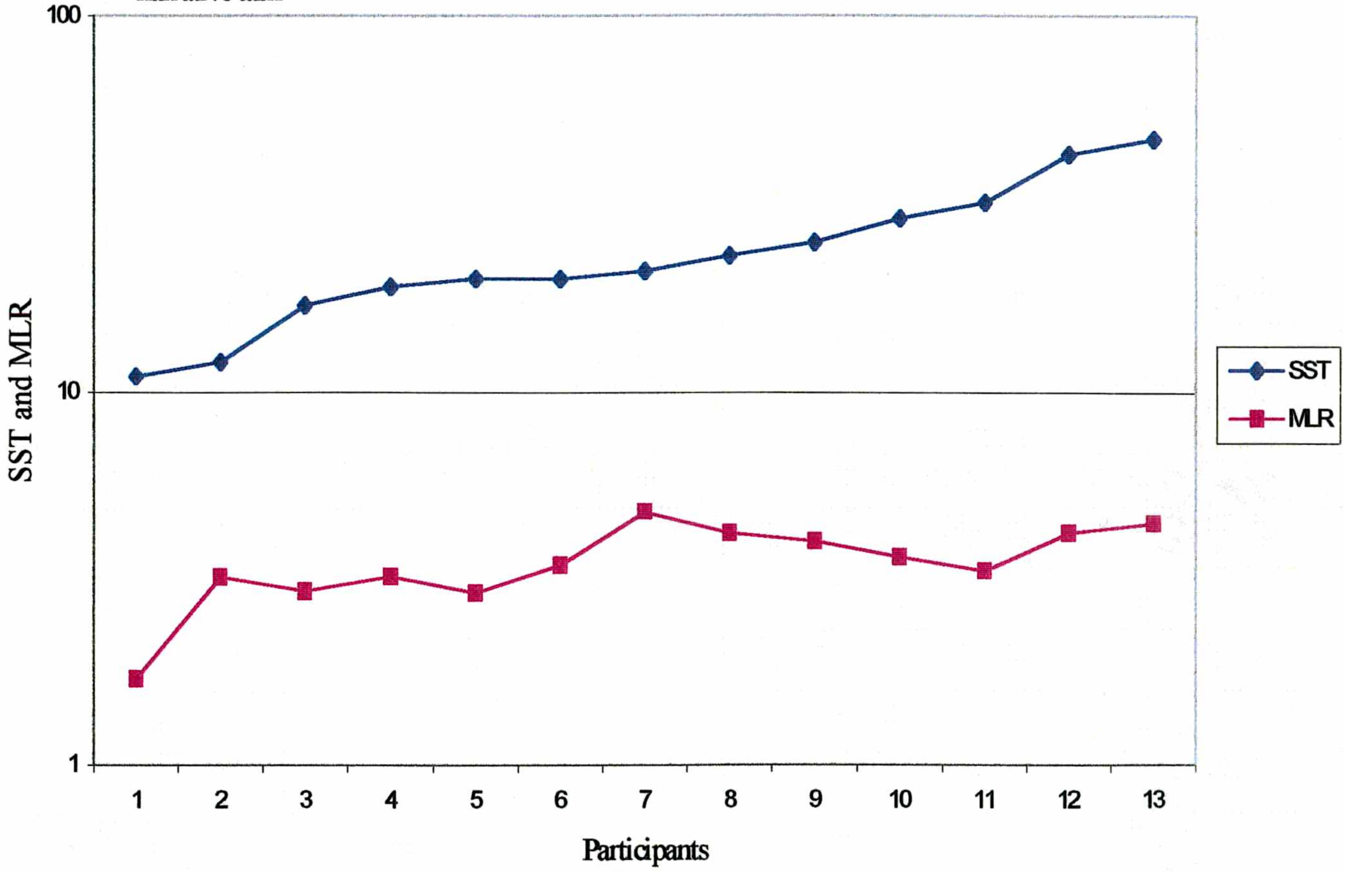


Figure 5.4 Participants' behavior on the speaking span test and their mean length of run in the narrative task



As participants' scores on the speaking span test increase, their mean length of run tends to increase as well. This increase in mean length of run is steadier in the picture description task than in the narrative task, although the range in the former is smaller. Hypothesis 1 also predicted that working memory capacity, as measured by the speaking span test, would correlate positively with number of silent pauses per minute and number of hesitations per minute. The results are presented in the next section.

5.3.1.2 The relationship between the speaking span test and number of silent pauses and hesitations per minute

Table 5.8 reports the results from the Pearson Product Moment Coefficient of Correlation--with and without the scores of participants 12 and 13 included--and from Spearman's' Rank Order Correlation with all participants' scores:

Table 5.8
Pearson Product Moment Correlations and Spearman's Rank Order Correlation Between the Speaking Span Test (SST) and Number of Silent Pauses per Minute (SPpm) and Number of Hesitations per Minute (Hpm) in the Picture Description and Narrative Tasks:

	Picture Description		Narrative	
	SPpm	Hpm	SPpm	Hpm
SST[a]	-.43	.20	-.22	.42
SST[b]	-.60*	.11	-.45	.62*
SST[c]	-.71**	.26	-.49*	.56*

[a] Pearson Product Moment Correlation (r), N=13
[b] Pearson Product Moment Correlation (r), N = 11
[c] Spearman's Rank Order Correlation (r_s), N = 13
* $p < 0.05$ (one-tailed)
** $p < 0.01$ (one-tailed)

As can be seen in Table 5.8, with **all** participants' scores included, the results from the Pearson Product Moment Coefficient of Correlation between working memory capacity, as measured by the speaking span test, and number of silent pauses and hesitations per minute were not statistically significant in either task. Thus, the prediction made in Hypothesis 1 that the speaking span test would correlate negatively with number of silent pauses per minute and number of hesitations per minute was not statistically supported. The relationship between the speaking span test and number of silent pauses per minute, both in the picture description and narrative tasks, is negative, as predicted, r (13) = -.43 and -.22, respectively. That is, individuals with a higher working memory capacity seem to be less prone to producing long silent pauses when speaking in the L2.

However, the relationship between the speaking span test and number of hesitations per minute, in both tasks, is positive, contrary to what was predicted, r (13) = 20 in the description task, and r (13) = .42, in the narrative task. That is, individuals with a higher working memory capacity seem to be more prone to hesitating--to producing silent pauses of .4 seconds or less, filled nonlexical pauses, and immediate repetitions--when speaking the L2. Although not statistically significant, these results might be taken as an indication of two trends. First, they might indicate that, as predicted, as working memory capacity increases, the number of silent pauses of .5 seconds (or longer) tends to decrease during L2 speech production. Second, they might be an indication of a trade-off between silent pauses and hesitations during L2 speech production. In other words, for the participants of the present study, the production of a smaller number of silent pauses was achieved through the production of a higher number of hesitations.

The results from the Pearson Product Moment Coefficient of Correlation **without** the outliers give stronger support to the trend that individuals with a higher working memory capacity, as measured by the L2 speaking span test, are less prone to exhibiting long silent pauses when speaking the L2 at the cost of producing a higher number of hesitations. Without the scores of participants 12 and 13, the relationship between the speaking span test and number of silent pauses reaches significance in the picture description task and improves considerably in the narrative task. However, the association between the span test and hesitations remains positive in the two tasks and, although decreasing in the description task, it reaches statistical significance in the narrative task, which reinforces the idea that the association between the two variables is positive and that an interaction between long silent pauses and different types of hesitations takes place during L2 speech production.

Finally, the results from Spearman's Rank Order Correlation, with **all** participants' scores included, bring the strongest support for a negative relationship between working memory capacity and silent pauses and for a trade-off between silent pauses and hesitations. When participants' scores are rank-ordered, the negative association between the span test and silent pauses is statistically significant, in both the description task r_s (13) = -.71, $p < 0.01$, and the narrative task, r_s (13) = -.49, $p < 0.05$. Thus, as predicted, working memory capacity, as measured by the speaking span test, correlates negatively with number of silent pauses per minute. However, the degree of the positive association between the span test and hesitations per minute increases in the picture description task, r_s (13) = .26, and reaches statistical significance in the narrative task, r_s (13) = .56, $p < 0.05$. That is, participants with a higher working memory capacity are more prone to hesitating during L2 speech production.

Taken together, the results from the Pearson Product Moment Coefficient of Correlation, with and without the scores of participants 12 and 13, and from Spearman's Rank Order Correlation lend substantial support to Hypothesis 1, which predicted that working memory capacity, as measured by the speaking span test, would be related to fluency in L2 speech production. As noted earlier, the notion of fluency in the present research project, is taken to reflect continuous performance in real time. To this extent, five variables were used that seem to measure continuous oral performance: speech rate unpruned, in which all semantic units produced are included in the count, speech rate pruned, in which immediate repetitions and partial words are left out, mean length of run, which reflects the average number of semantic units produced between pauses and hesitations, silent pauses of .5 second or longer, and hesitations, an umbrella term including silent pauses of .49 seconds or shorter, filled nonlexical pauses, and immediate repetitions.

The results from 3 statistical techniques aimed at determining the degree of relationship between the speaking span test and the variables of L2 oral fluency show that individuals with a higher working memory capacity, as measured by the speaking span test, speak faster, have a greater mean length of run between pauses and hesitations, tend to produce fewer long silent pauses, but more hesitations when engaged in the oral production of a picture description and a narrative task in the L2.

It is for the prediction that there would be a negative relationship between working memory capacity and hesitations that the results from the statistical techniques employed in the present study lend no support. What these results seem to show is that there is an interaction between silent pauses and hesitations: fewer silent pauses are accompanied by more hesitations, as shown in Table 5.5. To examine this interaction further, statistical analyses were performed on the two variables. Table 5.9 reports the results from the Pearson Product Moment Coefficient of Correlation, with and without the outliers, and from Spearman's Rank Order Correlation, with all participants' scores included:

Table 5.9
Correlations Between Number of Silent Pauses per Minute (SPpm) and Number of Hesitations per Minute (Hpm) in the Picture Description and Narrative Tasks

	Picture Description Hpm	Narrative Hpm
$SPpm_a$	-.68**	-.58*
$SPpm_b$	-.70**	-.59*
$SPpm_c$	-.67**	-.63*

[a] Pearson Product Moment Correlation (r), N=13
[b] Pearson Product Moment Correlation (r), N = 11
[c] Spearman's Rank Order Correlation (r_s), N = 13
* $p < 0.05$ (one-tailed)
** $p < 0.01$ (one-tailed)

As can be seen in Table 5.9, results from the three statistical tests show a significant negative relationship, of almost the same degree, between number of silent pauses per minute and hesitations per minute, in the two tasks. These results indicate that a lower number of silent pauses of .5 seconds or longer is related to a higher number of hesitations during the production of the description and narrative tasks used in the present study, and can be taken as evidence for a trade-off between the two variables.

The results obtained in the present study regarding silent pauses and hesitations are consistent with the literature both on speech production and working memory. As will be discussed later, the relationship among temporal variables is not a straightforward one (Lennon, 1990), with intricate patterns emerging from their interaction. In the case of the current results, any gains in one type of temporal variable (for instance, fewer silent pauses) seem to take place through losses in another (for instance, a higher number of hesitations). From the perspective of working memory theories, this finding is consistent with the view that the working memory system is capable of differentially allocating resources to different aspects of the task when task demands exceed its capacity. As far as temporal variables go, then, it appears that in the present study, individuals with a larger working memory span, in order to speak faster, longer (between pauses and hesitations), and with fewer silent pauses, relied extensively on the use of a number of hesitation phenomena, in contrast to those with a smaller working memory span, who relied more heavily on long silent pauses. The next section deals with Hypotheses 2-4.

5.3.2 Working memory capacity, as measured by the speaking span test, and accuracy, complexity, and weighted lexical density

Hypothesis 2 predicted that there would be a negative correlation between working memory capacity, as measured by the speaking span test, and accuracy in L2 speech production, as measured by the number of errors in syntax, morphology, and lexical choice per hundred words. Hypothesis 3 predicted that there would be a positive correlation between the span test and complexity in L2 speech production, as measured by the number of dependent clauses per minute. Hypothesis 4 predicted that there would be a positive correlation between the span test and weighted lexical density in L2 speech production. Table 5.10 presents the results from The Pearson Product Moment Coefficient of Correlation--with and without the outliers--and from Spearman's Rank Order Correlation:

Table 5.10

Pearson Product Moment Correlations and Spearman's Rank Order Correlation Between the Speaking Span Test (SST) and Accuracy (Acc.), Complexity (Comp.) and Weighted Lexical Density (WLD) in the Picture Description and Narrative Tasks

	Picture Description Task			Narrative Task		
	Acc.	Comp.	WLD.	Acc.	Comp.	WLD
SST[a]	-.53*	.76**	-.57*	-.48*	.54*	-.39
SST[b]	-.72**	.63*	-.61*	-.70**	.73**	-.68**
SST[c]	-.67**	.80*	-.68**	-.63*	.74**	-.49*

[a] Pearson Product Moment Correlation (r), N=13 — * $p < 0.05$
[b] Pearson Product Moment Correlation (r), N=11 — ** $p < 0.01$
[c] Spearman's Rank Order Correlation (r_s), N=13
Note. SST: speaking span test

As can be seen in Table 5.10, Hypothesis 2 seems to be supported. Results from The Pearson Product Moment Coefficient of Correlation, with **all** participants' scores included, show that working memory capacity, as measured by the number of sentences produced in the speaking span test, correlates negatively with the number of errors in syntax, morphology, and lexical choice per hundred words produced in the picture description task, r (13) = -.53, $p < 0.05$, and narrative task, r (13) = -.48, $p < 0.05$.

Results from the Pearson Product Moment Coefficient of Correlation **without** participants 12 and 13 improve the degree of association between working memory capacity and accuracy in both tasks, r (11) = -.72, $p < 0.01$ for the picture description task, and r (11) = -.70, $p < 0.01$, for the narrative task.. The association between the speaking span test and accuracy is further demonstrated by the significant results, to a degree midway between those of the two Pearson tests, from Spearman's Rank Order Correlation, with **all** participants' scores included: r_s (13) = -.67, $p < 0.01$ for the description task, and r_s (13) = -.63, $p < 0.01$, for the narrative task. Taken together, these results seem to suggest that individuals with a larger working memory capacity were also less prone to making syntactic, morphological, and lexical errors when describing a picture and narrating in their L2.

Figures 5.5 and 5.6 display, visually, a trend for number of errors to decrease as span size increases:

Figure 5.5 Participants' scores on the speaking span test (SST) and accuracy (Acc.) in the picture description task

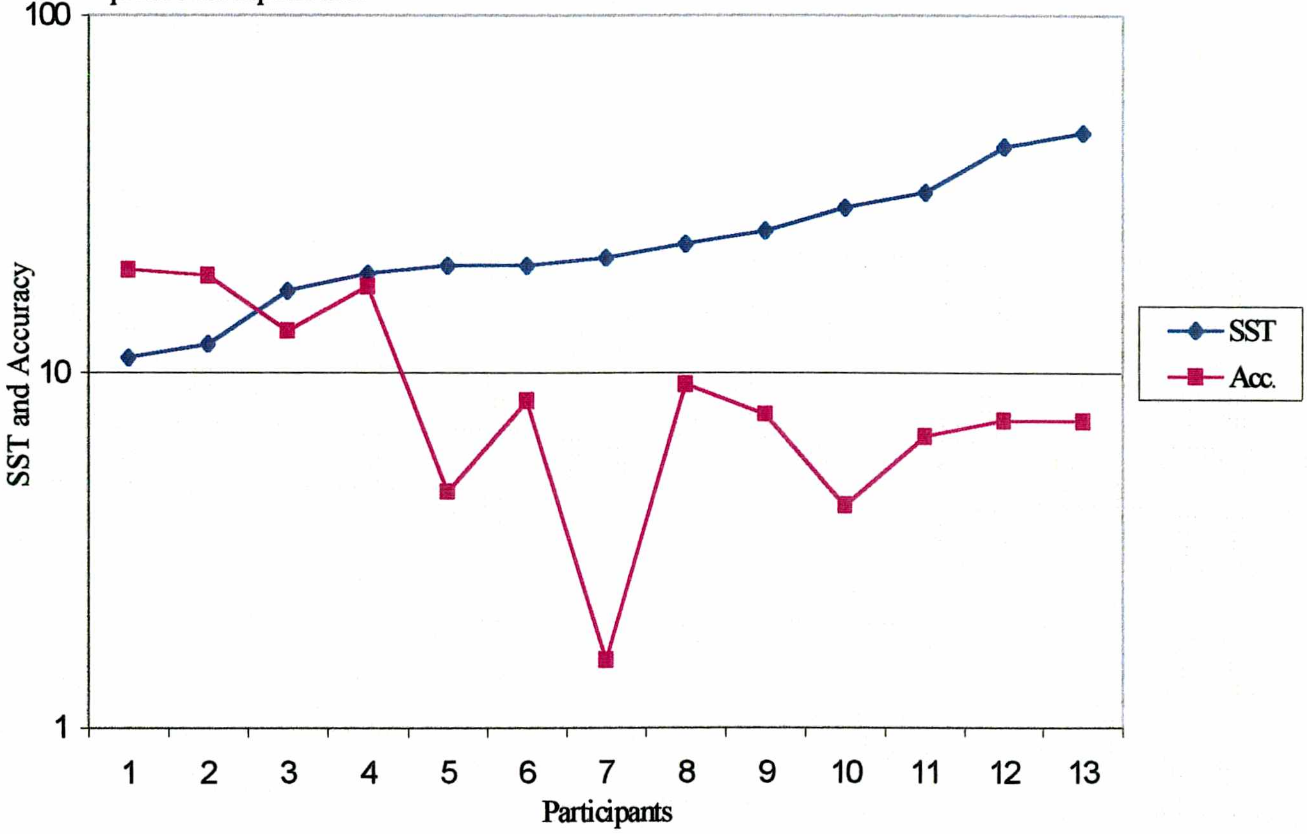
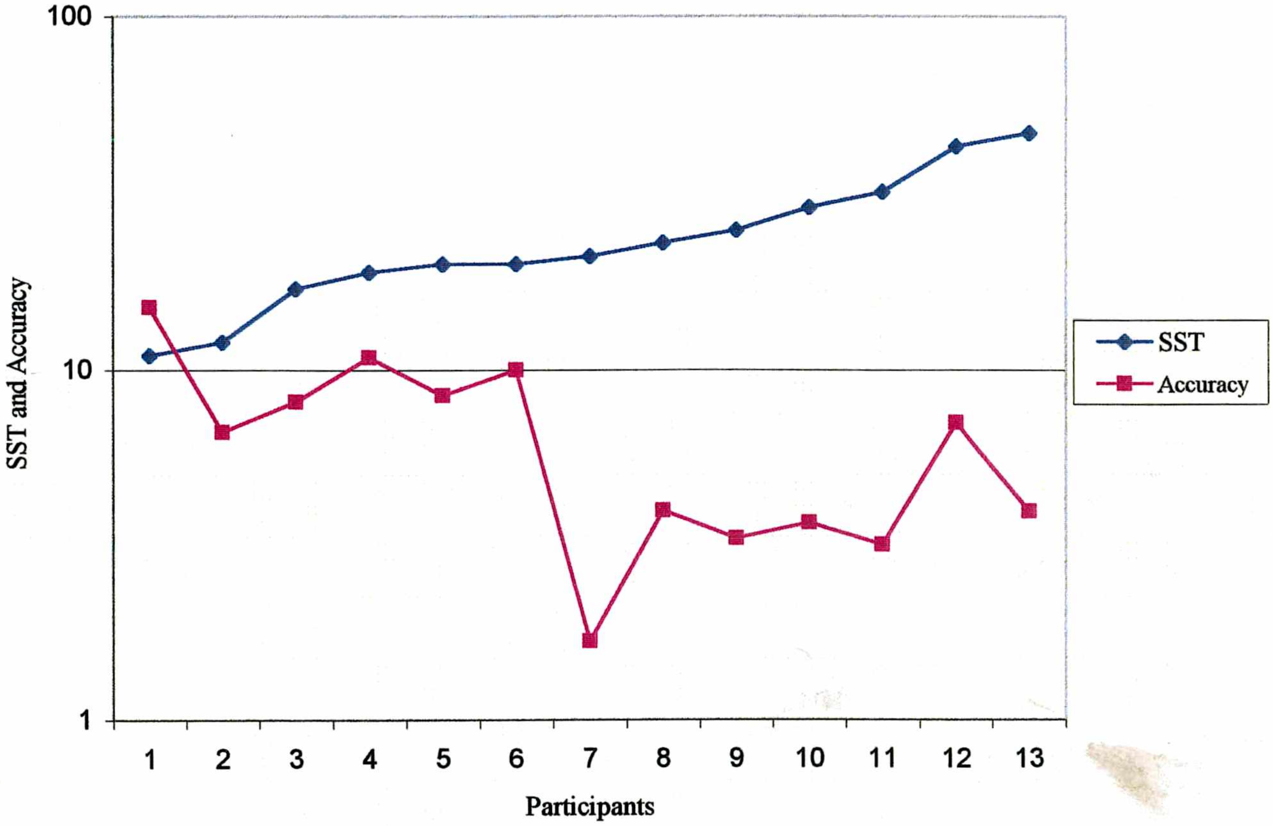


Figure 5.6 Participants' scores on the speaking span test (SST) and accuracy in the narrative task



The correlation coefficients reported in Table 5.10 indicate that there is a significant positive correlation between working memory capacity, as measured by the speaking span test, and complexity in L2 speech production, as measured by the number of dependent clauses per minute, both in the picture description and narrative tasks. Based on these results, Hypothesis 3 is supported. Results from the Pearson Product Moment Coefficient of Correlation with **all** participants' scores included are *r* (13) = .76,

$p < 0.01$, for the description task, and r (13) = .54, $p < 0.05$ for the narrative task, thus stronger for the former and weaker for the latter. Results from the Pearson correlations **without** participants 12 and 13 decreases the degree of association between the span test and complexity in the description task, r (11) = .63, $p < 0.05$, but increases their association in the narrative task, r (11) = .73, $p < 0.05$. Further statistical support for the positive association between working memory capacity and complexity in L2 speech production is provided by the results from Spearman's Rank Order Correlation, with **all** participants' scores included, which shows the strongest association in both tasks: r_s (13) = .80, p < 0.01, for the description task, and r_s (13) =.74, p < 0.01. Taken together, these results might suggest that individuals with a larger working memory capacity , as measured by the L2 speaking span test, are also more prone to producing L2 speech that is more grammatically complex, as measured by the number of dependent clauses per minute of talk. Figures 5.7 and 5.8 demonstrate graphically that, in both tasks, the number of dependent clauses seem to increase as working memory span increases.

Figure 5.7 **Participants' scores on the speaking span test (SST) and complexity in the picture description task**

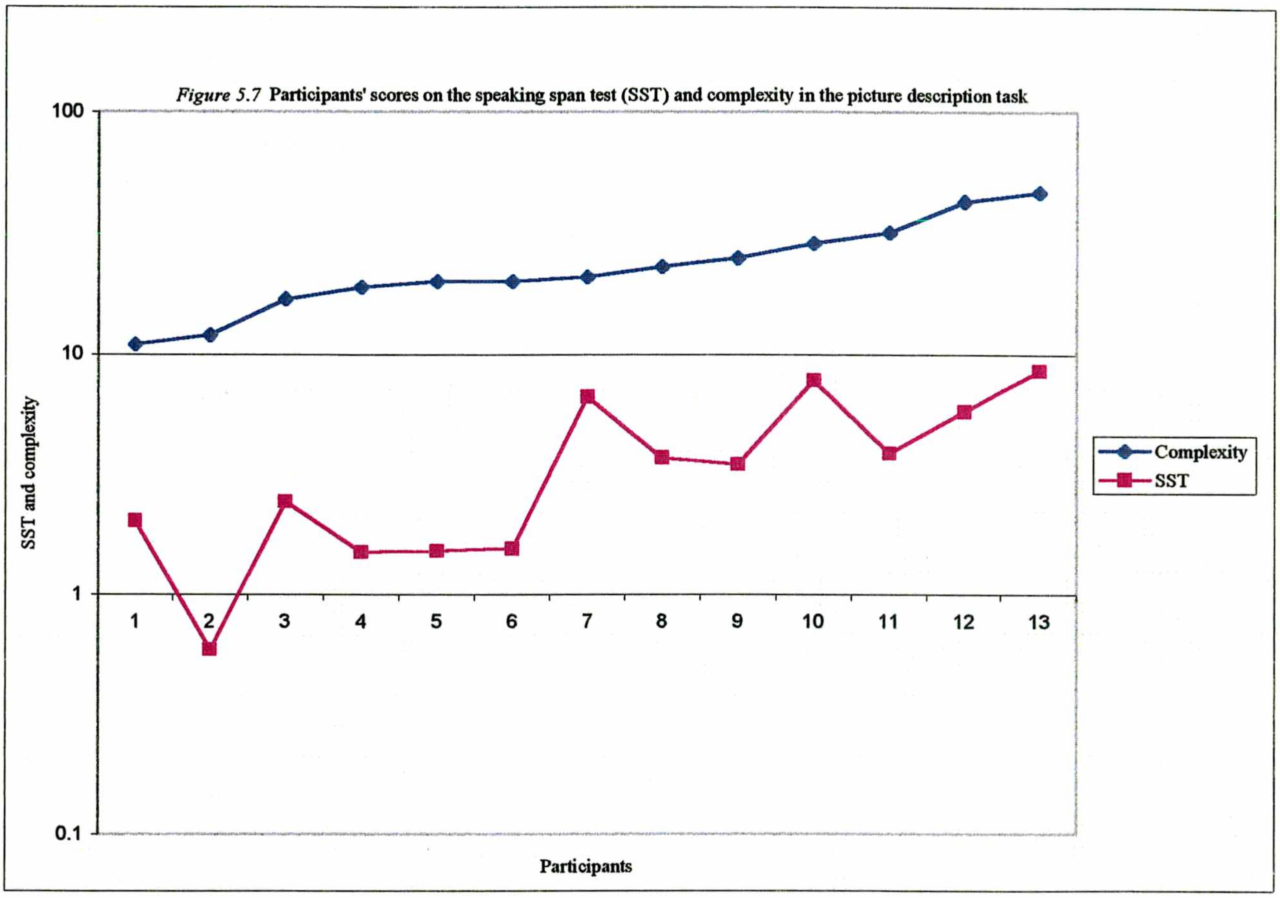
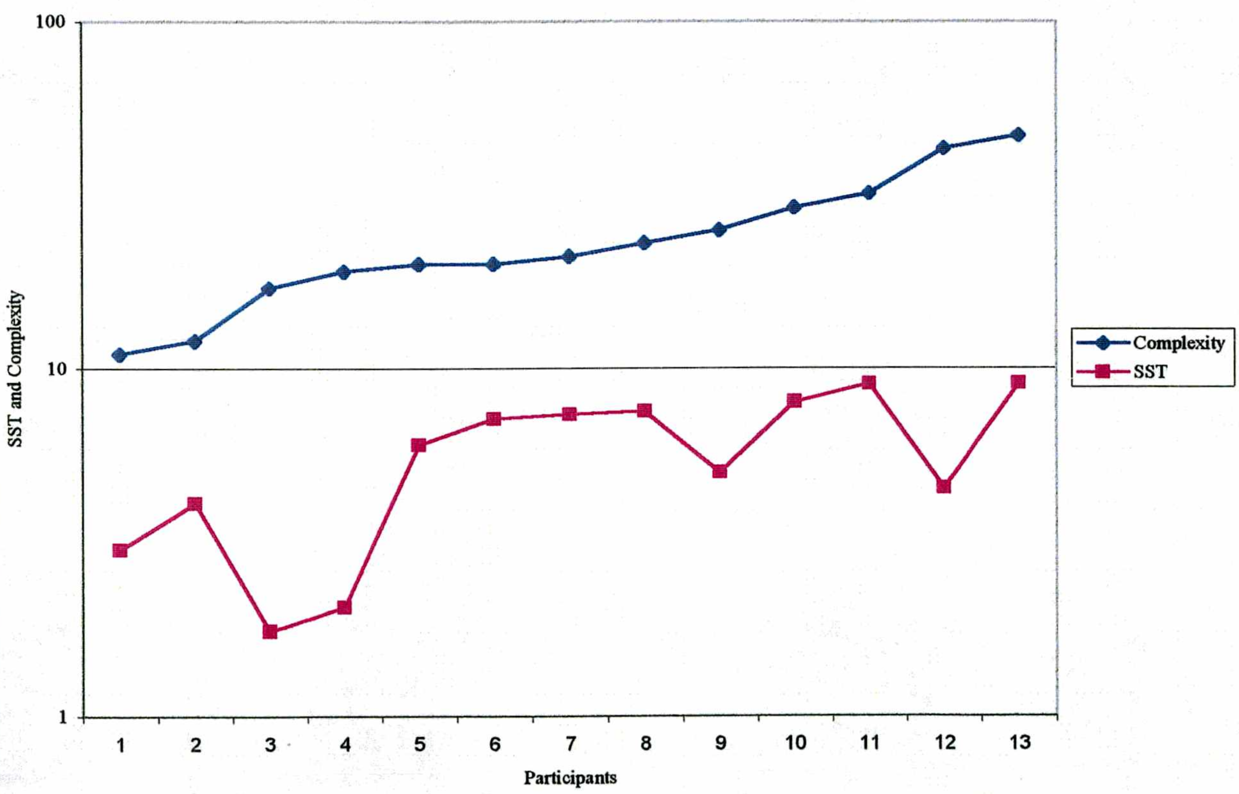


Figure 5.8 **Participants' scores on the speaking span test (SST) and complexity in the narrative task**



Finally, for Hypothesis 4, the results of all three tests, reported in Table 5.10 reveal an initially counter-intuitive finding: There is a significant correlation between working memory capacity, as measured by the speaking span test, and L2 weighted lexical density, as measured by the percentage of weighted (or low-frequency) lexical items over the total number of linguistic items, but in the opposite direction from that predicted. Thus, Hypothesis 4 is not supported. Contrary to what was predicted, the results from the Pearson correlation with **all** participants' scores show a *negative* association between the span test and weighted lexical density that is statistically significant in the description task, r (13) = -.57, p < 0.05, but not in the narrative task, *r* (13) = -.39. Results from the Pearson correlation, **without** the scores of participants 12 and 13, show that this negative association reaches significance in both tasks: *r* (13) = -.61, p < 0.05, for the description, and *r* (13) = -.68, p < 0.01, for the narrative. This unpredicted negative association is further supported by the results from Spearman's Rank Order Correlation, **with all** participants' scores: r_s (13) = -.68, p < 0.01, for the description task, and r_s (13) = -.49, p < 0.01, for the narrative task.

Taken together, these results might be interpreted as an indication that individuals with a larger working memory capacity were less prone to producing lexically dense L2 speech when this was measured by weighted lexical density. In other words, these participants tended to make use of a relatively small number of different lexical items, which made these items highly frequent in their speech samples, thus lowering the items' weight in comparison to weighted grammatical items and to the overall number of linguistic items. Although unexpected, these findings are consistent with the trade-off view of working memory capacity resources and of speech production processes. In line with results obtained by Foster & Skehan (1996), Mehnert (1998), and Ortega (1999), there seems to be, in the present study, an interaction among fluency, accuracy, complexity, and lexical density during L2 speech production, so that gains in some of these aspects result in losses in other aspects. This trade-off is also claimed by theories of working memory, which pose that the system makes use of a resource-allocation policy when task demands exceed its capacity.

At this point, an interim summary of the statistical analyses is necessary. First, the analyses showed that participants' results on the operation-word span test are at ceiling and do not exhibit the variation necessary to address the second research question of the present research project--whether working memory capacity is specific to the task in which it is being applied or is a more general phenomenon, maintained across tasks. The effect of this lack of variability can be seen in the absence of significant correlations with the speaking span test or any of the measures of L2 speech production. Together, these results show that Hypothesis 5 cannot be supported by this study.

Second, because the analyses revealed the presence of an outlier in the speaking span test and a second outlier in the association of this variable with L2 speech production measures, two additional correlational techniques were used: the Pearson Product Moment Correlation without the scores of participants 12 and 13 and Spearman's Rank Order Correlation with all participants' scores entered.

These complementary analyses revealed that the outliers demonstrate a somewhat unstable behavior which affects the data in different ways. Removing participants 12 and 13 or using all participants' scores in rank order sometimes strengthens, sometimes weakens, and sometimes has no effect on the relationship between the span test and variables of L2 speech production.

Third, the results from the Pearson and Spearman's correlations, overall, provided support for Hypotheses 1 through 3, but not for Hypothesis 4. These results might be interpreted as evidence for an association between working memory capacity, as measured by the speaking span test, and fluency, accuracy, complexity and, although in the opposite direction, weighted lexical density. In other words, individuals with a higher working memory capacity speak faster, have a higher mean length of run, are less prone to making errors, and present a more grammatically complex speech. Taxing working memory capacity with the complex cognitive processes necessary to speak faster, more grammatically correct and more grammatically complex might, however, affect the lexical dimension of L2 speech production. Together, the results from the correlations computed converge to the idea that the data from the present study might be taken as preliminary evidence for an association between working memory capacity and L2 speech production.

Finally, the analyses also indicated that this association is not straightforward, in the sense that higher working memory capacity means higher fluency, accuracy, complexity, and lexical density. Two trade-offs were identified in the present study: one between silent pauses and hesitations and the other among fluency, accuracy, complexity, and weighted lexical density.

To determine the extent to which working memory capacity, as measured by the speaking span test, can predict performance on L2 speech production, in the dimensions of fluency, accuracy, and complexity, simple linear regressions were computed . This is the focus of the next section.

5.4 Simple linear regressions

Simple linear regressions were computed so as to determine the predictive power of working memory capacity, as measured by the speaking span test, on measures of L2 speech production, both in the picture description task and in the narrative task. As Walsh (1990) points out, before computing linear regressions, it is useful to inspect the scattergram of the joint distribution of the two variables. The scattergram plots the values for the two variables and indicates the nature of the relationship between them. To illustrate, the scattergram shown in Figure 5.9 indicates that the relationship between working memory capacity, as measured by the speaking span test (SST), and speech rate unpruned (SRU) in the picture description task, is a positive, relatively strong relationship. On the other hand, the scattergram plotting the relationship between the operation-word span test (OWST) and speech rate unpruned (SRU) in the same task (Figure 5.10) shows that the association between the two variables is nonexistent:

Figure 5.9 Relationship between SST and SRU in the picture description task.

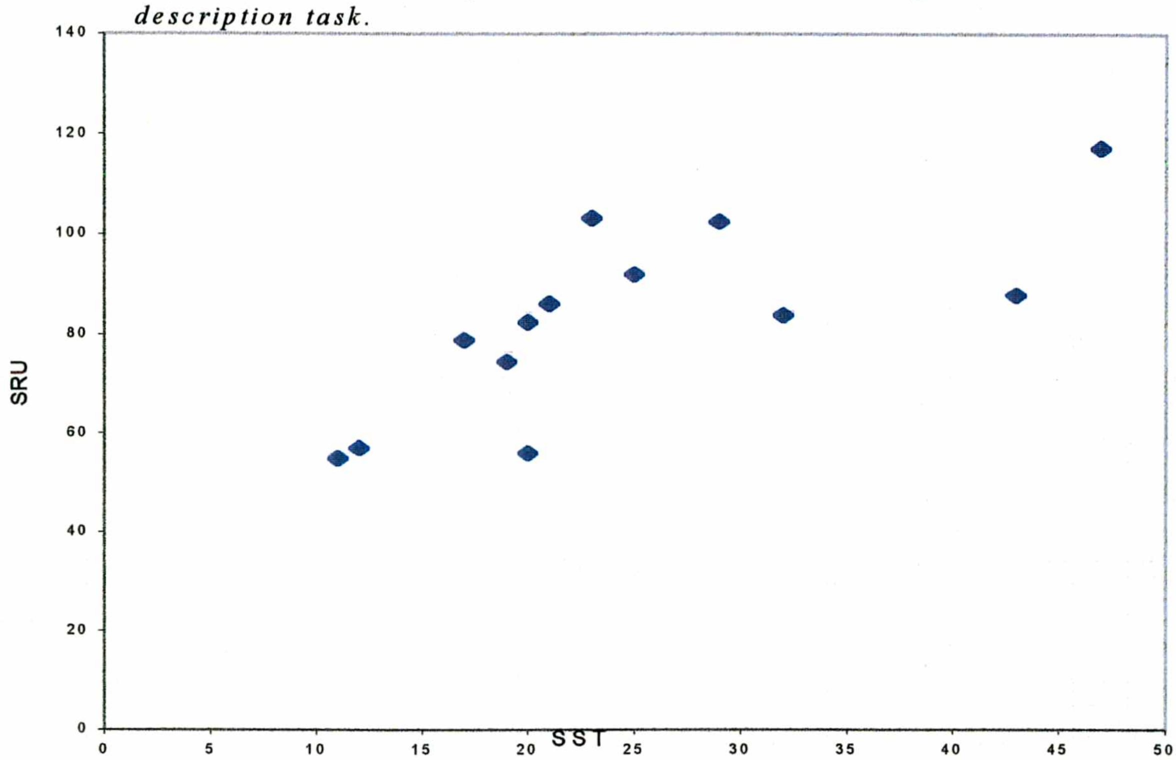
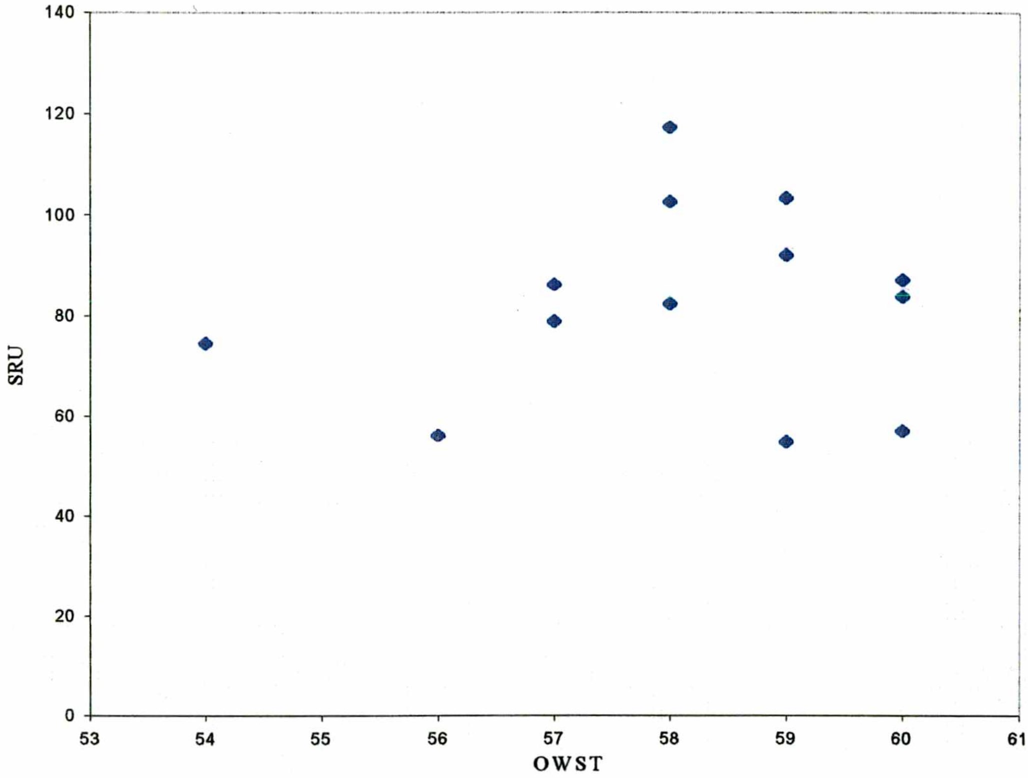


Figure 5.10 Relationship between the OWST and SRU in the picture description task



Thus, based on the inspection of the scatterplots of the association between each span measure and each L2 speech production variable and on the strength of the relationship between these variables as indicated by the Pearson Product Moment Correlation Coefficient, it was possible to take the informed decision of determining the predictive power of the speaking span test on those variables of L2 speech production with which it had a significant correlation, in the direction hypothesized. Values for the speaking span test and each variable of L2 speech production were entered in the following formula:

$$Y = \beta_0 + \beta_1 X$$

Where Y is the dependent variable--the one to be predicted--and X, the independent variable, the predictor--in this case, the speaking span test.[3]

Table 5.15 presents the results of the simple linear regressions computed to determine the predictive power of working memory capacity, as measured by the speaking span test (SST), (1) on the fluency measures speech rate unpruned (SRU), speech rate pruned (SRP), and mean length of run (MLR), (2) on accuracy, as measured by number of errors in syntax, morphology, and lexical choice per hundred words (Acc.), and (3) on complexity, as measured by number of dependent clauses per minute (Comp), in the picture description and narrative tasks:

Table 5.11
Summary of Simple Linear Regression Analysis for the SST Predicting Variables of L2 Speech Production

	Picture Description Task			Narrative Task		
Variable	β_0	β_1	r^2	β_0	β_1	r^2
SRU	51.14	1.2	.535	59.22	1.4	.485
SRP	48.53	1.2	.525	57.64	1.4	.382
MLR	2.1	4.5	.499	2.3	4.6	.386
Acc.	16.65	-.28	.288	10.88	-.17	.240
Comp.	-.73	.18	.587	2.6	.12	.293

N = 13
Note. For all analyses, $p < 0.05$

[3] As McClave et al (1997) note, in regression analysis the phrase *independent variable* refers to a predictor variable for *y*.

These results show that the speaking span test is a significant predictor of L2 speech production when this is assessed in terms of speech rate, mean length of run, accuracy, and complexity in monologic tasks, such as a picture description task and a narrative task. The values expressed by β_1 show the mean unit increase--or decrease--in the speech production variable for each unit increase in the speaking span test. For instance, the model shows that the estimated mean speech rate unpruned in the picture description task increases by 1.2 semantic units per minute for each additional sentence produced in the speaking span test. The values expressed by β_0 show that, when working memory span is equal to 0, speech rate unpruned has an estimated value of 51.14 semantic units per minute. Nevertheless, as McClave et al. (1997) note, β_0 values can be meaningfully interpretable only if $x = 0$ is within the range of the independent variable. Since participants' working memory span, as measured by the speaking span test, varied from 11 to 47, $x = 0$ does not apply to the present study and β_0 values are not subject to meaningful interpretation.

Predicting L2 speech production scores from scores on the speaking span test in the narrative task tends to reproduce the results obtained in the picture description task, thus indicating that working memory capacity, as measured by the speaking span test, is a significant predictor of L2 speech production in the dimensions of fluency, accuracy, and complexity in monologic tasks. Thus, in the narrative task, for instance, for each additional increase in the participants' working memory span, the estimated average increase in speech rate unpruned is of 1.4 semantic units per minute and 4.6 semantic units in the participants' mean length of run.

Table 5.12 reports the results from simple linear regressions without scores of participants 12 and 13:

Table 5.12
Summary of Simple Linear Regression Analysis for the SST Predicting Variables of L2 Speech Production

	Picture Description Task			Narrative Task		
Variable	β_0	β_1	r^2	β_0	β_1	r^2
SRU	37.84	1.9	.535	40.52	2.4	.419
SRP	35.27	1.9	.520	29.35	2.9	.488
MLR	1.91	5.6	.639	2.03	6.4	.267
Acc.	24.7	-.70	.527	16.25	-.45	.493
Comp.	-1.48	.22	.401	-.48	.28	.539

N = 11
Note. For all analyses, $p < 0.05$

As Table 5.12 shows, without the scores of participants 12 and 13, the estimate mean unit increase in each variable is slightly higher for all measures, in both tasks. Based on the results reported in Tables 5.11 and 5.12, it appears to be possible to reject the null hypothesis (H_0)--that is, that working memory capacity, as measured by the speaking span test, contributes no information for the prediction of fluency, accuracy, and complexity of L2 speech production in a description and a narrative task--and accept the alternative hypothesis (H_a)--that is, that working memory capacity, as measured by the speaking span test--contributes information for the prediction of fluency, accuracy, and complexity of L2 speech production in a description and a narrative task for a p-value of 0.05.

The general conclusion we can draw here is that working memory capacity, when measured by the speaking span test in the participants' L2, is at least linearly related to fluency, accuracy, and complexity in L2 speech production--when this is elicited by means of a description and narrative task--with mean performance on these aspects increasing as working memory capacity increases.

McClave et al. (1997) note that one way to measure the utility of a simple regression analysis and gain additional information on the linear relationship between two variables is to look at the coefficient of determination, r^2. The coefficient of determination indicates the amount of variance in the dependent variable--the predicted variable--that can be explained by or attributed to the independent variable--the predictor (Hatch, 1987; McClave et al, 1997; Walsh, 1990). Thus, as Table 5.11 shows, in the picture description task, the speaking span test accounts for 53% of the variation in speech rate unpruned, 52% of the variation in speech rate pruned, and 49% of the variation in the mean length of run. By the same token, working memory capacity, as measured by the speaking span test, accounts for 58% of the variation in complexity but for only 28% of the variation in accuracy. The implication here is that a great amount of the variation in each of these variables is left unexplained or is attributable to other variables. Table 5.11 displays lower percentages for the performance of the narrative task. In this task, working memory capacity accounts for 43% of the variance in speech rate unpruned, 38% of the variance in speech rate pruned, and 38% of the variance in the mean length of run. It explains only 24% of the variance in accuracy and only 29% of the variance in the complexity of L2 speech production.

In the picture description task, leaving out the scores of participants 12 and 13 does not change the amount of variance in speech rate that can be explained by working memory capacity: variation in the speaking span test accounts for 53% of the variation in speech rate unpruned and 52% in speech rate pruned. However, a greater amount of variation in mean length of run (63%) and accuracy (52%) can be explained by the speaking span test, without the outliers. Finally, 40% of the variation in complexity is attributed to variation in the speaking span test, without the outliers. In the narrative task, the r^2 tells us that 41% of the variation in speech rate unpruned and 48% in speech rate pruned are accounted for by the speaking span test without the scores of participants 12 and 13. Thus, for speech rate unpruned, this percentage is slightly smaller than when all participants' scores are entered for computation, but it is greater in speech rate pruned. The amount of variation in mean length of run that can be accounted for by working memory capacity decreases without the outliers, in the narrative task: here only 26% is attributed to the speaking span test. Finally, without the outliers the amount of variation that can be accounted for by working memory capacity doubles for accuracy and complexity, in the narrative task: 49% and 53%, respectively. Thus, again, the unstable behavior of outliers produces different effects on the regression results, at times increasing the coefficient of determination (r^2), but most times making it smaller.

The important point here, nevertheless, is that although a linear relationship seems to exist between working memory capacity, as measured by the speaking span test, and fluency, accuracy, complexity and, albeit, in an initially counter-intuitive direction, weighted lexical density, and that the speaking span test is a significant predictor of performance on L2 speech production, in terms of fluency, accuracy, and complexity, a great proportion of the variation in these dimensions is not explained by working memory capacity alone.

Summarizing, the results obtained in the present study indicate that working memory capacity, as measured by the speaking span test, is positively related to L2 oral fluency, as measured by speech rate unpruned, speech rate pruned, and mean length of run. Working memory capacity, as measured by the speaking span test, is negatively related to number of silent pauses per minute, but this relationship reaches significance only in the two alternative tests (the Pearson correlation without the scores of participants 12 and 13 and Spearman's Rank Order Correlation). Although, contrary to predictions, working memory capacity, as measured by the speaking span test, is positively related to number of hesitations per minute, the findings for the variables of fluency lend substantial support for the hypothesis that the speaking span test is related to L2 continuous performance in real time. The results also indicate, as predicted, that working memory capacity, as measured by the speaking span test, is negatively related to accuracy, as measured by number of errors per hundred words, and positively related to complexity, as measured by number of dependent clauses per minute. Finally, the results show that, contrary to predictions, working memory capacity, as measured by the speaking span test, is negatively related to weighted lexical density. These results together might suggest that individuals with a larger working memory capacity speak faster, have a higher mean length of run, tend to display fewer long silent pauses, and exhibit speech that is more error free and grammatically complex than those individuals with a smaller working memory capacity. However, this seems to be achieved through a greater number of hesitations and relatively low lexical density.

In the present study, the operation-word span test was affected by a lack of sufficient variation in scores and a ceiling effect, thus providing inadequate data to discuss whether individuals' working memory capacity is specific to a task or is maintained across domains. The speaking span test, in turn, was shown to be a significant predictor of fluency, accuracy, and complexity in L2 speech production. The findings of the present research project might be taken as tentative evidence for a linear relationship between working memory capacity and L2 speech production. Working memory capacity, however, accounts for only a portion of the variation in L2 oral performance. The next chapter presents a discussion of the results of the present study in light of existing research.

CHAPTER 6

DISCUSSION

The discussion in this chapter addresses the research questions and hypotheses of the present study in light of the results reported in the previous chapter and existing research in the areas of working memory capacity and L2 speech production. It is organized in 8 main sections.

Section 6.1 presents a summary of the statistical results. Section 6.2 presents a preliminary discussion of the need to determine the processes common to both working memory and L2 speech production. Section 6.3 presents a proposal for envisaging L2 grammatical encoding as a complex subtask which requires the control and regulation of several cognitive processes. Section 6.4 presents a proposal for conceptualizing working memory as the site of attentional resources. This section is further subdivided into 4 sections, each presenting evidence that is consistent with the findings obtained in the present study. Section 6.5 discusses the relationship between working memory capacity and weighted lexical density and Section 6.6, the relationship between working memory capacity and pauses and hesitations. Next, Section 6.7 discusses the problems found in the application of the operation-word span test. Finally, Section 6.8 discusses the amount of shared variance between the speaking span test and measures of L2 speech production.

The crux of my argument will be that individuals with a higher working memory capacity, as measured by the speaking span test, have a greater amount of attentional resources to be shared among 6 macro cognitive mechanisms involved in the encoding of an L2 message--activation, maintenance, suppression, strategic search and retrieval, and monitoring--and that there are trade-offs between fluency, accuracy, complexity, and lexical density, as well as between silent pauses and hesitations.

6.1 Summary of results

The present study was undertaken to investigate the relationship between working memory capacity and L2 speech performance. Working memory capacity was assessed by means of the speaking span test (Daneman, 1991; Fortkamp, 1998, 1999) and the operation-word span test (Cantor & Engle, 1993; Turner & Engle, 1989; La Pointe & Engle, 1990; Fortkamp, 1998), both in the participants' L2. Speech production in the L2 was assessed by means of a picture description task and a narrative task, and four aspects were examined: fluency, accuracy, complexity, and lexical density.

To reiterate, the first research question this study pursued was whether there was a statistically significant relationship between task-dependent working memory capacity, as measured by the speaking span test in the participants' L2, and measures of fluency, accuracy, complexity, and weighted lexical density in L2 speech productions. The results from Pearson and rank-ordered correlations as well as from simple linear regressions support the conclusion that working memory capacity is linearly related to fluency, accuracy, complexity and weighted lexical density in L2 speech production.

Hypothesis 1 predicted a positive relationship between working memory capacity, as measured by the speaking span test, and speech rate unpruned, speech rate pruned, and mean length of run, and a negative relationship between working memory capacity and number of pauses and hesitations per minute.

The Pearson Product Moment Coefficient of Correlation (*r*) yielded significant positive correlations in both tasks between the speaking span test and speech rate unpruned, speech rate pruned, and mean length of run. Without participants 12 and 13, the Pearson correlation yielded similar results in the description task for the variables speech rate unpruned and pruned, and a stronger correlation for the variable mean length of run. In the narrative task, the association between the speaking span test and speech rate unpruned does not change, but it decreases when the variable is speech rate pruned, and loses significance for mean length of run. Spearman's Rank Order Correlation yielded statistically significant results in both tasks for all three variables.

These results were interpreted as providing overall support for the hypothesis that individuals with a larger L2 working memory capacity would be more prone to producing, in real time, L2 speech that is faster and more continuous in both a picture description task and a narrative task.

However, the prediction that working memory capacity would correlate negatively with number of silent pauses per minute and number of hesitations per minute was not statistically supported. The results from the Pearson correlation between the speaking span test and number of silent pauses per minute in the picture description task were negative, as predicted, but not statistically significant. The correlation between the speaking span test and number of hesitations per minute, in both tasks, was positive, contrary to expectations, although not statistically significant. Without the outliers, the coefficient of Pearson correlation between the speaking span test and silent pauses reaches significance in the picture description task, but not in the narrative task, thus providing limited support for Hypothesis 1. The association between the speaking span test and hesitations per minute remains positive in both tasks, reaching significance in the narrative task. Spearman's Rank Order Correlation with all participants' scores yields a negative and statistically significant correlation between the speaking span test and number of silent pauses per minute in both tasks, as predicted in Hypothesis 1, but the relationship between the speaking span test and number of hesitations per minute is positive, gaining significance in the narrative task. Together, these results might indicate a trade-off between silent pauses and hesitations, with individuals with a larger working memory capacity tending to use fewer silent pauses, as predicted, but relying extensively on the use of hesitations.

Hypothesis 2 predicted a negative correlation between working memory capacity, as measured by the speaking span test, and accuracy in L2 speech production, as measured by the number of errors in syntax, morphology, and lexical choice per hundred words. The results from all three statistical tests indicate that this hypothesis is supported, in both tasks, and might show that individuals with a higher working memory capacity are more prone to producing speech that is lexicogramatically error-free.

Hypothesis 3 predicted a positive correlation between working memory capacity, as measured by the speaking span test, and complexity, as measured by number of dependent clauses per minute. Results from all three statistical tests indicate that this hypothesis is supported, in both tasks. These results might be interpreted as an indication that individuals with a higher working memory capacity also tend to produce L2 speech that is more grammatically complex.

Finally, Hypothesis 4 predicted a positive correlation between working memory capacity, as measured by the speaking span test, and weighted lexical density, as measured by the percentage of weighted (or low frequency) lexical items over the total number of weighted linguistic items--high- and low-frequency lexical items plus high-and low-frequency grammatical items--produced during the picture description and narrative tasks. Results from the statistical tests provided no support for this hypothesis. The Pearson Correlation yielded a significant but negative correlation between the speaking span test and weighted lexical density in the picture description task and a negative correlation, not statistically significant, in the narrative task. Without the outliers, the results for the narrative reaches significance. Results from Spearman's Rank Order Correlation tend to reproduce the same general pattern. In other words, there appears to exist an association between working memory capacity and weighted lexical density during L2 speech production, but this association is initially counter-intuitive and indicates that individuals with a larger working memory span tend to use a small number of different lexical items.

The second research question pursued in the present research project was whether working memory capacity was task-specific or a general, domain-free capacity. To address this question, participants were required to perform the operation-word span test, a working memory span test which has mathematical operation as the processing task (in contrast to the speaking span test, which has speech production as the processing task). Hypothesis 5 predicted a statistically significant relationship between working memory capacity, as measured by the operation-word span test, and fluency, accuracy, complexity, and weighted lexical density in L2 speech production, in the directions predicted for the speaking span test.

Hypothesis 5 was not supported. The operation-word span test did not correlate with any of the measures of L2 speech production. Although this finding might be interpreted as an indication that working memory capacity is specific and dependent on the task being performed, a more straightforward explanation might be that the ceiling effect and a restricted range of scores on this variable may have masked differences in performance capabilities. Thus, due to the these two methodological problems, the data obtained in the present research project are probably not appropriate to address the task-dependent/domain-free debate.

To determine whether working memory capacity, as measured by the speaking span test, was a significant predictor of L2 speech production measures, simple linear regression analyses were computed. Results indicate that the speaking span test is a significant predictor, in both tasks at the 0.05 level, of three of the variables of fluency--speech rate pruned, speech rate unpruned, and mean length of run--in addition to accuracy and complexity. This predictive power improves when the outliers are removed from the analyses. Thus, although the coefficient of determination r^2 showed that the amount of variation in L2 speech performance can be only partially accounted for by working memory capacity, as measured by the speaking span test, the results of this study might be interpreted as preliminary, albeit exploratory, evidence for a linear association between individual differences in working memory capacity and L2 speech production.

6.2 The speaking span test and the variables of L2 speech production: preliminary discussion

Rosen & Engle (1998) note that, although obtaining a correlation between performance on working memory span tests and a higher-level cognitive task is an important finding in and of itself, it does not tell us what mechanisms are responsible for this relationship. In other words, to understand how working memory capacity relates to performance we need to examine what it is that the complex span test measures that is also present in the higher-level cognitive task (Engle, Kane, & Tuholski, 1999; Conway & Engle, 1996): we have to look for the interface between working memory processes and, in the case of the present study, L2 speech production processes.

The claim of those working in the psychometric correlational approach is that limitations in working memory capacity differ among different individuals (Haberlandt, 1994) and that these individual differences may be related to variation in the performance of complex cognitive tasks (Carpenter & Just, 1989; Daneman & Carpenter, 1980, 1983; Daneman & Green, 1986; Daneman & Tardiff, 1987; Daneman, 1991; Engle, 1996; Engle & Turner, 1989; Just & Carpenter, 1992, among many others). The L2 speaking span test--devised after Daneman and Green (1986) and Daneman (1991)--taxes individuals' working memory processes during speech production. It requires that individuals keep a set of words stored in memory while simultaneously providing a sentence for each word in the set, in the order and form in which they were presented. L2 speaking span is thus defined as the maximum number of words for which sentences were generated.

In order to construct the sentences during the speaking span test, participants have to make use of processes that are at the heart of speech production: the conceptualization of a message, the construction of a surface structure and a phonetic plan, the articulation of the plan. Individuals who scored higher on this test were those able to provide a greater number of sentences containing the to-be-remembered words in their exact order and form of presentation. But why were they able to provide this greater number of sentences?

From a macro cognitive perspective, one explanation might be that they were better able to coordinate the processing and storage demands of the L2 speech production process *as a whole*; that is, they were better able to carry out the processing of each component as described by Levelt (1989) while also temporarily holding in memory the intermediate products of this processing, as well as the final product (the phonetic plan) until realizing it as overt speech. Their ability to coordinate processing and storage in L2 speech production processes freed their resources to maintain active the words presented in the set, resulting in a greater number of sentences and thus in a higher span. The span (the number of sentences produced) indicates how able an individual is in coordinating the processing and storage functions of his/her working memory. Coordination of processing and storage demands in working memory are also necessary, and will thus be reflected, in the higher-level cognitive task, L2 speech production.

However, span tests are general in the sense that they do not specify the mechanisms involved in the coordination of processing and storage in the background task. All they say is that an individual has greater or less ability to coordinate the overall processing and storage demands higher level tasks require. In this sense, working memory is the aggregate of several mental processes, summarized under the notions of computation (processing) and storage. Such a view of working memory does not seem sufficient to clarify the relationship between the capacity of the system and any higher-level task.

With regard to L2 speech production, the aspects investigated in the current study--fluency, accuracy, complexity, and lexical density--as well as the variables used to assess these aspects, seem to reflect the processes that take place in the formulator (using Levelt's terminology), more specifically the processes involved in the grammatical encoding of the message, the construction of a syntactic structure for the message. Obviously, participants also had to conceptualize and articulate their messages, but the variables used in the present study do not measure conceptualization and articulation per se. Thus, in trying to disentangle the relationship between working memory capacity and fluency, accuracy, complexity, and weighted lexical density in L2 speech production, it will be assumed that, in the case of the present study, the processes that were captured by the speaking span test that are relevant for both working memory and L2 speech performance are those involved in the grammatical encoding of the L2 message, that is, those that take place in the formulator.

In the discussion that follows, it will be argued that L2 grammatical encoding is a complex sub-task that qualifies as a *controlled processing activity*, in Engle & Oransky's (1999) terms, requiring, as such, the activation of information, temporary maintenance of activated information, suppression of irrelevant information, serial search and retrieval, and monitoring of information. In this sense, the results obtained in the present study--that working memory capacity, as measured by the speaking span test, is related to L2 speech production at the grammatical encoding level--are in line with current research showing that working memory capacity is related to the performance of a controlled processing activity: individuals with a higher working memory capacity are better able to activate and temporarily maintain information active, to suppress information, to carry out serial search for and retrieval of information, and to monitor information.

In addition, it will also be argued that working memory, as a central executive processor, is capable of allocating resources when the processes involved in a task exceed its capacity. In this sense, the negative correlation between the speaking span test and weighted lexical density, as well as the positive correlation between the speaking span test and hesitations, result from a trade-off between fluency, accuracy, complexity, and weighted lexical density, on one hand, and between long silent pauses and hesitations, on the other.

Before presenting the evidence that supports the findings obtained in the present study, it is first necessary to explain why L2 grammatical encoding is a complex sub-task that qualifies as a controlled processing activity and to propose a view of working memory that accounts for the relationship between capacity and L2 grammatical encoding.

6.3 L2 grammatical encoding as a complex subtask and as a controlled processing activity

Speaking is a complex cognitive behavior (Clark, 1996; Clark & Clark, 1977; Levelt, 1989; McLaughlin, 1987; Mehnert, 1998), possibly the "most complex skill of *homo sapiens*" (Levelt, 1995, p. 13). Most models of speech production divide speaking into two main phases: planning and execution (Akmajian, Demers, Farmer, & Harnish, 1995; Clark, 1996; Clark & Clark, 1977; Daneman, 1991; Dell, 1986; Levelt, 1989, 1992, 1995, Meyer, 1996). In the planning phase, a series of hierarchical levels of representation is constructed (Dell, Juliano, & Govindjee, 1993). Speakers first construct an internal conceptual representation of what they intend to say and then construct representations at the level of syntax and phonology. The execution phase, in turn, involves articulating what was planned as overt speech. However, the execution phase may start at any given moment of the planning phase, so that, as most models claim, planning and execution during speech production are carried out incrementally and in parallel (Daneman, 1991; Faerch & Kasper, 1983; Levelt, 1989; Meyer, 1996). These two macro-phases of speech production involve a number of subprocesses (McLaughlin, 1987) which take place by means of various mental mechanisms.

Recall from Chapter 3 that L2 investigators claim that L2 speech production shares many of the processes of L1 speech production. Thus, taking Levelt's (1989) model of L1 speech production as the basis for L2 oral production, De Bot (1992) suggests that L2 speech performance would involve the following general sequence of processes: (a) conceptualization of the message, in which its propositional content is developed ; (b) grammatical encoding of the message, where first lemmas and lexemes are accessed and then a surface syntactic structure for the message is built; and (c) phonological encoding of the surface structure followed by the construction of a phonetic plan, which is, in turn, transformed into overt speech.

In the present study, as noted before, the variables used to assess speech production tapped the processes in the formulator, more specifically, those involved in the construction of a surface structure for the message. The surface structure, quoting Levelt (1989, p. 11), is "an ordered string of lemmas grouped in phrases and subphrases of various kinds". To generate a surface structure in the L1, the speaker must conceptualize the preverbal message, which will then activate lemmas. The selection of lemmas that match the preverbal message takes place through the retrieval of those that are in a high state of activation (Levelt, 1989, 1995; Levelt, Roelofs, & Meyer, 1999). After the lemma is selected, its internal grammatical specifications become available to be worked on by automatic syntactic building procedures. Thus, the construction of a surface structure involves multiple steps of processing, constituting a complex task, as defined by Kintsch, Healy, Hegarty, Pennington, and Salthouse (1999). Thus, to the extent that L1 and L2 speech production have similar hierarchical macro-phases (conceptualization, formulation, and articulation) and that grammatical encoding is a subtask in one of these phases (formulation), involving, in itself, various other processes, it can be argued that L2 grammatical encoding is a complex subtask of L2 speech production.

According to Levelt (1989) and De Bot (1992), the formulator is specific to each language: That is, the morphological, syntactic, and phonological encoding processes of L2 speech production are particular to those of the L2. Currently, there seems to be no consensus in the L2 acquisition/use literature on how the L2 grammatical encoding processes take place (De Bot, 1992; De Both & Schreuder, 1993; Kroll, 1993; Poulisse, 1997, 1999; Poulisse & Bongaerts, 1994, among others), but Poulisse (1999) has recently made a proposal.

Based on the analysis of L2 speech errors elicited from 45 L2 learners at three different levels of proficiency, she suggested that L1 and L2 lemmas are organized in a single, multilingual network in the mental lexicon (the basis for grammatical encoding processes), as proposed in Poulisse and Bongaerts (1994), reviewed in Chapter 3. Thus, during L2 speech production both L1 and L2 lemmas are activated simultaneously. Activation spreads to the corresponding L1 and L2 word forms (recall that Levelt, 1989, postulates that word forms--lexemes--can only be activated after the lemma has been activated). Poulisse also suggests that L2 syntactic encoding is ideally language-specific, but that the wrong encoding procedure (probably based on the L1) might be chosen occasionally. Since her data are inconclusive in this respect, it could be that De Bot's (1992) proposal that two speech plans at the level of syntax are encoded simultaneously is correct. Poulisse claims that simultaneous activation of L1 and L2 information is necessary during L2 grammatical encoding because L2 speech production models need to take into account the fact that L2 speakers are able to mix, intentionally or not, the two languages.

Thus, if simultaneous activation of L1 and L2 takes place during L2 speech production, a mechanism is necessary to make it possible that only one language (the L2) be realized as overt speech. Poulisse (1999) and Green (1998) propose that this mechanism is the inhibition or suppression of the L1. Thus, in order for L2 lemmas and their respective syntactic specifications to be selected, they need to be in a high state of activation. Activation of lemmas that match the preverbal message increases as the activation of those that are not relevant for the message decreases through suppression.

One additional feature of L2 grammatical encoding is that, in contrast to L1, the necessary information will not be as automatically retrieved from long-term memory (De Bot, 1992; De Bot, Cox, Ralston, Schaufeli, & Weltens, 1995; De Bot & Schreuder, 1993; Kroll, 1993; Poulisse, 1997, 1999; Poulisse & Bongaerts, 1994; Snodgrass, 1993; Segalowitz, Segalowitz, & Wood, 1998; Schmidt, 1992, among others). Since formulator processes are language-specific, as claimed by Levelt (1989) and De Bot (1992), new mental representations and processes have to be formed in the L2 formulator. These new mental representations and processes will, as a rule, result in incomplete knowledge of the L2: the L2 mental lexicon has fewer words available and, for some of these words, syntactic information may not be fully specified (Poulisse, 1999). It is also quite likely, as suggested by Poulisse (p. 56), that the relationship between the lexical entries of an L2 mental lexicon is not as fully developed as in the L1 lexicon (Levelt, 1989).The L2 speaker, thus, has less linguistic information on which to draw when encoding a message in the L2. Together, these factors might interfere in the selection of lemmas and/or their corresponding syntactic information, leading the L2 speaker to perform a serial search for and retrieval of information that is not readily available.[1]

So far, L2 grammatical encoding processes have been described as requiring the simultaneous activation of L1 and L2 lemmas and lexemes, the suppression of L1 information, and the serial search for and serial retrieval of L2 information that is not immediately available. In the context of the present study, accuracy was one of the measures of L2 speech production adopted. It is possible, therefore, to argue, that the participants of the study, in order to speak accurately, also performed some monitoring to ensure that their output was error-free. Activation, suppression, serial search, serial retrieval, and monitoring are cognitive mechanisms that seem to be part of *a controlled processing activity* (Engle & Oransky, 1999; Engle, Kane, & Tuholski, 1999), that is, an activity which demands controlled processing.

In a series of publications, Engle and colleagues claim that individual differences in working memory capacity are more prone to being reflected in those activities that demand controlled processing (Engle, 1996; Engle & Oransky, 1998; Engle, Kane, & Tuholski, 1999; Engle, Tuholski, Laughlin, & Conway, 1999; Kane, Bleckley, Conway, & Engle, in press; Kane & Engle, in press; Rosen & Engle, 1997 and 1998). These activities, according to Engle and colleagues include situations (a) when it is necessary to apply *activation* to memory representations, bringing them into focus and maintaining them in focus, (b) when it is necessary to *maintain* information active in the face of distraction or interference (c) when it is necessary to *suppress* irrelevant information, (d) when *strategic search* and *retrieval* of information is necessary, (e) when *monitoring* for and correcting errors are necessary, and (f) when conflict among actions must be resolved

---

[1] Serial search, in this context, means purposeful or strategic search for information in long-term memory.

to prevent error in the output. Relating this description to the account of L2 grammatical encoding given above, it seems plausible to argue that L2 grammatical encoding includes, to a large extent, these situations, therefore qualifying as a controlled processing activity (Engle, 1999, personal communication).

6.4 An attention view of working memory

Engle, Kane, and Tuholski (1999) conceptualize working memory as "a system consisting of those long-term memory traces active above threshold, the procedures and skills necessary to achieve and maintain that activation, and limited-capacity, controlled attention" (p. 102). In their framework, working memory capacity is operationalized as the "capacity for controlled, sustained attention in the face of interference or distraction" (p. 104). Thus, when they talk about a controlled processing activity, it is attention that is being controlled--or regulated, for that matter--so that it can be divided among the processing the activity entails, some of which being activation, temporary maintenance of active information, suppression, serial search, serial retrieval, and monitoring.

This controlled-attention view of working memory has several advantages over the quite general and vacuous processing and storage view of working memory proposed by Daneman (1991). Although there is no consensus, in the literature, on what the mental energy or limited resources of working memory consist of (Tomitch, 2000, personal communication), it seems feasible to characterize these resources as attentional for several reasons.

First, it is in line with Baddeley's original conceptualization of the central executive as an attentional controller in his tripartite working memory model (e.g.

Baddeley & Logie, 1999), which, he claims, is conceptually similar to the Supervisory Attentional System (SAS) described by Norman and Shallice (1986), a system in charge of controlling resources during performance. The central executive, which is the object of study of the psychometric correlational approach to working memory, is fundamentally related to attention and control (Kintsch et al., 1999). Approached from an attentional viewpoint, working memory (the central executive) could very well be labeled "working attention" (Baddeley & Logie, 1999, p. 52). Second, it is in line with the notion of attention as a limited mental energy or mental resource (Kintsch et al., 1999; Just & Carpenter, 1992; Messick, 1996). Finally, the view that attention is a limited mental resource that has to be shared among various processes is also widely accepted in the L2 acquisition/use literature (VanPatten, 1994, 1996; Skehan, 1998). In the present study, thus, the limited resources of working memory are characterized as attentional, although the present researcher is fully aware of the fact that these resources have not been fully accounted for in the existing literature. In light of the above, it seems plausible to argue that L2 grammatical encoding processes require that attention be controlled and divided among activation, maintenance, suppression, serial search, serial retrieval, and monitoring.

Thus, in order to explain the statistical relationship between the speaking span test and fluency, accuracy, complexity and, although in the direction opposite from that predicted, weighted lexical density, the specification of the mechanisms that might be involved in this relationship are as follows. In terms of *language* production processes, those assessed by the measures used in the present study were the ones involved in L2 grammatical encoding, in the formulation phase of L2 speech production. In terms of *cognitive* processes, those that are proposed to operate during L2 grammatical encoding are the processes of simultaneous activation, suppression, temporary maintenance of activation, serial search, serial retrieval, and monitoring. These might have been the processes captured by the speaking span test that are also relevant for the higher-level task in question: L2 speech production.

Recent developments in working memory theories have shown that individual differences in working memory capacity are indeed related to an individual's ability for *controlled processing*, as defined above. Much of the evidence for this relationship is based on laboratory tasks which do not always address language processing directly. Extensions to L2 speech performance will, thus, only be metaphorical and tentative. The next four subsections present this evidence.

6.4.1 Working memory capacity and activation of information

Cantor and Engle (1993) provide evidence that individuals vary in the total amount of mental energy they have available to activate representations (knowledge units) in memory and maintain them active. They divided participants into high- and low-working memory span groups and asked them to perform a fan-effect task, a task that has been assumed to measure activation of information in long-term memory. Participants had to learn a series of unrelated sentences consisting of a subject and a predicate (e.g. The lawyer is in the house/The plumber is in the park). Each subject in the sentence was paired with more than one predicate (e.g., The lawyer is in the boat/The lawyer is in the park). Participants were also asked to perform a speeded recognition task in which they verified whether a given sentence belonged to the studied set or not (e.g. The lawyer is in the church/The teacher is in the park). In the fan-effect task, recognition times increase as a function of the number of times a subject in the studied set appears with a different predicate. That happens because a great number of knowledge units have to be activated at the same time for sentences with many different predicates. Cantor and Engle hypothesized that individual differences in working memory capacity would be related to recognition times. Results showed that these increases in recognition times occurred for both high- and low-span participants, but they were greater for the low-span participants. Cantor and Engle suggested that low-span individuals took longer to recognize whether a sentence whose subject had a large number of predicates belonged to the studied set or not because they had less overall resources to be shared among the activation and temporary maintenance of various knowledge units associated with that sentence.

As Green (1998) notes, current accounts of L2 speech production suggest that activation of L1 and L2 is a fundamental mechanism in L2 speech processes. Concurrent activation of L1 and L2 implies that a great number of knowledge units is active at the same time. Thus, an initial explanation for the results of the present study might be that the individuals with a higher working memory capacity, as measured by the speaking span test, had a greater amount of mental resources--a greater amount of attentional resources--available to activate and maintain temporarily active this great number of knowledge units. These knowledge units are the L1 *and* L2 lemmas that match the preverbal message.

Now, in order to be selected, the L2 lemmas need to be at a higher state of activation than the L1 lemmas. The mechanism that might allow that to happen is suppression or inhibition of L1 lemmas.

6.4.2 Working memory capacity and suppression

Inhibition or suppression of information seems to be a particularly important mechanism in L2 processes in the formulator. Green (1998) has recently proposed an inhibitory control (IC) model which postulates that L2 speech production requires multiple levels of control of the two language systems. By control, he means the regulation of the two language systems which would be modulated by the supervisory attentional system (SAS) proposed by Norman and Shallice (1986). In his view, during L2 speech production both languages are active and competing to be selected. Successful selection of L2 lemmas will require the inhibition (or suppression) of non-target lemmas.

Evidence that a greater working memory capacity is related to the suppression of information in the performance of higher-level cognitive tasks demanding controlled processing--that is, regulation of attention--comes from Rosen & Engle (1998). In this study, they investigated interference from intrusions and the suppression of intrusions using the paired associates task. Briefly, in paired-associates tasks participants learn lists of pairs of words such as "book-train", "bird-table", and are then asked to recall "train" when presented with "book", and "table" when presented with "bird" . In this case, "book" and "bird" are cue or A words, and "train" and "table" are response or B words (Rosen & Engle, 1998). To introduce interference, the experimenter asks participants to learn a second or third list in which A words are the same but B words are changed.

By manipulating the design of a paired-associates task, Rosen & Engle showed that individuals with a higher working memory capacity experienced fewer intrusions during second-list learning than those with a lower span. In addition, they showed that high span individuals were able to suppress first-list response items during second and third-list learning. Based on this evidence, it might be possible to suggest that, in the present study, individuals with a higher working memory capacity, as measured by the speaking span test, were better able to suppress L1 lemmas, that is, their greater attentional resources enabled them to suppress L1 lemmas while also maintaining the L2 lemmas necessary for the message temporarily active.

Now, with L1 lemmas suppressed, the degree of activation of L2 lemmas increases and their syntax becomes available for the construction of the surface structure. It is possible that in the activation/suppression activity learners need to perform some search for L2 knowledge items, either at the level of lemma selection or at the level of the syntax of the lemma. It is possible that these items are not readily available in long-term memory due either to the level of proficiency of the speaker (De Bot, 1992; Poulisse, 1997 and 1999; Poulisse & Bongaerts, 1994) or to momentary forgetting.

6.4.3 Working memory capacity and strategic search and retrieval

Evidence that individuals with a larger working memory capacity seem to be better able to serially search and retrieve relevant knowledge structures in tasks demanding controlled processing has been provided by Rosen and Engle (1997).

In a series of four experiments, they showed that there was a relationship between performance on the operation-word span test and the retrieval of animal names under different conditions (digit tracking and memory preload). Participants were divided into higher- and lower-span. In the first experiment, all participants were required to generate animal names for 15 minutes, avoiding repetitions. In the second, one half of the participants in each span group were asked to generate animal names while concurrently tracking digits, whereas the other half in each span group generated animal names without the concurrent task. In the third experiment, all participants were required to memorize and recall a 12-word list before generating animal names. The list contained either animal names or building-part names. Finally, in the fourth experiment participants were asked to generate animal names for 10 minutes while also tracking digits. This time, participants were encouraged to repeat animal names already retrieved. Together, these tasks require activation of information, suppression, strategic search, and monitoring for errors.

Findings showed that high-span individuals generated more animal names than low-span individuals. The concurrent digit tracking task reduced the number of animal names retrieved by the high-span participants, but had no effect on the retrieval by low-span participants, who tended to make more repetitions. In addition, the preload memory task (the list) reduced the number of animal names by the high-span participants, but not the low-span. Overall, low-span participants were, however, more likely to repeat already retrieved animal names than their high-span counterparts. Rosen and Engle interpreted these findings as evidence that in order to perform the tasks, all participants had to monitor responses as well as suppress information, in order to avoid repetitions. In doing that, only the high-span participants had enough working memory capacity to search for animal names while also suppressing already generated animal names and monitoring for errors. Participants in the low-span group were much more susceptible to the activation of animal names already generated, being less able to suppress this information and to search for new animal names.

Based on this evidence, it might be possible to suggest that, in the present study, participants with a higher working memory capacity, as measured by the speaking span test, had a greater amount of attentional resources available to strategically search and retrieve the L2 lemmas or their corresponding syntactic information when this was not readily available from their long-term memory, while also holding active previously activated items and suppressing interfering or non-relevant information. Their greater attentional resources might also have allowed them to activate new items and perform a faster search for and retrieval of items. This is particularly visible in the fact that individuals with a higher working memory capacity also tended to produce a greater number of hesitations--immediate repetitions, nonlexical pauses ("uh" and "uhm"), and short silent pauses (of .49 seconds or less), as shown by the results from the correlations.

Immediate repetitions imply that those individuals with a higher speaking span were indeed holding the repeated item active while also looking for a word or its syntactical information. While searching, they were also able to activate and maintain active new units, the nonlexical pauses. Although nonlexical pauses do not carry meaning in themselves, they are part of the speaker L2 lexicon, since languages realize nonlexical pauses in different ways. Furthermore, nonlexical pauses are a good indication to the listener that the speaker is not finished and more is to come, which is a proposition that might be stored in the form of declarative knowledge in the speaker's long-term memory and that could have been activated when s/he had to look for specific information. Finally, individuals with a higher speaking span also produced fewer short and long silent pauses, which indicates that they were also faster in their search for and retrieval of information that was not immediately accessible.

6.4.4 Working memory capacity and monitoring

The study mentioned in the previous section (Rosen & Engle, 1997) also provides evidence that individuals with a higher working memory capacity are better able to monitor for and correct errors. Indeed, Levelt (1989) argues that, in L1 speech production, the two tasks that draw most heavily on the speaker's attentional resources are the conceptualization of the message and self-monitoring. Self-monitoring can take place either after articulation, through self-corrections, or before articulation, when the segment to be expressed as overt speech is being temporarily held in working memory as internal speech. In Levelt's view, the speaker compares what is prepared to be said or what was said to what was intended. Speakers monitor for meaning as well as for linguistic adequacy, checking whether formulation processes yielded the appropriate outcome at the lexical, syntactic, and morphological levels (pp.13-14, 460-462).

In the present study, individuals with a higher working memory capacity, as measured by the speaking span test, were also better able to produce error-free L2 speech than their lower-working memory capacity counterparts. One possible explanation might be that their greater attentional resources allowed them to engage in self-monitoring of output to prevent syntactic, morphological, and lexical choice errors. In other words, in addition to activating L1 and L2 lemmas relevant for their message, deactivating L1 lemmas, increasing the activation of L2 lemmas, and activating the syntactic information related to these lemmas, higher-span speakers also had enough attentional resources left to allocate to verifying whether those were the appropriate lexicogrammatical items for the message they wanted to verbalize. In contrast, the lower-span speakers might have allowed either L1 lexical and syntactical information or the inappropriate L2 lexicogrammatical information to be selected, thus incurring in L2 error, because they did not have enough attentional resources to both suppress inadequate information and verify whether L2 lemmas and their accompanying syntactic information were accurate.

Further evidence that individual differences in working memory capacity are important in controlled processing is presented by Kane, Bleckley, Conway, and Engle (in press). They conducted two experiments in which participants, divided into high- and low-working memory span groups, were asked to perform a visual-orienting task in which they had to detect an abrupt-onset visual cue in the environment and use that cue to move attention and eyes to a spatial location that would present a target. In the prosaccade version, the signal appeared in the same location as the target. In the antisaccade version, the signal appeared in the opposite location as the target. Participants were required to identify the target (a letter) as quickly and as accurately as possible in the two versions. The prosaccade version predictably required attention to be directed towards the abrupt-onset signal, allowing responses based on relatively automatic processes. The antisaccade version, on the other hand, required attention to be directed away from the signal, thus requiring participants to control their attention and prevent the use of automatic processes, in this case, eye movements. Kane et al. showed that in the prosaccade task high- and low-span individuals performed equivalently. However, in the antisaccade task, which required the control of attention, high-span individuals were able to detect targets more quickly and more accurately than low-span individuals.

In experiment 2, the researchers explored the effects of practice on the antisaccade task, for both groups, by adding further antisaccade trials and monitoring eye-movement. Results replicated those of the first experiment--individuals with a higher working memory capacity, as measured by the operation-word span test, were faster and more accurate in identifying the target than their lower-span counter-parts. In addition, this difference was maintained across 360 trials, showing that the performance of the low-span individuals was not related to practice. Kane et al. interpreted these results as evidence that working memory capacity is related to controlled attention, "an ability to effectively maintain stimulus, goal, or context information in an active, easily accessible state in the face of interference, and/or to effectively inhibit goal-irrelevant stimuli or responses" (p.27).

The results obtained in the current research project might be interpreted as an indication that individuals with a larger working memory capacity, as measured by the speaking span test, spoke faster and longer (between pauses and hesitations), with fewer errors, and more complexly, while producing a picture description task and a narrative task. In light of the evidence provided above, it might be argued that individuals with a larger span, as measured by the speaking span test in their L2, had a greater amount of attentional resources to be shared among the activation of L1 and L2 lemmas that corresponded to their intended message, the suppression of L1 lemmas, the temporary maintenance of active L2 lemmas for the building of the surface structure and the phonetic plan for the message. A greater amount of attentional resources might also have contributed to their being better able to search for and retrieve L2 knowledge units needed to speak, when these were not immediately available. Rosen and Engle (1997, 1998) have also provided evidence that individuals with a higher working memory capacity are more prone to monitoring for errors. In the case of the present study, it could be that those participants with a higher working memory capacity were also better able to monitor for incorrect output, thus producing more accurate speech.

Now, one question that has intrigued me throughout this research project is why it is that they have this greater amount of attentional resources available. One very tentative explanation might be that some of their formulation processes are automatized, thus freeing their attention for the attention-consuming tasks of activating, suppressing, maintaining active, monitoring, strategically searching and retrieving knowledge units.

The distinction between automatic and controlled processing, together with the notions that humans are limited-capacity processors, is now a consensus in cognitive psychology and is fundamental to the understanding of human cognitive behavior. It is also widely recognized that all complex skills involve a mixture of controlled and automatic processing. As a rule, automatic processing does not demand processing resources, thus freeing the cognitive system for the more complex, higher-level processing of the task. Because it does not share cognitive resources, automatic processing is highly efficient and can be carried out in parallel. Controlled processing, on the other hand, is assumed to make great demands on the capacity of the system--which can attend to only a few things at a time--tending to be carried out serially and, therefore, slowly.

Determining what is under automatic processing and controlled processing during the performance of complex tasks is still a challenge for cognitive theorists (McLaughlin & Heredia, 1996). The present study does not offer an empirical basis for determining what, in the L2 formulator of the participants of the present study, is automatized and what is not. However, as Schmidt (1992) suggests, variables such as speech rate and length of speech run reflect automaticity rather directly. In the present study, there was a statistically significant correlation between working memory capacity and speech rate, unpruned and pruned, and mean length of run, in both tasks. Thus, although these variables do no tell us what aspects are automatized, it could be that individuals with a higher working memory capacity, as measured by the speaking span test, have greater attentional resources available because they also have a greater amount of automatized knowledge of the L2 than those individuals with a lower working memory capacity, who have to allocate a great part of their attentional resources to the controlled processing of linguistic knowledge. That is, the automatized linguistic knowledge of the higher-span participants frees their attentional resources for the control of the macro cognitive activities involved in the grammatical encoding of the L2 message--activation, maintenance, suppression, strategic search and retrieval, and monitoring. As a result, their performance is more successful than that of the individuals whose L2 knowledge is not as automatized.

The degree of automatized and controlled processing of L2 knowledge might contribute to the greater amount of attentional resources of higher-span individuals, but this possibility should be taken with caution for two reasons. First, as Shiffrin (1997) has recently pointed out, although it is generally agreed that automatic processing does not entail attentional resources, there is evidence, mainly from visual studies, that automatic processing can draw on attentional resources. In addition, Shiffrin notes that, depending on the cognitive load imposed by the task, automatic processing can be carried out serially, as most controlled processing. In other words, it is not clear in the literature whether automatic processing does not require attention. Second, current theories of working memory capacity have not been able to determine exactly why some individuals have greater resources than others. Automatic processing is one of several factors (Carpenter & Just, 1989).

In the present research project I have proposed that the grammatical encoding of L2 messages be envisaged as a sub-task of the overall L2 speech production process. As a sub-task, it involves the control of attention to be divided among the processes of activating and suppressing information, of bringing information into focus and maintaining it in focus (Engle & Oransky, 1999), of carefully searching for and retrieving information, and of monitoring for the correct output. Doing all this draws on the resources of working memory and those who have more of it available will probably go about the task better. To illustrate, compare the following speech samples produced by a lower-span individual (participant 1) to the samples produced by a higher-span individuals (participant 9), in the picture description and narrative task (see Chapter 4 for coding):

Picture description task

Participant 1:

Uhm: (0.6) the left one (.) it is a (.) in Egypt . **(0.7) uh: (1.1)** it has a big pyramid **(0.9)** and (.) **uh (1.2)** *ma-* maybe inside **(1.5)** it has a mummy: . **(0.8)** and **uh (1.3)** I think that **(1.8) uhm (1.4)** many people **(1.2)** take some trips to visit mummy . **(0.8)** in these days **(1.6) uhm (0.6)** I (.) left side means a **(1.8) uhm (1.6)** people **(1.8) uhm (1.6)** astronaut **(1.3)** it means a: **(1.4) uh (1.9)** how people **(0.72)** people's power **(2.3)** its' growing . **(0.7)** and a **(1.)** and **(4.6)** last year **uh** last year ? No **(0.7)** *s-* **(0.6)**maybe couple of years ago (.) people already **(0.5) uh** reached the: (.) Mars . **(1.7)** so (.) **uhm** also then (.) for messages they want to **(1.1)** send us . **(1.4)** but **uh (2.7)** people is always **uh: (2.2)** look for the new thing and **uh: (1.)** how people **(1.3)** people's power (.) is growing (.) *growing* up **(1.7)** and **(1.5) uhm (2.4)** *and* **uh (0.6)** maybe long standing *ci- civirili-* civilization **(1.6)** I think .

Total time: 1m57s80 **(117,80s)**
No. of words: 108/103

Participant 9:

I see **uh:** a picture: **(.)** that is **uh** divided into two halves . **(0.7) uh (.)** the left hand **(.) uh** side **(1.1) uh** shows an **(.)** Egyptian **uhm** temple: **(.)** with a: **(2.)** very big **uh (0.9)** symbol of an Egyptian god in the: very center of the picture . **(0.7)** and **uh** the right handed side **uh** shows **(.) uh** a flying astronaut in the sky . **(1.1) uh (.)** the connection of the two: halves **(.) uh** is given **uh** through that the two: **(0.7) uh** figures the Egyptian god and **uh** the astronaut **(.)** are in the: **uh (1.1)** very front and splitted in two halves . **(.)** as *as* if each **(.)** as if they were one person **(0.8)** but still the two different sides give a very different impression and a different mood . **(.) uh** and **uh** give the *the:* **uhm (1.8)** the impression of really two very different pictures that are just **(.)** thrown together **(1.) uh** the **(0.5)** *the* left hand side ? of the Egyptian **uh** temple **(.) uh (.)** sh:ows this very big **uh** god *sh-* **uh** from **(.)** the top to: the: **(0.8) uh** basic of **uh** the picture: **(.)** and **uhm (0.6)** on the left hand side ? shows **uh** a row o:f **uhm (.)** columns **uh** of **uh** six columns **(0.6)** that **uh** lead to the end **(.) uh** there's no move on *on* **uh** the temple ? you can look into the: **uh (.)** kind of **uh:** very bright ? sounding sky **(0.6)** right hand side oh is it finished ?

Total time: 2m **(120s)**
No. of words: 184/180

Narrative task:

Participant 1:

It was a **(1.5) uhm (3.5)** man which: **(1.5)** try to: **(.)** save the: **(1.5)** earth . **(1.6)** he **uhm (2.5) uhm (0.8)** some **(1.2)** I'm not sure the **(0.8)** correct *b-* **(.)** word but **uh (1.)** meteorite **(.)** *meteorite* **(1.3)** almost **uh (.)** hits **uh** us . **(.) uh (.) uhm (0.6)** so **(0.7) uhm (2.5)** American army **(1.2)** send a **(.)** *a* special person to: **(0.8) uhm (0.5)** send a **(1.3)** *send a* special person **(2.1)** save the **(.)** earth . **(1.7)** and **uh (2.0)** they **(0.8)** *they* success **(0.6)** *success* but **uh (1.7)** they are proud **(.)** but **uh: (1.5) uh (.)** successful **(1.3)** or but **uh (0.6) uh (0.6)** someone had to: **(1.0)** leave the **(1.7)** meteorite **(.)** *meteorite* **(.)** *meteorite* so **uhm (4.1)** do you know the **(.)** actor ? Bruce Willis **uhm (2.9)** was a **(1.)** *a* **(3.7)** *a* **(1.4)** Bruce Willis **(2.3) uh (1.2)** decided to: **(1.3)** live the **(1.7)** live there and **uh (3.5)** stay there.

Total time: 1m59s87 **(120s)**
N. of words: 86/77

Participant 9:

Yesterday night **uh** (.) I saw **uh** (.) the film Gold Rush by **uh** Charles Chaplin . **(0.5)** and **uh** I saw that a long time ago (.) **uh** but I couldn't I could hardly (.) *hardly* remember it anymore . **(0.8)** you don't know the film ? so **uh** (.) **uh** (.) the story of the film (.) is very easy (.) it's **uh** basically is about this *li-* fellow (.) **uh** played by Charles Chaplin who goes out into the world (.) to **uh** make his money to make a fortune and **uh** to: **uh:** (.) get **uh** the woman **uh** that he loves . **(0.7)** so **uh** the story is very simple and there is a *a* happy end (.) in that story as well . **(0.8)** but **uhm (1.9)** yeah **(0.7)** first of all in that (.) **uh** whole **uh** story the language is not important at all . (.) like in <u>all</u> Chaplin films **(0.6)** **uh** it's more this kind of basic (.) **uh** playing he's doing . **(2.3)** and **uhm (0.5)** I liked it especially (.) **uh** because of **uh** the (.) I liked it especially because of **uh** **(0.7)** the: **uh** possibility of Chaplin to express (.) really basic skills (.) and basic feelings **(0.6)** of all people in a very inter-national (.) language . **(0.9)** so **uh** I always feel very touched when I see him **(0.8)** **uhm** emotionally: **(0.9)** **uh** and **uh** he touches me (.) especially that (.) he can be so: **uh** show so opposite so different things in on in *in in* one scene . (.) so he can be **uh** poor and rich (.) at the same time and **uh uh** especially use a lot to be weak **(2m)**...

Total time: 3m5429 (**only 120s counted)**
No. of words (up to 2 m): 211/206

As can be seen from the samples, the speech performance of participant 9--a higher working memory capacity speaker--when compared to that of participant 1, in both tasks, was relatively more fluent, with a faster speech rate (pruned and unpruned and longer runs). Participant 9 was also more accurate, as indicated by the relatively small number of errors in syntax, morphology, and lexical choice, as well as more complex, as indicated by the higher number of subordinate clauses, as compared to Participant 1.

In sum, these examples show that a higher L2 speaking span corresponds to higher scores in three aspects of L2 speech production. Together, the results of the current study might, thus, be an indication of a relationship between L2 working memory capacity and fluency, accuracy, and complexity in L2 oral performance in monologic tasks. Those who have greater ability to control activation and suppression, to maintain items activated, to search for and retrieve not readily accessible information, and to monitor output seem better able to deal with the cognitive demands of encoding a message in the L2.

This, however, has its costs. Hypothesis 4, in the present study, predicted that individuals with a higher working memory capacity, as measured by the speaking span test, would also produce more lexically dense L2 speech. There was no support for this prediction in the present study and this will be discussed in the next section in terms of a trade-off between fluency, accuracy, complexity, and lexical density.

6.5 Working memory capacity and lexical density

Theories of working memory claim that the mental computations involved in the performance of a complex task compete for the limited capacity of the system, so that disruptions in performance may occur when several mental processes have to be carried out concurrently (e.g., Baddeley, 1990; Baddeley, 1992a, 1992b, 1992c; Baddeley & Logie, 1998). In order to deal with concurrent mental computations, the system is capable of differentially allocating resources (Gathercole & Baddeley, 1993; Baddeley, 1996; Just & Carpenter, 1992; Saariluoma, 1998).

By the same token, interactions have been found across the various levels or aspects of speech production in various studies (Ratner, 2000). For instance, Nelson and Bauer (1991, cited in Gathercole & Baddeley, 1993) analyzed spontaneous speech samples of a group of two-year-old children and found a trade-off between the complexity of word combinations and the phonetic complexity of individual words. In

Gathercole and Baddeley's (1993) view, this type of evidence may be indicative of the allocation of resources in working memory so that the system can handle the various concurrent processing demands during speech production.

Hypothesis 4 in the present research predicted that individuals with a higher working memory capacity, as measured by the speaking span test, would also produce more lexically dense speech both in the picture description and narrative tasks. This hypothesis was not supported. In fact, there was a significant negative correlation between working memory span and weighted lexical density. This finding might be explained in terms of a trade-off in working memory between fluency, accuracy, complexity, and weighted lexical density.

As noted before, the variables used in the present study to assess speech production tapped the processes in the formulator, more specifically grammatical encoding processes. Following McLaughlin (1987), grammatical encoding was suggested to be a sub-task (or a sub-goal) in the hierarchical process of L2 speech production. As such, it requires what Engle et al (1999a, 1999b) have termed controlled processing, or the control and regulation of attention to orchestrate activation, temporary maintenance of relevant active items, suppression, the search for and retrieval of items that are not readily available, and monitoring of output. It is clearly, thus, an activity that overloads the naturally limited capacity of working memory, regardless of individual performance, requiring that the system prioritize some aspects to the detriment of others.

It might have been the case, then, that in order to speak faster, more accurately, and more complexly, the participants of the present study had to rely on the use of those lexical items that were more easily available from long-term memory, that is, those which corresponded directly to the concepts in the preverbal message, which were in a high state of activation, and about which they had language specific syntactic information more easily available. In other words, they might have used the L2 lemmas whose level of activation could more easily be kept above that of the correspondingly L1 lemmas and whose lexemes were well-developed, accurate and easily accessible. The use of the same lexical items throughout either the description or the narrative task increased the frequency of these items, which had to be given half the weight of a lexical or grammatical item appearing only once in the speech sample, thus affecting the lexical density of their oral production.

From the perspective of speech production theories, resource allocation in working memory seems motivated by the fact that the formulation of a message is initiated by first activating lemmas--that is, formulation processes are lexically driven (Levelt, 1989). All the information necessary for surface structure generation is contained in the lexical entries of the mental lexicon. Given the amount of processing preceding and following the activation of lexical entries--that is the conceptualization of the message, the construction of a phonetic plan, and the articulation of the message--and the speed with which this processing takes place, it is likely that the speaker will rely primarily on using lemmas that are already highly activated.

The use of a great number of different lexical items to express the contents of a message is a task that, in itself, requires a large amount of controlled search and retrieval of information even in L1 spontaneous speech production, when speakers are not under time pressure, strictly speaking. Levelt (1989) states that lexical access is such a high-speed process that there is probably no other cognitive task demanding such high rates of decision making. For instance, in L1 speech production an average speaker has to select a word every 400 milliseconds. Although in L2 speech production more time is needed to select words (De Bot & Schreuder, 1993), there must be a mechanism within the process of speech production, in general, that allows for rapid selection. Griffin and Bock (1998) suggest that words that have been retrieved and used during a given oral task tend to be re-used by speakers, probably as a means to deal with processing demands. Since these words are in a relatively high state of activation, they can be selected more easily, therefore freeing the speaker's attentional resources for other aspects of the task. Just and Carpenter (1992), Just, Carpenter and Hemphill (1996), and Carpenter and Just (1989) call this the resource allocation policy. Allocation of attentional resources during demanding cognitive tasks has been consistently shown to be related to working memory capacity, with higher span individuals being better able to dynamically reallocate their resources when these cannot meet the task's demands (King & Just, 1991; Carpenter & Just, 1989; Just & Carpenter, 1992; Tomitch, 1995).

A related factor that might have contributed to the negative correlation between working memory capacity and lexical density is that to produce more grammatically complex language we need a larger number of grammatical items. It seems that in limiting their utterances to relatively simple and short structures, participants with a smaller working memory capacity, as assessed by the speaking span test, decreased their number of grammatical items while at the same time increasing the weight of lexical items. They did not seem to have sufficient attentional resources to activate and maintain active L2 lemmas and their corresponding lexemes and simultaneously work on the internal grammatical structure of their message, by selecting alternative syntactic information which allowed for the embedding of sentences into one another. Although an exact explanation is lacking, it seems that their strategy to deal with the demands of the oral tasks was to construct mainly short sentences in the subject + verb + object order and link these either by coordinated conjunctions (*and*, *but*) or silent and nonlexical pauses. Participants with a larger working memory capacity, on the other hand, had sufficient attentional resources available to, in addition to controlling the macro cognitive activities of activating, temporarily holding active, suppressing, searching, retrieving, and monitoring, also deal with specific linguistic information such as the ordering of sentences and their internal relations.

In the present study, time pressure was an extrinsic load in the picture description task that might have made it more demanding than the narrative task. The degree of the relationship between the speaking span test and speech rate (pruned and unpruned), mean length of run, accuracy, and complexity was higher for this task than for the narrative task, (see Tables 5.7 and 5.10). This finding is in line with Just and Carpenter's (1992) claim that individual differences in working memory capacity are more evident in tasks that exceed the capacity of the system. The more the task exceeds capacity limitations, the greater the possibility of disruptions in performance, making the task an optimal environment for the emergence of individual differences. The narrative task, in this sense, might have been less demanding, since participants were free to speak as long as they wanted. In this case, the narrative might not have taxed working memory capacity as heavily as the picture description task, thus the slightly lower correlations.

In general, nevertheless, and based on the fact that there seems to be a trade-off between aspects of speech production and that in this trade-off allocation of resources is a necessary mechanism, it could be tentatively argued that, in trying to deal with the demands of describing a picture and narrating a story, participants with a higher working memory capacity, as measured by the speaking span test, were also better able to allocate their cognitive resources in order to accomplish the tasks.

The view that there is a competition for cognitive resources during L2 speech production is also shared by SLA researchers investigating the effects of planning time on L2 oral performance. Foster and Skehan (1996a, 1996b), Mehnert (1998), and Ortega (1999), reviewed in Chapter 3, have provided evidence for interaction effects resulting from processing capacity limitations.

Foster and Skehan (1996), for instance, showed that when oral tasks do not require learners to express complex ideas, fluency is given priority. They also showed that the more complex the speech output, the less accurate it is. Along the same lines, Mehnert showed that when tasks do not require the expression of complex ideas, priority is given to accuracy. If complex language needs to be used, however, complexity will affect fluency and/or accuracy. Ortega, in turn, showed that gains in fluency were accompanied by gains in complexity, with no changes in lexical range and nonlinear gains in accuracy. In the present study, the degree of association between working memory capacity and fluency measures--the two types of speech rate and mean length of run--accuracy, and complexity, as shown by Pearson Product Moment Coefficient of Correlation taking the results of all participants, seems to indicate that cognitive resources were first allocated towards fluency, then complexity, and to a much lesser extent, accuracy--in both the picture description and narrative task[2].

These results are in line with Ortega's study in that higher fluency is accompanied by higher syntactic complexity, an interaction also noticed by Wolfe-Quintero, Inagaki, and Kim (1998). Wolfe-Quintero et al., reviewing 39 studies of L2 writing, found that gains in fluency progressed linearly with gains in complexity, whereas gains in accuracy were much more unstable, being observed for some grammatical forms

---

[2] This pattern is maintained in Spearman's Rank Order Correlation as well as Pearson Product Moment Correlation without the outliers.

but not others. That is, when fluency, accuracy, and complexity are examined in L2 output, accuracy is the variable that seems to suffer the greatest disruption. It could be, as Crookes (1989) and Foster and Skehan (1986) note, that the longer, faster, and more complex the utterances in the L2 become, the greater the chances of errors, probably due to limitations in attentional resources. In the case of the present study, it might have been the case that participants with a higher working memory capacity were better able to monitor for these errors while also producing faster, longer, and more complex speech.

Whereas the relationship between fluency, accuracy, complexity, lexical density (in Mehnert's study) and lexical range (in Ortega's study) is clearly complex, the evidence converges, nevertheless, on the fact that increases in one aspect take place through decreases in other competing aspects, due to constraints on resources. Although these studies do not refer to the concept of working memory and do not address individual differences in attentional resources, the results of the present study seem to be compatible with the general notion that these aspects of speech production interact during performance as a result of limited cognitive resources.

Producing L2 speech that is fluent, accurate, and complex, albeit lexically simple, has its costs. Hypothesis 1, which is related to temporal measures of fluency, predicted that there would be a negative correlation between the speaking span test and number of silent pauses as well as number of hesitations per minute, both in the description and the narrative tasks. This aspect of the Hypothesis was not supported. The next section discusses pauses and hesitations in L2 speech production.

6.6 Pauses and hesitations in L2 speech production

Traditionally, studies of speech production have taken silent pauses and hesitations as evidence of transient problems during speech production. Garret (1982), for instance, suggested that silent and filled pauses reflect increases in processing load and temporary inaccessibility of a unit of knowledge necessary to convey the message in any of the three macro-level stages--conceptualization, formulation, articulation.

This view is shared by a number of L2 speech production researchers. For instance, Dechert (1987) states that temporal variables and speech errors indicate planning problems L2 learners have when producing speech. Likewise, Rohde (1985) suggests that breakdowns in the production of L2 speech reflect linguistic problems the speaker encounters. Möhle (1984) proposes that by investigating temporal variables such as speech rate and silent pauses, as well as hesitation phenomena such as filled pauses, repetitions, and self-corrections, it is possible to draw conclusions about the problems of L2 speech production. The same position is advocated by Raupach (1984) and Lennon (1984). Wiese (1984) argues that hesitation phenomena such as filled pauses, repetitions, corrections, and prolongations constitute overt expression of difficulties in the planning and execution of speech. In his view, pauses and hesitations are the means whereby L2 speakers gain time to solve speech production problems. Finally, Dörnyei and Kormos (1998) conceptualize different types of pauses and hesitations as time-gaining mechanisms and problem-solving mechanisms to deal with the demands of the L2 speech production process. Thus, the general picture is that pauses and hesitations in L2 oral performance are directly related to the notion of problem in processing. Based on this reasoning, researchers claim that the more efficient processing is, the smaller the number of pauses and hesitations in speech.

Indeed, pauses and hesitations (defined in different ways but always meant to reflect disruptions in speech) have been shown to be salient features of the lack of fluency (fluency defined either as speech that is continuous in time or, more globally, as competence in the production of L2). Several studies have demonstrated that more fluent speakers speak faster, longer, and more, with fewer silent pauses and fewer hesitations (Ejzenberg, 1992; Freed, 1995; Lennon, 1990; Riggenbach, 1989; Temple, 1992; Towell, Hawkins, & Bazergui, 1996), although findings have not reached statistical significance in some of these studies.

The rationale for Hypothesis 1 was that if working memory capacity, as measured by the speaking span test, is positively related to speech rate and speech rate is inversely related to pauses and hesitations, then working memory capacity would be inversely related to pauses and hesitations. That was not the case.

Results from the Pearson Product Moment Coefficient of Correlation indicate a tendency toward a negative relationship between the speaking span test and number of silent pauses (those equal to or longer than 0.5 seconds) in both oral tasks, but the association is not statistically significant. It becomes significant without the outliers only in the picture description task. In the narrative task, leaving the outliers out of computations improves the degree of association but does not make it significant. Results from Spearman's Rank Order Correlation show a negative and significant association between the span test and silent pauses per minute in both the description task and the narrative task, as predicted. Thus, although the evidence is not strong, there seems to be a tendency for individuals with a larger working memory capacity to be less prone to producing long silent pauses per minute during L2 speech production.

On the other hand, the association between the speaking span test and number of hesitations--silent pauses of .49 seconds or shorter, filled nonlexical pauses, immediate repetitions, and partial words per minute--is always positive, reaching significance without the outliers in the narrative task, a result that is maintained in Spearman's Rank Order Correlation. In other words, there is not the slightest indication that individuals with a higher working memory capacity are less prone to hesitating during L2 oral performance.

A tentative explanation for these results might be that there is an interaction in this study between silent pauses of .5 seconds and hesitations. Given the demands of the task, individuals with a higher working memory capacity, as indexed by the speaking span test, were better able to reduce the number of times they would pause silently for .5 seconds or longer but at the cost of hesitating more, as if to gain time to carry out further processing.

The use of hesitations as a device to alleviate the exceeding demands of L2 speech production is, yet, further evidence for the greater amount of attentional resources of higher-span individuals. More specifically, in the case of filled nonlexical pauses, immediate repetitions, and partial words, their greater attentional resources allowed them to engage in two concurrent activities: realize speech while also activate information and start the building of a surface structure. In the case of shorter silent pauses, these speakers, aware of the break in the chain of speech, might have sped up their processing. Participants with lower speaking spans, on the other hand, seemed to have relied more heavily on the use of long silent pauses instead of shorter or non-silent devices, while carrying out processing. This might indicate that their more limited attentional resources did not allow them to activate information while also realizing speech, and that activating information and building a surface structure to go on with the message they wanted to deliver required longer amounts of time.

In a way, the positive relationship between working memory span and number of hesitations per minute might reinforce the idea that pauses and/or hesitations are multifunctional in nature and do not necessarily imply a problem (Lennon, 1984), as if the speaker were not able to accomplish the task (Chafe 1998, 1985). Indeed, as Clark (1996) points out, interruptions in the stream of speech are the rule, rather than the exception and also happen in L1 speech production. They signal the speaker's effort to overcome the natural limitations of a cognitive system that can deal with only a limited number of things at a time. Speakers pause and hesitate for a variety of reasons, including rhetorical purposes, individual styles, L1 pause/hesitation pattern influence and, fundamentally, because "speech production is an act of creation" (Chafe, 1985, p.79), demanding such breaks so that talk continues smoothly. Thus, any expectations to eliminate the trade-off between silent pauses and hesitations, as operationalized in the present study, might actually be far-fetched. It is more likely that working memory capacity will be negatively related to one *or* the other.

To summarize, the results of the present research project might indicate that there is a relationship between working memory capacity, when this is measured by means of a speaking span test in the participants' L2, and fluency, accuracy, and complexity, in the L2, implying that individuals with a higher speaking span seem better able to manage the control over the processing involved in the sub-task of formulating a message in the L2. Formulation of an L2 message was suggested to be a sub-task in the hierarchical process of L2 speech production. As such, it seems to possess the basic features of what Engle and his associates call a controlled processing activity: the activation of information, temporary maintenance of active information, suppression of information, search for and retrieval of unavailable information, and monitoring for errors, requiring, thus, an interplay among the processes of bringing information into focus, maintaining information in focus, and keeping information out of focus. This interplay demands the control or the regulation of attention.

Ideally, such a demanding task would require a great number of more specific non-attention-demanding processes in dealing with the L2--for instance, having L2 lemmas readily available, as well as the syntactic rules applied to them. However, it is generally agreed that the L2 formulator is language specific and because of the knowledge base that has to be formed containing specific L2 information, an interaction between automatic and nonautomatic processes at the linguistic level will occur. This interaction might entail strategic search and retrieval of information, thus imposing an extra load on the already demanding task of formulating a message in the L2. Since it is a task that seems to exceed capacity demands to a large extent, individuals with a higher working memory capacity were better able to control their attention to be shared among activation, maintenance, suppression, and monitoring, in addition to strategic search and retrieval--thus producing speech that was more fluent, accurate, and complex--at the cost of producing speech that was less lexically dense and with a higher number of hesitations.

One of the questions motivating the present research question was whether individual differences in working memory capacity were task-specific, as suggested by Daneman (1991), Daneman and Carpenter (1980, 1983), and Daneman and Green (1986), or a more general phenomenon that could be observed across several tasks, as suggested by Cantor and Engle (1993), Engle, Cantor, and Carullo (1992), and Conway and Engle (1996), among others. To investigate this question, the operation-word span test, an alternative working memory span test, was applied to the same participants and the results correlated to their measures of speech production. Thus, Hypothesis 5 of the present study, predicted that the operation-word span test would correlate to measures of L2 speech production in the directions predicted in Hypotheses 1-4. Hypothesis 5 was not supported because, as already indicated, the operation-word span test did not yield adequate data. A discussion of the possible reason for this is the topic of the next section.

6.7 The operation-word span test and fluency, accuracy, complexity, and lexical density in L2 speech production

One of the current debates in the working memory community is whether individual differences in the limited capacity of the system vary across tasks--that is, whether each individual will have one working memory capacity for reading and another for speaking in the L1, and still different other capacities in the L2--or whether these differences reflect capacity that remains the same for each individual across tasks.

Daneman and Carpenter (1980, 1983) suggested that working memory capacity was task-dependent. They devised the reading span test and showed that performance on this test related to reading measures. In their view, the reading span test taxed the processing and storage functions of working memory. Those with a higher reading span were better able to recall to-be-remembered words because they processed the sentences of the span test more efficiently. Their efficient reading processes left more of their capacity to be used for storage of sentence-ending words. Their efficient reading processes as captured by the reading span test are reflected in reading tasks. However, Daneman and Carpenter maintained that more and less efficient readers could have equivalent capacities in other, non-reading tasks: For instance, a lower-span reader could reveal a higher-span speaker.

Engle and colleagues (Turner and Engle, 1989; Engle, Cantor, & Carullo, 1992; Cantor & Engle, 1993; Cantor & Engle, 1993) have proposed an alternative view, holding that working memory capacity reflects a domain-free ability to carry out higher-level cognitive tasks. To test this hypothesis, Turner and Engle devised the operation-word span test, which has mathematical processing as the background task. They showed that, like the reading span, the operation-word span test is a good predictor of performance in reading tasks. Since 1989, the operation-word span test has been consistently used to predict performance in various other tasks, such as taking notes (Kiewra & Benton, 1988), playing bridge (Clarkson-Smith & Hartley, 1990), and learning a computer programming language (Kylonnen & Christal, 1990).

As indicated in 6.1 and in Chapter 5, in the present research project the operation-word span test did not correlate with any of the measures of L2 speech production, apparently because of the low variability among scores on the operation-word span test in addition to a ceiling effect. Furthermore, participants scored high also in the background task (solving the operation and confirming the result as true or false), which shows that they indeed engaged in processing (see Appendix B, item II for their results).

One possible explanation for both the lack of variability and the ceiling is that the operation-word span test was actually performed as a digit-span task. Their scores on the background task (solving the math operation) were relatively high, indicating that they did engage in processing. Recall that this test, unlike the speaking span test, was participant-paced, allowing participants to take as long as they wanted on an operation until they decided to move on to the next operation in the set. This might have caused the span test, which was designed to tax the dynamic nature of working memory, to be, in fact, two separate tasks, in which case, the recall task was performed as a traditional word-span task and the background task, as an independent processing task. In other words, the participants were not actually performing the two tasks (processing and storing) concurrently.

Being participant-paced, the test allowed for rehearsal and that is what participants might have done to be able to recall so many to-be-remembered words. They might have engaged in rehearsal of previous words in the set and would only press True or False on the keyboard after making sure these words were kept in memory. If this was the case, then the load imposed by the simultaneous concurrent processing and storage demands of the task was considerably decreased by the aid of rehearsal. That was an unpredicted methodological failure, also observed in Fortkamp (1998). In that paper, the lack of correlation between the operation-word span test and temporal variables of L2 speech production was interpreted as evidence for a task-specific view of working memory. However, in that study participants' scores were also close to the maximum, which would not allow for such a conclusion.

Unaware of the fact that the problem could be in how participants were performing the task, I decided to reapply this test in the present research. Here, too, the same pattern was observed. It is noteworthy that the procedures for this test were taken from Turner and Engle (1989). In later publications (e.g. Cantor & Engle, 1993; Kane & Engle, in press; Kane et al., in press), however, researchers explicitly state that pausing is not permitted during the performance of this test, which might indicate that being participant-paced, the test leaves open the possibility for rehearsal. Furthermore, as suggested by Kane & Engle (in press), potential ceiling effects in addition to small sample sizes, as is the case in the present study, limit the power to detect significant correlations, in general. Finally, although Klein and Fiss (1999) have recently reported that the operation-word span test reliably assesses individual differences in working memory capacity, they also make it clear that the rate of presentation has to be controlled in order for these differences to emerge.

Therefore, in the light of the above discussion, the present results can not be taken as adequate evidence for a task-dependent view of working memory capacity and do not allow for an empirically-based discussion of the task-dependent/domain-free debate[3]. Nevertheless, I would like to point out that the debate over the nature of working memory capacity is not a simple question and is one of the most complex issues in current research.

Due to criticism raised by Waters and Caplan (1996) and Engle and his group, Daneman and Merikle (1996) ran a meta-analysis of the data from 6.179 participants in 77 studies investigating the relationship between working memory capacity and language comprehension. Their main objective was to compare the predictive power of the test developed by Daneman and Carpenter (1980) with other measures of working memory. They concluded that working memory span tests, which require concurrent processing and storage, are better predictors of comprehension performance than tests that tap only the storage function of the system, such as word span tests. They also found that tests that include math as the background task are also valid predictors of comprehension.

[3] However, the finding that a simple digit-span test does not correlate with a complex cognitive task is consistent with previous research showing that the construct of working memory reflects a dynamic cognitive system and not simply the storage of information.

In their view, then, it is "the individual's efficiency at executing a variety of symbolic manipulations that is related to comprehension ability. In other words, although Daneman sticks to claims about language comprehension, she accepts that the operation-word span test is a valid predictor of performance, thus agreeing that the background task of the span test need not be task-specific. Furthermore, a second group of researchers (e.g. Just and Carpenter, 1992) propose that working memory capacity varies across domains, with language processing (production and comprehension) being one domain. Thus, it seems that the punchline is that working memory capacity is specific, not to a task (reading as opposed to writing and speaking), but to a domain (language as opposed to visuals-spatial processing).

Although, if strictly taken, the results obtained in the present study have to be interpreted as evidence for a specific view of working memory capacity, I would like to leave this question open. Following Miyake and Shah (1999), it may be that we are approaching the issue in a too simplistic dichotomous way. For one thing, the notion of domain is not specified in the literature. For another, we do not know to what extent domains interact. Finally, we still need to investigate how factors such as age, aptitude, previous experiences, extended practice, organization of knowledge in long-term memory, motivation, among many others, affect performance. The next section discusses to what extent working memory capacity predicts performance in the present study, as indicated by the results from simple linear regression analysis.

6.8 Individual differences in working memory capacity: a pimple on the face of cognition (Reason, in Baddeley, 1990, p.87)?

It goes without saying that correlational data are open to all kinds of alternative interpretation. To try to overcome this problem, researchers in cognitive psychology make use of various sophisticated statistical analyses such as factorial and latent variable analyses. In the present study, one way of going a little beyond the pure correlational results was to run a regression analysis in order to find out the extent to which the variance in measures of L2 speech production can be explained by the variance in working memory capacity. As can be seen in Table 5.11, the speaking span test is a significant predictor of L2 speech production in the dimensions of fluency, accuracy, and complexity, both in the picture description task and the narrative task. In the picture description task, approximately half of the variation in speech rate unpruned, speech rate pruned, and mean length of run--50%--might be explained by the variation in working memory capacity. In the narrative task, these values fall between 48% and 43%. The speaking span test accounts for 58% of the variation in complexity and 28% of the variation in accuracy in the picture description task. These values are much lower in the narrative task, however, with working memory capacity accounting for only 29% of the variance in complexity and 24% in accuracy.

Removing the outliers does not improve the coefficient of determination, r^2, on which these results are based. The most straightforward explanation for this finding might be that the variation left unaccounted for by working memory capacity in each measure of L2 speech production, in the two tasks, is attributable to other variables. Although this is totally speculative in nature, such variables might comprise myriad aspects such as attitude towards the tasks (monologic in nature) as well as towards the whole context in which the participants had to perform (a testing situation), anxiety, familiarity with the tasks, motivation to speak, and--most probably--the level of L2 proficiency of the speaker.

One aspect of the correlational research on individual differences in working memory capacity that has been the target of a great amount of criticism is that even in studies with a large number of participants, statistically significant correlations account for a relatively small amount of the variance in the performance of complex tasks (Roberts & Gibson, 1999). That was also the case in the present study. Although from an information processing theory perspective it might be uninteresting to say simply that other things may account for the unexplained variance, knowing that certain things are unexplained can be better understood if we frame the results obtained in the present study, conducted in a language learning setting, within a broader perspective, perhaps in an educational perspective, that takes into consideration the complexities of being a learner--a cognitive, but also social and affective entity. There are many factors contributing to L2 speech performance, and working memory capacity is only one of them.

The next chapter outlines the conclusions drawn from carrying out the present research project and points out its limitations. Suggestions for further research as well as pedagogical implications are also presented.

CHAPTER 7

Final remarks, limitations, suggestions, and implications

7.1 Final remarks

The objective of the present research project was to examine (a) whether there was a relationship between working memory capacity and L2 speech production and (b) whether this relationship was task-specific or domain-free. Working memory was assessed by means of the speaking span test (Daneman, 1991; Daneman & Green, 1986) and the operation-word span test (Turner & Engle, 1989). Both tests require that participants manipulate while also holding information and therefore tax the dual function of working memory: processing and storage. L2 speech production was elicited by a picture description task and a narrative task and four dimensions of performance were assessed: fluency, accuracy, complexity, and weighted lexical density.

The statistical techniques used to address these questions were the Pearson Product Moment Coefficient of Correlation with and without the scores of participants 12 and 13, Spearman's Rank Order Correlation, and simple regression analyses also with and without the scores of participants 12 and 13. While the results from the Pearson Correlation and regression analyses without participants 12 and 13 as well as the results from Spearman's served the purpose of inspecting more strictly the association between working memory capacity and L2 speech production, the most relevant results for the present study are those from the Pearson Correlation with all participants' scores because they follow mainstream research on individual differences in working memory capacity, methodologically speaking. With only one exception (the correlation between the speaking span test and mean length of run in the narrative task), the alternative correlational and regression techniques reinforce the idea that working memory capacity is linearly associated with fluency, accuracy, complexity, and weighted lexical density, although the correlation with the last variable was in the opposite direction from that predicted. Thus, the following findings were obtained in the present study, as indicated by the results from the Pearson Product Moment of Correlation computed on the scores of the 13 participants.

(1) Finding 1: Working memory capacity, when measured by the speaking span test in the participants' L2 (English), correlates significantly with three measures of oral fluency (speech rate unpruned, speech rate pruned, and mean length of run), in the two tasks. This finding can be taken as preliminary evidence that individuals with a higher L2 working memory capacity are also more fluent when describing a picture and narrating in the L2 and provides substantial (although partial) support for Hypothesis 1.

(2) Finding 2: Working memory capacity, when measured by the L2 speaking span test, correlates significantly with accuracy. This finding can be taken as preliminary evidence that individuals with a higher L2 working memory capacity are also more prone to making fewer errors in syntax, morphology, and lexical choice when describing a picture and narrating in the L2 and thus provides support for Hypothesis 2.

(3) Finding 3: Working memory capacity, when measured by the L2 speaking span test, correlates significantly with complexity. This finding can be taken as preliminary evidence that individuals with a higher L2 working memory capacity are also more

prone to producing a greater number of dependent clauses when describing a picture and narrating in the L2 and provides support for Hypothesis 3.

(4) Finding 4: Working memory capacity, when measured by the L2 speaking span test, does not correlate significantly with number of silent pauses (of .5 seconds or greater) and hesitations per minute in either task, but there is a tendency toward a negative relationship with silent pauses, as predicted, and a positive relationship with hesitations. This finding can be taken as very tentative preliminary evidence that individuals with a higher L2 working memory capacity have to rely on the use of silent pauses and, to a much greater extent, hesitations, in order to be able to speak faster, with longer runs, fewer errors, and more complexity. This finding fails to provide support for Hypothesis 1 and for this reason this Hypothesis in only partially supported.

(5) Finding 5: Working memory capacity, when measured by the L2 speaking span test, correlates significantly with weighted lexical density. However, this correlation is negative and contrary to what was expected. This finding can be taken as preliminary evidence that individuals with a higher L2 working memory capacity tend to produce speech with lower lexical density; that is, they use a smaller number of different lexical items when describing a picture or narrating in the L2. The lower lexical density in their speech samples might be related to the fact that L2 speech production is lexically driven and speakers tend to reuse those lexical items that are already activated so that resources can be directed at other aspects of the production process. This finding fails to support Hypothesis 4.

(6) Finding 6: Findings 4 and 5, nevertheless, can be taken as evidence for two trade-offs in L2 speech production. The first trade-off is that between silent pauses and

hesitations. In the present study, those individuals who produced a smaller number of silent pauses tended to produce a greater number of hesitations, and vice-versa. The second trade-off is that between fluency, accuracy, complexity, on one hand, and weighted lexical density, on the other. Individuals with a higher L2 working memory capacity spoke more fluently, accurately, and complexly, but their speech presented lower lexical density and tended to present a greater number of hesitations. This finding likely reflects the fact L2 grammatical encoding, the process more directly assessed by the variables used in the present study, draws heavily on the limited resources of working memory, and for this reason the system has to allocate its resources. This interpretation is congruent with the claim in the cognitive psychology literature that working memory makes use of a resource allocation policy when the task exceeds its capacity (Carpenter & Just, 1989; Carpenter, Miyake, & Just, 1994; Just & Carpenter, 1992; Just, Carpenter, & Hemphill, 1996).

(7) <u>Finding 7:</u> The speaking span test is a significant predictor of speech rate unpruned, speech rate pruned, mean length of run, accuracy, and complexity. This finding adds to earlier research showing that the speaking span test is a valid instrument to assess working memory capacity and to predict L1 as well as L2 oral performance (Daneman, 1991; Daneman & Green, 1986; Fortkamp 1998, 1999).

(8) <u>Finding 8:</u> Working memory capacity, as measured by the L2 speaking span test, accounts only partially for the variation in fluency, accuracy, and complexity in L2 speech production. This finding indicates that a great part of the variation in L2 speech performance is attributable to other variables, probably including the level of L2 proficiency of the participants of the present study.

(9) Finding 9: The operation-word span test suffered a methodological failure and did not yield adequate data to address the issue of whether individual differences in working memory capacity are task-specific or domain-free. For this reason, while the results obtained in the present study are in favor of the task-specific view, the most cautious approach seems to be to consider the discussion on the nature of working memory capacity still open, at least in the realm of L2 research.

The relationship between working memory capacity and L2 speech production needs to be further specified in terms of the processes which are common to both the speaking span test and L2 speaking. For this, it is necessary to determine which processes are attributed to working memory that are also relevant to the processes assessed by the variables in the present study. The proposal advanced was that the L2 speech production process in evidence in the present study was the grammatical encoding of the message, which was conceptualized as a controlled processing activity. As such, L2 grammatical encoding requires that the limited resources of working memory be shared among various processes. These limited resources were characterized as attentional resources and the processes were proposed to be activation, maintenance, suppression, strategic search and retrieval, and monitoring. Thus, the covariance of working memory capacity and L2 grammatical encoding can be explained by suggesting that individuals with a higher L2 speaking span have greater attentional resources to be shared among the activation of L1 and L2 lemmas, the maintenance of activated L2 lemmas and suppression of L1 lemmas, the search for and retrieval of L2 lemmas when these are not readily available, and the monitoring of L2 output. In addition, individuals with a higher working memory capacity also tend to make adjustments in the allocation of their resources, thus appropriately using short periods of silence together with sound while carrying out processing and taking advantage of lexical items that were already in focus to maintain the flow of speech.

The proposal made in the present study goes beyond the general notion that working memory capacity is the capacity to process and store information to specify what, in L2 speaking, is meant by process and storage. It is also congruent with the view in L2 research that speech performance is constrained by the operations of a limited capacity information-processing system (Foster & Skehan, 1995, 1996; Mehnert, 1998; Skehan, 1992, 1998). The present research project identifies this system as being equivalent to the construct of working memory, qualifies the limited resources as being attentional resources, and goes one step further by showing that individual variation in the amount of these resources is related to variation in L2 speech performance.

The results obtained in this study are compatible with both the processing efficiency and the total capacity views of working memory. In the context of research on individual differences in working memory capacity and reading comprehension, Just & Carpenter (1992) have proposed that individuals vary in the amount of activation they have available for information processing and maintenance during language processing. Thus, they qualify the resources of working memory as activation, claiming that the term is more refined than the general label "attentional resources". However, they also state that individual differences in working memory capacity can be explained in terms of the efficiency of mental processes; that is, individuals with a higher working memory capacity have more efficient language processes. Just and Carpenter claim that the two views are mutually compatible. The experiment carried out in this study does not allow for discriminating the two views and the results obtained might favor both explanations.

Thus, it might be that individuals with a higher working memory capacity, as measured by the speaking span test, have greater attentional resources (or activation, in Just and Carpenter's terminology) and (or because of) more efficient processing when speaking in the L2.

The results of the present study are also in line with previous research showing that there are trade-offs in the human information-processing system. The claim for competition among sub-tasks in the performance of complex tasks is raised by researchers in the area of working memory (e.g., Just & Carpenter, 1992; Gathercole & Baddeley, 1993) in general, and in the area of L2 speech performance, in particular (e.g., Skehan, 1992, 1996, 1998; Mehnert, 1998). The present study presents further evidence for this trade-off.

Thus, in the main, the results of the present study support research in cognitive psychology showing that working memory is involved in the performance of a complex cognitive task, L2 speaking, and that individual differences in working memory capacity are related to performance. They also support research in L2 speech performance by showing how the constraints of the so-called limited capacity information-processing system might influence performance.

7.2 Limitations of the study and suggestions for further research

The present study is the result of a systematic effort to examine the role of working memory capacity in L2 speech production. As stated in the Introduction, this study was carried out for the purpose of understanding better L2 speech performance from a cognitive perspective. It was theoretically and methodologically based on the existing literature on working memory capacity and L2 speech production. However, research on working memory deals mostly with L1 processing and, within this, mostly with reading comprehension. In the same manner, research on L2 speech production rarely, if ever, mentions explicitly the construct of working memory and the idea of individual differences in the capacity of the system. Thus, the present study is totally tentative and exploratory and suffers from a number of severe limitations in both the theoretical and practical aspects. Next, I present its limitations as well as suggestions for much-needed future research.

(1) Sample size: The results of the present study can not be generalized to the L2 population from which the sample was taken (adult foreign language learners of English recently arrived in the American university context) due to the size of the sample: 13 participants. Published studies on working memory capacity do not have a sample size smaller than 50 participants: most are conducted with hundreds of them. Although in the L2 area sample sizes are usually smaller than those in cognitive psychology, the tendency in the field is to invest research efforts in obtaining sample sizes of at least 30 participants if the analysis involves statistical procedures. Thus, the present results are limited to this sample and can, at best, be taken as a trend for the speakers who took part in the study. Future research in the L2 area dealing with psychometric issues such as individual differences in working memory capacity needs to involve a much greater number of participants.

(2) Participants: The participants of the present study were speakers of various native languages, many of them speakers of Asian languages. It is widely accepted in the L2

literature that L1 influences L2 performance, including oral performance. For example, disruptions in the flow of speech have been claimed to be largely related to L1 patterns (Olynyk, Sankoff, & d'Anglejan, 1983; Riggenbach, 1991). Thus, L1 background was one additional variable that could not be controlled in the present study and that probably influenced the participants' performance. Although the relationship between the speaking span test and speech rate replicates the findings obtained in Fortkamp (1999), which had a larger sample size of participants with the same native language, the ideal situation would be to have participants with the same L1 background, especially since the study is so exploratory and so limited in sample size. Future research addressing the relationship between working memory capacity and L2 speech production should be carried out with participants of the same L1, which would allow for running the working memory span tests and the oral tasks in their L1 and L2 and would thus make it possible to go further into the question of whether capacity is task-specific and domain-free. Controlling for participants' L1 would also allow for educated inferences as to the possible influence of L1 on L2 speech performance.

(3) Participants' level of L2 proficiency: In the present study the only instrument used to assess the participants' level of proficiency in the L2 was an in-house placement test to which this researcher was not allowed access. Based on their results on the test, they were placed in an advanced oral skills class and were thus judged to be at approximately the same level of proficiency in the language. Clearly, this is not enough. As Carpenter, Miyake, & Just (1994) point out, individual differences in performance may arise from differences in long-term knowledge. Many researchers rule out the possibility that differences in degree of knowledge or skill in the complex cognitive task account for variation in performance, qualifying this possibility as "uninteresting" (e.g., Rosen &

Engle, 1997). However, Miyake & Shah (1999) have recently noticed that the working memory community is starting to take into consideration the fact that long-term knowledge and degree of skill are important factors in working memory performance. Part of the reason why these factors have been overlooked in mainstream research on working memory and individual differences in working memory capacity is that the complex cognitive tasks dealt with in most studies are those that are well developed in the adult college population: language comprehension, reasoning, problem-solving. Thus, these skills are almost a given. In the case of second/foreign languages, however, significant differences in learners' level of proficiency are likely to be more prominent due mainly to differences in educational backgrounds, in general, and type and length of exposure to the L2, in particular. Assessing L2 proficiency with a valid and reliable instrument is still an obscure area in L2 research. The very concept of "proficiency" is problematic in the field. However, every effort should be made to minimize differences in L2 knowledge among participants of quantitative (especially correlational) studies because, once the experimenter has an index of each participant's level of proficiency, s/he can partial this value out of statistical analyses and thus be more confident about the effects of L2 knowledge on the variables examined. In the present study, L2 level of proficiency stands as an invisible variable likely affecting the relationship between working memory capacity and L2 speech performance.

(4) <u>Elicitation of L2 speech:</u> In the present study, L2 speech was elicited by means of two monologic tasks widely used in studies on L2 speech performance. In an exploratory study like the present one it is probably a sound choice to use these types of task, since more interactive oral tasks are much more open to extraneous variables. However, the

picture to describe as well as the topic of the narrative were selected on an intuitive basis. The criterion to select the picture was that it should be open to both concrete description and subjective comments on the part of the participant, in which case the "topic" of the picture should be of general interest and knowledge. The criterion for selection of the topic of the narrative was that it should motivate participants to speak while also minimizing the burden of an experimental situation in which the tape-recorder was their interlocutor. For both tasks, these criteria were generated from the researcher's perspective. Future research should develop a set of specific criteria for the choice of monologic tasks that takes into consideration (a) the participants' perspective and level of L2 proficiency, (b) the control of extraneous variables, and (c) the fact that language use is a social and joint action (Clark, 1996). That can be possibly achieved by conducting a pilot study in which the oral tasks are evaluated by a pool of judges on the basis of a rubric containing operational definitions of the points outlined above. For the purpose of the present study, monologic tasks seemed to be the most adequate technique of speech elicitation. However, future research also needs to focus on L2 speech production elicited through dialogic tasks in more naturalistic contexts, for that approaches language use in real life.

(5) Dimensions of speech production: The present study drew on existing research on L2 speech production for the selection of the dimensions of speech production examined. It followed Skehan's (1996) and Mehnert's (1998) suggestion that language performance be evaluated in terms of the pedagogical goals of fluency, accuracy, complexity, and weighted lexical density. This seems a sound research strategy if we take into consideration that before their studies, research on L2 speech performance was much

more fragmented, focusing on the general notion of "fluency", and that, for an exploratory quantitative study, it is more cautious to rely on tasks and concepts to which there is some background. Nevertheless, there are many other dimensions of speech production that need to be addressed by future research. These include L2 pronunciation and the suprasegmental aspects of L2 speech production--for instance, intonation, stress, prominence--as well as the interactive features that characterize conversation.

(6) Pauses and hesitations: In the present study, again for its exploratory nature, the variables used to asses L2 speech production were taken from previous published studies reporting successful use of these variables. With regard to variables of fluency, one limitation of the present study was the measurement of silent pauses, which was made through the use of a stopwatch and software. Griffith (1991) states that reliable measures of pauses requires timing in milliseconds, perceptual analysis, and detailed evaluation of spectrographic printouts. Thus, although in the present study the location of pauses and hesitations in part of the participants' speech samples was submitted to an interrater analysis, the values obtained can be considered, at best, an approximation. Future research should use more specialized equipment to assess pause length in L2 speech production. More research is also needed to determine criteria for the selection of a cut-off point for silent pauses in L2 speech production. The literature on L2 speech production uses a number of different cut-offs, varying from 0.1 second to 0.5 seconds, the reason for choice generally being related to the instruments the researcher has available to measure pauses. In addition, pauses and hesitations in L2 speech production need to be approached from a more qualitative perspective, in which their function can be assessed. As argued in the present study, L1 speech is full of disruptions. It is possible

that the difference between L1 and L2 pauses and hesitations is in the function they realize in the two languages.

(7) Accuracy, complexity, and weighted lexical density: In the present study, accuracy was assessed in terms of number of errors in syntax, morphology, and lexical choice per hundred words. Complexity and weighted lexical density were assessed in terms of number of dependent clauses per minute and the percentage of weighted lexical items over all linguistic items, respectively. Errors were located on the basis of the prescriptive grammar of English as a native language. The identification of dependent clauses and lexical items followed Mehnert's (1998) and O'Loughlin's (1995) sets of criteria. Although native language performance can serve as a baseline against which to compare L2 performance, evaluation of L2 speech performance should take into consideration the characteristics of learners' production, in the sense that learners will have less information, in general, at their disposal. Thus, criteria for assessing accuracy, complexity, and weighted lexical density should be set taking into account not only native-like performance but also general features of nonnative performance. In addition, these criteria have to be selected according to the characteristics of the oral task. For instance, if the task is too difficult to perform, it is possible that learners will yield simple language because their attention is being shared among higher-level cognitive processes such as reasoning and problem-solving. In this case, the task is unlikely to generate a desirable number of dependent clauses or lexical items that are suitable or relevant for quantitative analysis. Particularly in the case of weighted lexical density, it is questionable the extent to which the method used to distinguish lexical from grammatical items was suitable for the present study. Assessing lexical density in L2 learners' speech

samples is still a problem in the L2 acquisition and use area (Meara, Lightbown, & Halter, 1997). For one thing, the method to determine the weight of lexical items taxes repetition of items. Repetition, as observed by Baptista (2000, personal communication) is an element of cohesion in English. For another, Ortega (1999) notes that the concepts of lexical richness, lexical density, and lexical range are vague. These concepts are also problematic in the sense that L2 learners have idiosyncratic mental lexicons that might have little to do with traditional native-language-driven lexical measures. Finally, measurement of lexical density is affected by the length of the speech sample. Thus, future research needs to focus on how to improve assessment of lexical density, in general, in order for the relationship between working memory capacity and lexical density to be further specified.

(8) Assessment of working memory capacity: In the present study, working memory capacity was assessed by means of the speaking span test in the speakers' L2--English. The results obtained in this study contribute to establishing the speaking span test as a valid measure of working memory capacity. Further research is needed, however, to increase the validity and reliability of the test. For example, the reading span test first proposed by Daneman & Carpenter (1980) is a widely used and much respected measure of working memory capacity whose validity and reliability was constructed through cumulative research by a number of independent researchers interested in language processing. With regard to the operation-word span test, it is imperative that anyone applying this test control for participant's rehearsal to guarantee the quality of data. In the speaking span test, the software has a built-in system that controls the exposure of stimulus on the computer screen. In future studies, the same can be done with the

operation-word span test by determining how long each pair of operation-word will be shown on the screen.

(9) Participants' motivation: As Just & Carpenter (1992) note, performance can be affected by motivational factors. Therefore, it could be that, in the present study, individuals with a higher working memory span were intrinsically motivated to complete the tasks whereas those with a lower working memory span simply did not want to try as hard. Future research can help find ways, especially in countries which lack an experimental research tradition, of controlling for individual motivation in language studies conducted in an artificial manner.

(10) Conceptualization, phonological encoding, articulation in L2 speech production: The present study dealt with only one aspect of L2 speech production--the grammatical encoding of the message. However, since this was attained by giving participants a task in which they had to produce speech at the discourse level, it is clear that the other various processes of speech production were also operating and drawing on their attentional resources. The effects of the conceptualization of the message as well as of articulation were expected to be minimized by giving participants time to plan what to say, in which they also had the chance to check for the pronunciation of words. Nevertheless, only one portion of the whole speech production process was looked at here. Future research needs to address how the relationship between working memory capacity and L2 speech production is affected by the conceptualization of the message, the construction of the phonetic plan, and the articulation of the message as overt speech.

(11) Statistical techniques: In the present study, the statistical techniques used were the Pearson Correlation, Spearman's Rank Order Correlation, and simple linear regression analyses. Although these are valid techniques much used in early research on individual differences in working memory capacity, future research needs to employ more powerful statistical techniques in order to better determine the factors contributing to the relationship between capacity and performance. Recently published studies on working memory capacity have employed structural equation modeling and factorial analyses (Engle, Tuholski, Laughlin, & Conway, 1999; Kane, Bleckley, Conway, & Engle, in press; Kane & Engle, in press; Miyake & Friedman, 1998).

7.3 Pedagogical implications

As stated in the Introduction, this doctoral research project emerged from an interest in learning more about L2 speech production to gain insights on how to best deal with this skill, from both a learner's and teacher's perspective. The results of the present study show that part of the reason L2 speech production is a complex skill to deal with is that it draws heavily on our attentional resources since both L1 and L2 knowledge units are activated from the very beginning of the grammatical encoding process. The dual activation requires suppression of the interference of L1 information and the increase of activation of L2 items. Because the learner's lexicon is generally more limited in the L2 than in the L1, a purposeful and effortful search-and-retrieval activity is added to the process. Finally, the L2 speaker also has to deal with monitoring, to check whether what is being articulated follows the conceptual, lexicogrammatical, and articulatory patterns of the language. Thus, a task that is already complex in itself becomes even more complex when it is performed in the L2.

The present study assessed more directly the speech production processes taking place in the formulator. In Levelt's (1989) model of L1 speech production, the processes of the formulator are highly proceduralized or automatic. It is in this aspect that L2 speech production differs more critically from L1 speech production. Thus, while the present study does not bring evidence for a difference in the level of proceduralized or automatic knowledge among the participants of the present study (that is, it was not possible to determine what is procedural what is not), it seems plausible to suggest that proceduralization contributes to L2 speech production and will possibly affect the distribution of attentional resources, thus likely affecting its relationship with working memory capacity.

Indeed, Towell, Hawkins, and Bazergui (1996) have been able to show that increases in L2 speech production can be accounted for by increases in the degree of proceduralization in the L2 formulator. Towell et al. (1996) showed that gains in the L2 speech production of their participants were noticed in the length and complexity of runs between pauses. That is, production became faster and more complex and, they claim, possibly resulted from the proceduralization of processes in the formulator. Thus, the challenge for the L2 teacher is to provide situations in which learners have the chance to proceduralize L2 knowledge, that is, in which L2 lexicogrammatical information can be easily (rapidly) accessed and retrieved. Proceduralized knowledge minimizes the consumption of attentional resources, which can be directed towards other aspects of the production task.

Now, how can we help learners build procedural knowledge in L2 speech production? One proposal has been made by DeKeyser (1998), in the context of the focus on form debate. Following the work of Anderson (e.g. 1995), DeKeyser suggests that practice allows for the reorganization of knowledge units into chunks of combined elements. Practice leads to increases in speed of access to information and reductions in error rate. Procedural knowledge, thus, emerges as a function of amount of practice. DeKeyser suggests, then, that quality practice leading to procedural knowledge can be promoted in the classroom with communicative drills, in which learners have to convey real meaning, while also focusing on some aspect of the language.

As noted in Chapter 6, there was an interaction among fluency, accuracy, and complexity, and lexical density. The relationship between the speaking span test and L2 speech production was strongest for fluency, followed by complexity, and then accuracy. Of particular interest is the interaction between fluency and complexity. As in previous studies (Mehnert, 1998; Ortega, 1999; Skehan, 1992; Wolfe-Quintero et al.,1998), fluency and complexity seem to be related, with gains in one being accompanied by gains in the other. Thus, it is possible as Wolfe-Quintero et al. (1998) suggest, that fluency and complexity are related constructs. Indeed, Pawley & Syder (1983) point out that native speakers' speech production entails a great number of habitually spoken sentences consisting of memorized clauses and sentences retrieved as chunks. Most times, these sequences are longer than a single, simple clause, they claim. Examples are "there's nothing you can do about it now", "I see what you mean", "You just never know what they'll do next" (p.207). Along the same lines, Levelt (1989) claims that native speakers have whole messages available in long-term memory, these possibly being retrieved as chunks.

In a way, this goes in favor of Towell et al.'s (1996) claim, noted earlier, that what increases in learners' speech production is the length and complexity of runs. The memorized chunks native speakers make use of have some internal complexity and are produced in a run, as a chunk. Thus, one pedagogical suggestion that can be made based on the present study is that, to enhance the proceduralization of knowledge, teachers can create communicative tasks in which these chunks are practiced. In doing this, teachers can raise learners' awareness that native speech is full of memorized sequences, which account for a high proportion of what is said in real life situations. The assumption underlying this pedagogical implication is that, from a cognitive perspective, being able to retrieve chunks as opposed to retrieving word by word, the speaker's attentional resources are free for the cognitive mechanisms of speaking instead of being busy dealing with linguistic elements. The use of communicative tasks for the practice of these chunks might involve a focus on form, as DeKeyser suggests, and can be complemented by the task-repetition approach advocated by Bygate (1996), who has shown that performance of an oral task on a second occasion leads to better fluency, accuracy, and vocabulary choice.

The idea that memorizing sequences enhances oral performance seems intuitively appealing, at least from the point of view of the participants of the present study. In informal talks with these learners, it became apparent that they found their grammar classes more effective than the speaking classes as regards the improvement of their oral performance. When asked why they had this feeling, their usual answers would be that in the grammar classes they were learning "how to say it", and would give examples such as "this is the first time I've ever...", "how long does it take to go from here to .... by ....?" or "I'd love to...but I can't because...", produced in a run. Their feeling that grammar classes enhanced their oral performance might be related to the perception that, in addition to the grammar knowledge they were gaining, they were also learning whole sequences that could be accessed, retrieved, and produced as one unit, and not word by word.

Another pedagogical implication comes from the research on working memory capacity. The argument was raised in the present study that working memory is capable of allocating resources when the task exceeds its capacity. Following Just & Carpenter (1989), it might be that individual differences in working memory capacity are at least partly related to the effectiveness with which individuals allocate their resources. In this case, it might be possible to modify the relationship between working memory capacity and performance on complex cognitive tasks, including L2 speech production, by raising learners' awareness about how to strategically allocate resources. It might be possible to tell them that L2 speech performance has four salient aspects--fluency, accuracy, complexity, and lexical density--and that by using memorized chunks they gain in fluency and complexity and are thus able to focus on accuracy, for instance.

The objective of the present study was to investigate the relationship between working memory capacity and L2 speech production. Due to the sample size, the results cannot be generalized to the population of learners investigated. These results can be taken, thus, as only suggestive of a trend between capacity and production. Further research is necessary to understand better L2 speech performance. One way to do that is from the perspective of individual differences in working memory capacity. This perspective seems a promising one. As Perlow, Jattuso, & De Wayne Moore (1997) point out, one of the objectives of contemporary cognitive science is to explain how human beings learn and perform complex activities and why there is variance in performance. Working memory, one of the most intensively studied areas in contemporary cognitive psychology and cognitive neuroscience (Miyake & Shah, 1999), is at the heart of complex behavior and has been shown to be a source of individual differences in learning and performance of complex cognitive tasks (Baddeley, 1999: Daneman & Carpenter, 1980 & 1983; Shute, 1991). In addition, it seems to be a growing tendency in our contemporary society to require multiskilled individuals for more and more complex and cognitively demanding jobs (Howell & Cook, 1989). In this process, being able to perform in an L2 has become a necessary skill in many professional areas and speaking is, most times, the skill chosen by evaluators and recruiters as representative of L2 performance. While a much greater effort needs to be made until we can fully grasp the complexities of this skill, it is hoped that the present work constitutes a step towards understanding L2 speech production.

REFERENCES

AKMAJIAN, A., DEMERS, R.A, FARMER, A.K, HARNISH, R.M. (1995). *Linguistics: an introduction to language and communication.* Cambridge, MA: The MIT Press.

ANDERSON, J.R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

ANDERSON, J.R. (1995). *Learning and memory.* Singapore: Wiley.

AREVART, S., & NATION, P. (1991). Fluency improvement in a second language. *RELC Journal, 22,* 84-94.

ASHCRAFT, M.H. (1994). *Human memory and cognition.* New York: Harper Collins.

ATKINS, P.B., & BADDELEY, A.D. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics, 19,* 537-552.

ATKINSON, R.C., & SHIFFRIN, R.M. (1968). Human memory: A proposed system and its control processes. In K.W. Spencer (Ed.), *The psychology of learning and motivation: advances in research and theory* (Vol. 2, pp. 89-195). New York: Academic Press.

ATKINSON, R.C., & SHIFFRIN, R.M. (1971). The control of short-term memory. *Scientific American,* 225, 82-90.

BACHMAN, L.F., & PALMER, A.S. (1996). Language testing in practice: designing and developing useful language tests. Oxford: Oxford University Press.

BADDELEY, A.D. (1966a). Short-term memory for word sequences as a function of acoustic, semantic and formal similarity. *Quarterly Journal of Experimental Psychology, 18,* 362-365.

BADDELEY, A.D. (1966b). The influence of acoustic and semantic similarity on long-term memory for word sequences. *Quarterly Journal of Experimental Psychology, 18,* 302-309.

BADDELEY, A.D. (1981). The concept of working memory: a view of its current state and probable future development. *Cognition, 10,* 17-23.

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

BADDELEY, A.D. (1990). *Human memory: Theory and Practice.* Hove, UK: Lawrence Erlbaum.

BADDELEY, A.D. (1992a). Working memory. *Science, 255,* 556-559.

BADDELEY, A.D. (1992b). Is working memory working? The fifteenth Bartlett lecture. *The Quarterly Journal of Experimental Psychology, 44,* 1-31.

BADDELEY, A.D. (1992c). Working memory: the interface between memory and cognition. *Journal of Cognitive Neuroscience, 4,* 281-288.

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

BADDELEY, A.D. (1999). *Essentials of human memory.* East Sussex: Psychology Press.

BADDELEY, A.D. & HITCH, G. (1974). Working memory. In G. A. Bower (Ed*.), The psychology of learning and motivation: advances in research and theory* (Vol. 8, pp. 47-90). New York: Academic Press.

BADDELEY, A.D., & LOGIE, R.H. (1999). Working memory: the multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp.28-61). Cambridge: Cambridge University Press.

BADDELEY, A.D., PAPAGNO, C., & VALLAR, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language, 27,* 586-595.

BADDELEY, A.D., & WARRINGTON, E.K. (1970). Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior, 9,* 176-189.

BADDELEY, A.D., & WILSON, B. (1988). Comprehension and working memory: a single case neuropsychological study. *Journal of Memory and Language, 27,* 479-498.

BENTON, S.L., KRAFT, R.G., GLOVER, J.A., & PLAKE, B.S. (1984). Cognitive capacity differences among writers. *Journal of Educational Psychology, 76,* 820-834.

BERQUIST, B. (1998). *Individual differences in working memory span and L2 proficiency: capacity or processing efficiency?* Paper presented at the American Association for Applied Linguistics 1998 Annual Conference, Seattle, WA.

BOCK, K. (1996). Language production: methods and methodologies. *Psychonomic Bulletin & Review, 3,* 395-421.

BROADBENT, D.E. (1958). *Perception and communication*. London: Pergamon Press.

BROWN, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10,* 12-21.

BRUMFIT, C., & JOHNSON, K. (1979). Communicative methodology in language teaching. Cambridge: Cambridge University Press.

BUTLER-WALL, B.A. (1986). *The frequency and function of disfluencies in native and non-native conversational discourse*. Unpublished doctoral dissertation, University of California at Los Angeles.

BYGATE, M. (1996). Effects of task repetition: appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and Change* (pp.136-146). Oxford: Heinemann.

CANTOR, J., & ENGLE, R. (1993). Working memory capacity as long-term activation: an individual differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1101-1114.

CANTOR, J., ENGLE, R., & HAMILTON, G. Short-term memory, working memory, and verbal abilities: How do they relate? *Intelligence, 15,* 229-246.

CARPENTER, P.A., & JUST, M.A. (1989). The role of working memory in language comprehension. In D. Klahr & K. Kotovsky, K. (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp.31-68). Hillsdale, N.J.: Lawrence Erlbaum.

CARPENTER, P.A., & JUST, M.A. (1999). Computational modeling of high-level cognition versus hypothesis testing. In R.J. Sternberg (Ed.), *The nature of cognition* (pp. 245-293). Cambridge, MA: The MIT Press.

CARPENTER, P.A., MIYAKE, A., & JUST, M.A. (1994). Working memory constraints in comprehension: Evidence from individual differences, aphasia, and aging. In M. Gernsbacher (Ed.), *The handbook of psycholinguistics* (pp.1075-1122). San Diego: Academic Press.

CARROL, J.B. (1975). *The teaching of French as a foreign language in eight countries*. New York: John Wiley & Sons.

CHAFE, W. (1980). The deployment of consciousness in the production of a narrative. In W. Chafe, (Ed.), *The pear stories: cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex.

CHAFE, W. (1985). *Some reasons for hesitating*. In D. Tannen & M. Saville-Troike (Eds.), *Perspectives on silence* (pp.77-89). Norwood, NJ: Ablex.

CHAFE, W. (1998). Language and the flow of though. In M. Tomasello (Ed.), *The new psychology of language: cognitive and functional approaches to language structure*. Mahwah, NJ: Lawrence Erlbaum.

CLARK, H. (1996). *Using language*. Cambridge: Cambridge University Press.

CLARK, H.H., & CLARK, E.V. (1977). *Psychology and language*. New York: Harcourt, Brace, Jovanovich.

CLARKSON-SMITH, L., & HARTLEY, A.A. (1990). The game of bridge as an exercise in working memory and reasoning. *Journal of Gerontology: Psychological Sciences, 45,* 233-238.

CONRAD, R., & HULL, A.J. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, 55, 429-432.

CONWAY, A.R.A, & ENGLE, R.W. (1996). Individual differences in working memory capacity: more evidence for a general capacity theory. *Memory, 4,* 577-590.

COWAN, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. Psychological Bulletin, 104, 163-191.

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

CRAIK, F.I.M., & WATKINS, M.J. (1972). Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

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

CROOKES, G. (1991). Second language speech production research: a methodologically-oriented review. *Studies in Second Language Acquisition, 13*, 113-132.

CROOKES, G. & GASS, S. M. (Eds.). (1993a). *Tasks and language learning*. Clevedon: Multilingual Matters.

CROOKES, G. & GASS, S. M. (Eds.). (1993b). *Tasks in a pedagogical context*. Clevedon: Multilingual Matters.

DANEMAN, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research, 20*, 445-464.

DANEMAN, M., & CARPENTER, P. A. (1980). Individual differences in working memory and reading. *Journal or verbal learning and verbal behavior, 19*, 450-466.

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

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

DANEMAN, M., & MERIKLE, P.M. (1996). working memory and language comprehension: a meta-analysis. *Psychonomic Bulletin & Review, 3*, 422-433.

DANEMAN, M. & TARDIFF, T. (1987). Working Memory and reading skill re examined. In M. Coltheart (Ed.), *Attention & performance* (Vol. 12, pp. 491-508). Hillsdale, NJ: Lawrence Erlbaum.

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

De BOT, K., COX, A., RALSTON, S., SCHAUFELI, A., WELTENS, B. (1995). Lexical processing in bilinguals. *Second Language Research*, 11, 1-19.

De BOT, & SCHREUDER, R. (1993). Word production and the bilingual lexicon. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 191-214). Amsterdam/Philadelphia: John Benjamins.

DECHERT, H.W. (1983). How a story is done in a second language. In C. Faerch & g. Kasper (Eds.), Strategies in interlanguage communication (pp. 175-195). London: Longman.

DECHERT, H.W. (1984). Second language production: six hypotheses. In H.W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp.211-230). Tübingen: Gunter Narr Verlag.

DECHERT, H.W. (1987). Understanding producing. In H.W. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 229-237). Norwood: Ablex.

DECHERT, H.W., MÖHLE, D., & RAUPACH, M. (Eds.). (1984). Second language productions. Tübingen: Gunter Narr Verlag.

DECHERT, H.W., & RAUPACH, M. (Eds.). (1980). *Towards a cross-linguistic assessment of speech production*. Frankfurt: Lang.

DECHERT, H.W., & RAUPACH, M. (Eds.). (1987). *Psycholinguistic models of production*. Norwood: Ablex.

DEESE, J. (1980). Pauses, prosody, and the demands of production in language. In H.W. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 69-84). The Hague: Mouton Publishers.

DELL, G. (1986). A spreading-activation theory or retrieval in sentence production. *Psychological Review, 3*, 283-321.

DELL, G., JULIANO, C., & GOVINDJEE, A. (1993). Structure and content in language production: a theory of frame constraints in phonological errors. *Cognitive Science*, *17*, 149-195.
DESCHAMPS, A. (1980). The syntactical distribution of pauses in English spoken by French students. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 255-262). The Hague: Mouton Publishers.
DIXON, P., LeFEVRE, J., & TWILLEY, L.C. (1988). Word knowledge and working memory as predictors of reading skill. *Journal of Educational Psychology*, *80*, 465-472.
DÖRNYEI, Z., & KORMOS, J. (1998). Problem-solving mechanisms in L2 communication: a psycholinguistic perspective. *Studies in Second Language Acquisition*, *20*, 349-385.
DOUGHTY, C., & PICA, T. (1986). Information gap tasks: Do they facilitate second language acquisition? *TESOL Quarterly*, *20*, 305-325.
DUFF, P. (1986). Another look at interlanguage talk: taking task to task. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 147-181). Rowley, MA: Newbury House.
EJZENBERG, R. (1992). *Understanding nonnative oral fluency: The role of task structure and discourse variability*. Unpublished doctoral dissertation, State University of New York.
EJZENBERG, R. (1995). *Discourse markers and context-dependency in oral fluency assessment*. Paper presented at the American Association for Applied Linguistics 1995 Annual Conference, Long Beach, CA.
ELLIS, R. (1985). Sources of variability in interlanguage. *Applied Linguistics*, *6*, 118-131.
ELLIS, R. (1987). Interlanguage variability in narrative discourse: style shifting in the use of the past tense. *Studies in Second Language Acquisition*, *9*, 1-20.
ELLIS, R. (1993). The structural syllabus and second language acquisition. *TESOL Quarterly*, *27*, 91-113.
ENGLE, R.W. (1996). Working memory and retrieval: an inhibition-resource approach. In J.T.E. Richardson, R.W. Engle, L. Hasher, R.H. Logie, E.R. Stoltzfus, & R.T. Zacks (Eds.), *Working memory and human cognition* (pp. 89-119). Oxford: Oxford University Press.
ENGLE, R.W., CANTOR, J., CARULLO, J.J. (1992). Individual differences in working memory and comprehension: a test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 972-992.
ENGLE, R.W., CARULLO, J.J., & COLLINS, K.W. (1991). Individual differences in the role of working memory in comprehension and following directions in children. *Journal of Educational Research*, *84*, 253-262.
ENGLE, R.W., & CONWAY, A.R.A. (1998). Working memory and comprehension. In R.H. Logie & Gilhooly (Eds.), *Working memory and thinking* (pp.67-91). East Sussex: Psychology Press.
ENGLE, R.W., KANE, M.J., & TUHOLSKI, S.W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp.102-134). Cambridge: Cambridge University Press.
ENGLE, R.W., NATIONS, J.K., & CANTOR, J. (1990). Is working memory capacity just another name for word knowledge? *Journal of Educational Psychology*, *82*, 799-804.
ENGLE, R.W., & ORANSKY, N. (1999). Multi-store versus dynamic models of temporary storage in memory. In R.J. Sternberg (Ed.), *The nature of cognition* (pp.515-555). Cambridge, MA: The MIT Press.
ENGLE, R.W, TUHOLSKY, S.W., LAUGHLIN, J.E., CONWAY, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309-331.
ERICSSON, K. A., & DELANEY, P. F. (1998). Working memory and expert performance. In R. Logie & K. Gilhooly (Eds.), *Working memory and thinking* (pp.93-114). East Sussex: Psychology Press.

ERICSSON, K. A., & DELANEY, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp.257-297). Cambridge: Cambridge University Press.

ERICSSON, K. A., & KINTSCH, W. (1995). Long-term working memory. *Psychological Review, 102*, 211-245.

FAERCH, C., & KASPER, G. (1983). Plans and strategies in foreign language communication, In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication* (pp.20-60). London: Longman.

FILLMORE, C. (1979). On fluency. In C. Fillmore, D. Kempler, & w. Wang (Eds.), *Individual differences in language ability and language behavior* (pp.85-102). New York: Academic Press.

FITZGERALD, F.S. (1926). *The Great Gatsby*. Middlesex: Penguin Modern Books.

FORTKAMP, M.B.M. (1998). Measures of working memory capacity and L2 oral fluency. *Ilha do Desterro, 35*, 201-238.

FORTKAMP, M.B.M. (1999). Working memory capacity and aspects of L2 speech production. *Communication & Cognition, 32*, 259-296.

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

FREED, B. (1995). What makes us think that students who study abroad become fluent? In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123-148). Amsterdam/Philadelphia: John Benjamins.

GARRET, M.F. (1982). Production of speech: observations from normal and pathological language use. In A. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19-76). London: Academic Press.

GASS, S., & VARONIS, E. (1985). Variation in native speaker speech modification to non-native speakers. *Studies in Second Language Acquisition, 7*, 37-57.

GATBONTON, E., & SEGALOWITZ, N. (1988). Creative automatization: principles for promoting fluency within a communicative framework. *TESOL Quarterly, 22*, 473-492.

GATHERCOLE, S.E., & BADDELEY, A.D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: a longitudinal study. *Journal of Memory and Language, 28*, 200-213.

GATHERCOLE, S.E., & BADDELEY, A.D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*, 336-360.

GATHERCOLE, S.E., & BADDELEY, A.D. (1993). *Working memory and language*. Hillsdale: Lawrence Erlbaum.

GILHOOLY, K.J. (1998). Working memory, strategies, and reasoning tasks. In R.H. Logie & Gilhooly (Eds.), *Working memory and thinking* (pp. 7-22). East Sussex: Psychology Press.

GOLDMAN-EISLER, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. New York: Academic Press.

GOLDMAN-RAKIC, P.S. (1992). Working memory and the mind. *Scientific American, 267*, 72-79.

GOODWIN, C.J. (1995). *Research in psychology: methods and design*. New York: John Wiley & Sons.

GRAF, P., & SCHACTER, D.L. (1985). Implicit and explicit memory for new association in normal and amnesic subjects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 501-518.

GRAF, P., & SCHACTER, D.L. (1987). Selective effects of interference on implicit and explicit memory for new associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 45-53.

GREEN, D.W. (1986). Control, activation and resource: a framework and a model for the control of speech in bilinguals. *Brain and Language, 27*, 210-223.

GREEN, D.W. (1998). Mental control of the bilingual lexico-semantic system. Bilingualism: *Language and Cognition*, 1, 67-81.

GREENE, J., & D'OLIVEIRA, M. (1982). *Learning to use statistical tests in psychology: a student's guide*. Buckingham: Open University Press.

GRIFFIN, Z.M., & BOCK, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language, 38*, 313-338.

GRIFFITHS, R. (1991). Pausological research in an L2 context: a rationale, and review of selected studies. *Applied Linguistics, 12*, 345-364.

GUILLOT, M.-N. (1999). *Fluency and its teaching*. Clevedon: Multilingual Matters.

HABERLANDT, K. (1994). *Cognitive psychology*. Massachusetts: Allyn and Bacon.

HARRINGTON, M. (1992). Working memory capacity as a constraint on L2 development. In R.J. Harris (Ed.), *Cognitive processing in bilinguals* (pp.123-135). Amsterdam: Elsevier.

HARRINGTON, M., & SAWYER, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14*, 25-38.

HATCH, E., & FARHADY, H. (1982). *Research design and statistics for applied linguistics*. Cambridge, MA: Newbury House.

HIEKE, A.E. (1985). A componential approach to oral fluency evaluation. *The Modern Language Journal, 69*, 135-142.

HOWELL, W.C., & COOKE, N.J. (1989). Training the human information processor: A review of cognitive models. In I.L. Goldstein (Ed.), *Training and development in organizations* (pp.121-182). San Francisco: Jossey-Bass.

JEFFERSON, G. (1979). Error correction as an interactional resource. *Language and Society, 3*, 181-199.

JURDEN, F.H. (1995). Individual differences in working memory and complex cognition. *Journal of Educational Psychology, 87*, 93-102.

JUST, M.A., & CARPENTER, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological Review, 99*, 122-149.

JUST, M.A., CARPENTER, P.A., & HEMPHILL, D.D. (1996). Constraints on processing capacity: Architectural or implementational? In D. Steier & T.M. Mitchell (Eds.), *Mind matters: a tribute to Allen Newell* (pp.141-178). Mahwah, NJ: Lawrence Erlbaum.

KANE, MJ., BLECKLEY, M.K., CONWAY, A.R.A., ENGLE, R.W. (in press). A controlled-attention view of working memory capacity. *Journal of Experimental Psychology: General*.

KANE, M.J., & ENGLE, R.W. (in press). Working memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

KIEWRA, K.A., & BENTON, S.L. (1988). The relationship between information-processing ability and notetaking. *Contemporary Educational Psychology, 13*, 33-44.

KING, J., JUST, M.A. (1991). Individual differences in syntactic processing: the role of working memory. *Journal of Memory and Language, 30*, 580-602.

KINTSCH, W., & van DIJK, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review, 93*, 163-182.

KINTSCH, W., HEALY, A., HEGARTY, M., PENNINGTON, B., & SALTHOUSE, T. (1999). Models of working memory: eight questions and some general issues. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp.412-441). Cambridge: Cambridge University Press.

KLEIN, K., & FISS, W. (1999). The reliability and stability of the Turner and Engle working memory task. *Behavior Research Methods, Instruments, & Computers, 31*, 429-432.

KROLL, J.K. (1993). Accessing conceptual representations for words in a second language. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 53-81). Amsterdam/Philadelphia: John Benjamins.

KYLLONEN, P.C., & CHRISTAL, R.E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*, 389-433.

LA POINTE, L.B., & ENGLE, R. W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 6, 1118-1133.

LENNON, P. (1984). Retelling a story in English as a second language. In H.W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp.50-68). Tübingen: Narr.

LENNON, P. (1990). Investigating fluency in EFL: a quantitative approach. *Language Learning*, *40*, 387-417.

LEVELT, W.J.M. (1989). *Speaking. From intention to articulation*. Cambridge, MA: Bradford/MIT Press.

LEVELT, W.J.M. (1992). Accessing words in speech production: stages, processes and representations. *Cognition*, 42, 1-22.

LEVELT, W.J.M. (1995). The ability to speak: from intentions to spoken words. *European Review*, *3*, 13-23.

LEVELT, W.J.M., ROELOFS, A, & MEYER, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1-38.

LOGIE, R.H. (1986). Visuo-spatial processing in working memory. *Quarterly Journal of Experimental Psychology*, *38A*, 229-247.

LOGIE, R.H. (1995). Visuo-spatial working memory. Hove: Lawrence Erlbaum.

LOGIE, R.H. (1996). The seven ages of working memory. In J.T.E. Richardson, R.W. Engle, L. Hasher, R.H. Logie, E.R. Stoltzfus, & R.T. Zacks (Eds.), *Working memory and human cognition* (pp. 31-65). Oxford: Oxford University Press.

LOGIE, R.H., ZUCCO, G., & BADDELEY, A.D. (1990). Interference with visual short-term memory. *Acta Psychologica*, *75*, 55-74.

LONG, M., & CROOKES, G. (1991). Three approaches to task-based syllabus design. *TESOL Quarterly*, *26*, 27-55.

LUCENA, M.I.P., FORTKAMP, M.B.M, BRAGA, M.C., & ALMEIDA, M.S. (2000). *EFL language learning strategies: teaching beginners*. Paper presented at the III Convention of the Associations of English Teachers of the States of Rio Grande do Sul (APIRS), Santa Catarina (APLISC), and Paraná (APLIEPAR).

MacDONALD, M.C., JUST, M.A., & CARPENTER, P.A. (1992). Working memory constraints on the processing of syntactic ambiguity. *Cognitive Psychology*, *24*, 56-98.

MASSON, M.E.J., & MILLER, J.A. (1983). Working memory and individual differences in comprehension and memory of text. *Journal of Educational Psychology*, *75*, 314-318.

McCLAVE, J. T., DIETRICH, F.H., SINCICH, T. (1997). *Statistics*. Upper Saddle River, NJ: Prentice Hall.

McLAUGHLIN, B. (1987). *Theories of second language learning*. London: Edward Arnold.

McLAUGHLIN, B. (1998). Second language learning revisited: the psycholinguistic perspective. In A.F. Healy & Bourne, L.E., Jr. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp.399-411). Mahwah, NJ: Lawrence Erlbaum.

McLAUGHLIN, B., & HEREDIA, R. (1996). Information-processing approaches to research on second language acquisition and use. In W.C. Ritchie & T.K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 231-228). San Diego: Academic Press.

McLAUGHLIN, B., ROSSMAN, T., & McLEOD, B. (1983). Second-language learning: an information processing perspective. *Language Learning*, *33*, 135-158.

MEARA, P., LIGHTBOWN, P.M., HALTER, R. H. (1997). Classrooms as lexical environments. *Language Teaching Research*, *1*, 28-47.

MEHNERT, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, *20*, 83-108.

MESSICK, S. (1996). Human abilities and modes of attention: the issue of stylistic consistencies in cognition. In I. Dennis & P. Tapsfield (Eds.), *Human abilities: their nature and measurement* (pp. 77-96).Mahwah, NJ: Lawrence Erlbaum.

MEYER, A.S. (1996). Lexical access in phrase and sentence production: results from picture-word interference experiments. *Journal of Memory and Language, 35*, 477-496.

MILLER, G.A., GALANTER, E., & PROBRAM., K.H. (1960). *Plans and the structure of behavior*. New York: Holt.

MIYAKE, A., & FRIEDMAN, N.P. (1998). Individual differences in second language proficiency: working memory as language aptitude. In A.F. Healy & Bourne, L.E., Jr. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp.339-364). Mahwah, NJ: Lawrence Erlbaum.

MIYAKE, A., JUST, M.A., & CARPENTER, P. A. (1994). Working memory constraints on the resolution of lexical ambiguity: maintaining multiple interpretations in neutral contexts. *Journal of Memory and Language, 33*, 175-202.

MIYAKE, A, & SHAH, P. (1999). Toward unified theories of working memory: emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp.442-481). Cambridge: Cambridge University Press.

MÖHLE, D. (1984). A comparison of the second language speech production of different native speakers. In H.W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp.26-49). Tübingen: Narr.

MOTA, M.B. (1995). *Working memory capacity and L2 fluent speech production*. Unpublished master's thesis, Universidade Federal de Santa Catarina, Florianopolis.

MOTLEY, M.T., & BAARS, B.J. (1976). Laboratory induction of verbal slips: a new method for psycholinguistic research. *Communication Quarterly, 24*, 28-34.

NELSON, L.K., & BAUER, H.R. (1991). Speech and language production at age 2: evidence for tradeoffs between linguistic and phonetic processing. *Journal of Speech and Hearing Research, 34*, 879-892.

NORMAN, D., & SHALLICE, T. (1986). Attention to action: willed and automatic control of behavior. In R.J. Davidson, G.E. Schwartz, & D.E. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1-18). New York: Plenum.

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

OLYNYK, M., SANKOFF, D., & d'ANGLEJAN, A. (1983). Second language fluency and the subjective evaluation of officer cadets in a military college. *Studies in Second Language Acquisition*, 5, 213-249.

OLYNYK, M., d'ANGLEJAN, A., & SNAKOFF, D. (1990). A quantitative and qualitative analysis of speech markers in the native and second language speech of bilinguals. In R. Scarcella, E. Andersen ,& S. Krashen (Eds.), *Developing communicative competence in a second language* (pp.139-155).. New York: Newbury House.

O'LOUGHLIN, K. (1995). Lexical density in candidate output on direct and semi-direct versions of an oral proficiency test. *Language Testing, 12*, 217-237.

ORMROD, J.E., & COCHRAN, K.F. (1988). Relationship of verbal ability and working memory to spelling achievement and learning to spell. *Reading Research and Instruction, 28*, 33-43.

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

PAPAGNO, C., VALENTINE, T., & BADDELEY, A.D. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory & Language, 30*, 331-347.

PARADIS, M. (1985). On the representation of two languages in one brain. *Language Science*, 7, 1-39.

PAWLEY, A., & SYDER, F. (1983). Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In J.C. Richards & R. Schmidt (Eds.), *Language and communication* (pp. 191-225). London: Longman.

PERFETTI, C.A., & LESGOLD, A.M. (1977). Discourse comprehension and sources of individual differences. In M.A. Just & P.A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 141-183). Hillsdale: Lawrence Erlbaum.

PERLOW, R., JATTUSO, De WAYNE MOORE, D. (1997). Role of verbal working memory in complex skill acquisition. *Human Performance, 10,* 283-302.

PETERSON, L.R., & PETERSON, M.J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology,* 58, 193-198.

POULISSE, N. (1997). Language production in bilinguals. In A.M.B. de Groot, & J.K Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp.201-225). Mahwah, NJ: Lawrence Erlbaum.

POULISSE, N. (1999). *Slips of the tongue: speech errors in first and second language production.* Amsterdam: John Benjamins.

POULISSE, N., & BONGAERTS, T. (1994). First language use in second language production. *Applied Linguistics, 15,* 36-57.

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

QUIRK, R., & GREENBAUM, S. (1973). *A university grammar of English.* Essex: Longman.

RATNER, R.B. (2000). Elicited imitation and other methods for the analysis of trade-offs between speech and language skills in children. In L. Menn & N. B. Ratner, *Methods for studying language production* (pp.291-311). Mahwah, NJ: Lawrence Erlbaum.

RATNER, R.B., & MENN, L. (2000). In the beginning was the wug: forty years of language-elicitation studies. In L. Menn & N. B. Ratner, *Methods for studying language production* (pp.1-23). Mahwah, NJ: Lawrence Erlbaum.

RAUPACH, M. (1980). Temporal variables in first and second language speech production. In H.W. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 263-270). The Hague: Mouton.

RAUPACH, M. (1984). Formulae in second language speech production. In H.W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp.114-137). Tübingen: Gunter Narr Verlag.

RICHARDS, J.C., PLATT, J., & PLATT, H. (1993). *Dictionary of language teaching an applied linguistics* (3rd Ed.). Essex: Longman.

RICHARDSON, J.T.E. (1993). Commentary: developing the model of working memory. In G.M. Davies & R.H. Logie (Eds.), *Memory in everyday life* (pp. 219-173-218). Amsterdam: Elsevier.

RICHARDSON, J.T.E. (1996a). Evolving concepts of working memory. In J.T.E. Richardson, R.W. Engle, L. Hasher, R.H. Logie, E.R. Stoltzfus, & R.T. Zacks (Eds.), *Working memory and human cognition* (pp. 3-30). Oxford: Oxford University Press.

RICHARDSON, J.T.E. (1996b). Evolving issues in working memory. In J.T.E. Richardson, R.W. Engle, L. Hasher, R.H. Logie, E.R. Stoltzfus, & R.T. Zacks (Eds.), *Working memory and human cognition* (pp. 120-154). Oxford: Oxford University Press.

RIGGENBACH, H. (1989). Nonnative fluency in dialogue versus monologue speech: a microanalytic approach. Unpublished doctoral dissertation, University of California at Los Angeles.

RIGGENBACH, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes,* 14, 423-441.

ROBERTS, R., & GIBSON, E. (1999). *Individual differences in sentence memory.* Unpublished manuscript, Massachusetts Institute of Technology.

ROHDE, L. (1985). Compensatory fluency: a study of spoken English produced by four Danish learners. In E. Glahn & A. Holmen (Eds.), *Learner Discourse (Anglica et Americana)* (pp. 43-69). University of Copenhagen, Department of English.

ROSEN, V.M., & ENGLE, R. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General, 126*, 3, 211-227.

ROSEN, V.M., & ENGLE, R. (1997). Working memory capacity and suppression. *Journal of Memory and Language, 39*, 418-436.

RYLE, G. (1949). *The concept of mind*. London: Hutchinson.

SAARILUOMA, P. (1998). Adversary problem-solving and working memory. In R. Logie & K. Gilhooly (Eds.), *Working memory and thinking* (pp.115-138). East Sussex: Psychology Press.

SALAMÉ, P., & BADDELEY, A.D. (1982). Disruption of short-term memory by unattended speech: implications for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior, 21*, 150-164.

SCHACTER, D.L. (1987). Implicit memory: history and current status. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 501-518.

SCHMIDT, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14*, 423-41.

SEARLEMAN, A., & HERMANN, D. (1994). *Memory from a broader perspective*. New York: McGraw-Hill.

SEGALOWITZ, S.J, SEGALOWITZ, N., WOOD, A. (1998). Assessing the development of automaticity in second language recognition. *Applied Psycholinguistics*, 19, 53-67.

SHAH, P., & MIYAKE, A. (1999). Models of working memory: an introduction. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp.1-27). Cambridge: Cambridge University Press.

SHALLICE, T., & WARRINGTON, E.K. (1970). Independent functioning of the verbal memory stores: a neuropsychological study. *Quarterly Journal of Experimental Psychology, 22*, 261-273.

SHIFFRIN, R.M. (1997). Attention, automatism, and consciousness. In J. Cohen & J. Schooler (Eds.), *Scientific approaches to consciousness* (pp. 49-64). Mahwah, NH: Lawrence Erlbaum.

SHIFFRIN, R.M., & DUMAIS, S. (1981). The development of automatism. In J.R. Anderson (Ed.), *Cognitive skills and their acquisition* (pp.111-140). Hillsdale: Lawrence Erlbaum.

SHOHAMY, E. (1988). A proposed framework for testing the oral language of second/foreign language learners. *Studies in Second Language Acquisition, 10*, 165-179.

SHOHAMY, E. (1994). The validity of direct versus semi-direct oral tests. *Language Testing, 11*, 99-124.

SHUTE, V.J. (1991). Who is likely to acquire computer programming skills? *Journal of Educational Computing Research, 7*, 1-24.

SKEHAN, P. (1989). *Individual differences in second language learning*. London: Edward Arnold.

SKEHAN, P. (1992). Second language acquisition strategies and task-based learning. *Thames Valley University Working Papers in English Language Teaching*, 1, 178-208.

SKEHAN, P. (1996a). Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.), *Challenge and Change* (pp.17-30). Oxford: Heinemann.

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

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

SMYTH, M.M., COLLINS, A.F., MORRIS, P.E., LERRY, P. (1994). *Cognition in action*. Hillsdale: Lawrence Erlbaum.

STERN, H.H. (1983). *Fundamental concepts of language teaching*. Oxford: Oxford University Press.

STILINGS, N., FEINSTEIN, M., GARFIELD. J., RISSLAND, E., ROSENBAUM, D., WEISLER, S., & BAKER-WARD, L. (1987). *Cognitive science: an introduction*. Cambridge, MA: the MIT Press.

STOLTZFUS, E.R., HASHER, L., ZACKS, R.T. (1996). Working memory and aging: current status of the inhibitory view. In J.T.E. Richardson, R.W. Engle, L. Hasher, R.H. Logie, E.R.

Stoltzfus, & R.T. Zacks (Eds.), *Working memory and human cognition* (pp. 66-88). Oxford: Oxford University Press.

SWANSON, H.L. (1993). Individual differences in working memory: a model testing and subgroup analysis of learning-disabled and skilled readers. *Intelligence, 17,* 285-332.

TANNENBAUM, P.H., WILLIMA, F.; & HILLER, C.S. (1965). Word predictability in the environment of hesitations. *Journal of Verbal Learning and Verbal Behavior, 4,* 134-140.

TEMPLE, L. (1992). Disfluencies in learner speech. *Australian Review of Applied Linguistics,* 15, 29-44.

TEMPLE, L. (1997). Memory and processing modes in language learner speech production. *Communication & Cognition, 30,* 75-90.

TIRRE, W.C., & PENA, C.M. (1992). Investigation of functional working memory in the reading span test. *Journal of Educational Psychology, 84,* 462-472.

TOMITCH, L.M.B. (1995). *Reading: text organization perception and working memory capacity.* Unpublished doctoral dissertation, Universidade Federal de Santa Catarina, Florianopolis.

TORRES, A.C.G. (1998). *Prior knowledge, L2 working memory capacity, and L2 reading comprehension: How do they relate?* Unpublished master's thesis, Universidade Federal de Santa Catarina, Florianópolis.

TOWELL, R., HAWKINS, R., & BAZERGUI, N. (1996). The development of fluency in advanced learners of English. *Applied Linguistics, 17,* 84-119.

TOWSE, J.N., & HITCH, G.J. (1995). Is there a relationship between task demand and storage space in tests of working memory capacity? *Quarterly Journal of Experimental Psychology, 48A,* 108-124.

TULVING, E. (1966). Subjective organization and effects of repetition in multi-trial free-recall learning. *Journal of Verbal learning Verbal Behavior, 5,* 193-197.

TULVING, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (eds.), *Organization of memory* (pp.381-403). New York: Academic Press.

TULVING, E. (1983). *Elements of episodic memory.* Oxford: Oxford University Press.

TULVING, E. (1985). How many memory systems are there? *American Psychologist, 40,* 385-398.

TURNER, M. & ENGLE, R. (1989). Is working memory capacity task-dependent? *Journal of Memory and Language, 28,* 127-154.

VanPATTEN, B. (1994). Evaluating the role of consciousness in SLA: terms, linguistic features, and research methodology. *AILA Review, 11,* 27-36.

VanPATTEN, B. (1996). *Input processing and grammar instruction in second language acquisition.* Norwood, N.J.: Ablex.

WALSH, A. (1990). *Statistics for the social sciences: with computer applications.* New York: Harper & Row.

WIESE, R. (1984). Language production in foreign and native languages: same or different? In H.W. Dechert, D. Möhle, & M. Raupach (Eds.), *Second language productions* (pp.11-25). Tübingen: Gunter Narr Verlag.

WATERS, G.S., & CAPLAN, D. (1996). The capacity theory of sentence comprehension: critique of Just and Carpenter (1992). *Psychological Review,* 103, 761-772.

WOLFE-QUINTERO, K., INAGAKI, S., & KIM, H.-Y. (1998). *Second language development in writing: measures of fluency, accuracy, and complexity* (Technical Report #17). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.

WOLTZ, D.J. (1988). An investigation of the role of working memory in procedural skill acquisition. *Journal of Experimental Psychology: General, 117,* 319-331.

APPENDIX A
SPEAKING SPAN TEST

I. Stimuli:

First trial:
2
cake
hand
3
duck
pen
gas
4
desk
road
glass
brain
5
sun
mouth
key
bag
file
6
clock
wave
tool
coat
map
year

Second trial:
2
week
rain
3
club
spring
knife
4
table
sky
deer
ball
5
bank
shirt
egg
date
hair
6
cow
pair
church
sea

bus
dinner

Third trial:
2
bird
cup
3
snow
paper
cheese
4
blouse
class
farm
letter
5
day
arm
water
box
mall
6
dog
room
night
spoon
woman
butter

II. Subjects' responses organized from lower to higher working memory capacity as measured by the Speaking Span Test

Subject 1: S

2
I bought one slice of cheesecake.
My hands are big.
3
I went to a gas station.
I bought a red pen yesterday.
I like Donald duck.
4
I bought a desk made of wood.
This country have big roads.
I have a beauty glass.
I have a bad brain.
5
My mouth is big.
I like this bag.
6
The alarm clock is ringing.
People dies engulfed by a big wave.
2

I work 5 days a week.
I love rain.
3
In Japan young generation like going to club.
I cut my finger with a knife.
Spring came after a long winter.
4
I like catch ball.
Sky was clear.
I have a big table.
5
I have long hair.
My husband bought a sweatshirt.
I went into the West Bank.
6
The Japanese usually have dinner.
I have a pair of trousers.
Holly cow!
2
I'm feeding a beautiful bird.
I broke a cup this morning.
3
My sister loves Parmesan cheese.
There's red paper on the desk.
I made a snowman of snow.
4
I wrote a long letter to my friend yesterday.
I skipped a pronunciation class today.
5
I bought a basket ball.
It's a beautiful day, isn't it?
I went to the mall yesterday.
6
I don't like butter.
I saw beautiful woman walking that street.
That dog is mean.

Span: 11/60

Subject 2: MJ
2
I like eating cake.
My hands are small.
3
I saw a duck near my house.
I have a pen.
4
This is a big desk.
I drank a glass of juice this morning.
5
I have a file.
I saw the sun set.
6
I have a heavy coat.
This is 1999 year.
I look at the map sometimes in the car.

2
I go to the supermarket once a <u>week</u>.
3
I sometimes go to the golf <u>club</u>.
I use a <u>knife</u> when I eat.
4
Today we can see blue <u>sky</u>.
I saw a <u>deer</u> near my hose.
I have a big <u>table</u> in the house.
5
I have not had my <u>hair</u> cut for 4 months.
I ate an <u>egg</u> this morning.
6
I'm going to cook <u>dinner</u> this evening.
I don't go to <u>church</u> because I'm not Christian.
I have a <u>pair</u> of shoes.
2
I have kept a <u>bird</u> for years in Japan.
I drank a <u>cup</u> of coffee.
3
I like <u>cheese</u>.
There's a lot of <u>snow</u> here.
4
I wrote a <u>letter</u> to my friend.
I attended my English <u>class</u> this morning.
I have a <u>blouse</u>.
5
I'm going to the <u>Mall</u> of America this weekend.
6
I need <u>butter</u> to bake a cake.
I have a nice <u>bedroom</u>.
I know that <u>woman</u>.
I use <u>spoon</u> when I eat soup.

Span: 12/60

Subject 3: Ay

2
I like <u>cake</u>.
3
The <u>duck</u> is in the pond.
I forgot my <u>pen</u> at home.
4
The <u>desk</u> is made of wood.
The <u>glass</u> is on the table.
The <u>road</u> is closed.
5
The <u>sun</u> is rising.
This <u>key</u> opens the door.
6
I need a <u>tool</u>.
I need a <u>coat</u>.
2
This <u>week</u> I have lots of things to do.
The <u>rain</u> is falling.
3

Do you have a knife?
I never go to the club.
The spring is the best season of the year.
4
Hunters like to hunt deer.
5
The bank is closed.
You hair is black.
6
I go to church to pray.
2
That's a beautiful bird.
Before coming here I had a cup of tea.
3
I need to write a paper.
4
You're wearing a nice blouse.
5
The day is beautiful.
6
The woman is crying.
The dog is barking.

Span: 17/60

Subject 4: Ben
2
I made my birthday cake.
I washed my hands.
3
There are a lot of ducks around the park.
I used a red pen to write.
The gas is dangerous material.
4
Can I drink a glass of milk?
5
The earth go around the sun.
I have the key.
6
I used my tools.
I watched my map.
2
What are you going to do this week?
It's raining.
3
They went to the club.
Do you like the spring?
4
She is sat on the table.
I like looking at the sky.
The children are playing with the ball.
I like the deer.
5
I went to the bank.
I like fried eggs.

I'm wearing a shirt.
6
He called me cow.
I went to the church.
2
I like this bird.
What a beautiful cup it is.
3
There are much snow.
Write your name on this paper.
4
She's wearing a nice blouse.
I went to my grandfather's farm.
I like cheese.
5
It is a heavy box.
There's a famous mall here.
6
I like butter.
We can not eat dog meat.
It's dangerous to walk here at night.
There are a lot of rooms here.

Span: 19/60

Subject 5: Flavia

2
I baked a cake.
I shake hands.
3
I love Donald Duck.
4
My arm is aching.
5
I love the sun.
I have a big mouth.
6
It's ten o' clock.
2
This week I'm going to Brazil.
I don't like the rain.
3
I 'd like to go to a club.
I need a knife.
4
This is my desk.
5
Let's go to the bank.
This is my shirt.
I don't like eggs.
6
The cow makes moo.
2
The bird is flying.
May I have a cup of coffee.
3

I don't like the snow.
This is a piece of paper.
This is a chair.
4
My blouse is blue.
I love my English classes.
5
Today is a great day!
I love Mall of America.
6
This dog is white.
I need a spoon.

Span: 20/60

Subject 6: Veronica

2
I liked to bake some cakes when I was at home in Indonesia.
In Indonesia there's a costume that we have to take anything with the right hand.
3
When I was a child my dad got me a doll.
4
Next week I have to buy a desk.
When I was in secondary school I felt from my bicycle on the road.
I have never broke a glass.
Brain is the most important part of the human being.
5
People with a thin mouth usually like to talk.
6
When I go to the beach I like watching the waves.
2
I've been here for three weeks and I already feel bored.
In Indonesia there are no days without any rain.
3
In Indonesia I joined a basketball club for 4 years.
I hurt my hand with a knife when I was peeled the orange.
4
I usually study with the table.
Deer is a very cute animal.
5
I have to go to the bank.
Last week I bought a new shirt.
If there's nothing I can eat I usually boil an egg.
I used to have long hair but now I have short hair.
6
Meat came from cow.
I have two pairs of earrings.
I usually go to church.
2
Bird is very interesting animal.
Everyday I have a cup of tea.
3
This is the first time I see the snow.
4
I hardly ever wear a blouse.
I want to have a vacation on a farm.

I usually write <u>letters</u> to my friend in Indonesia.
5
The <u>day</u> after tomorrow I'll meet my tandem partner.
I usually have <u>water</u>. I don't like soft drink.
Almost every weekend I go to the <u>Mall</u> of America with my sister.
6
I used to have a <u>dog</u> in Indonesia.
I used to study at <u>night</u>.
I usually use a <u>spoon</u> to eat, not a fork.
Isn't she an adorable <u>woman</u>?

Span: 20/60

Subject 7: MG

2
I ate a <u>cake</u> last night.
I shake my brother's <u>hands</u>.
3
I saw a <u>duck</u> today.
I filled <u>gas</u> yesterday in the car.
4
I'm sitting near a <u>desk</u>.
5
The <u>sun</u> is shining.
I have the <u>key</u> to the exercises.
There are twelve months in the year.
6
I was in the army last <u>year</u>.
2
This is my third <u>week</u> here.
It <u>rains</u> a lot here.
3
I went to the <u>club</u>.
I cut vegetables with a <u>knife</u>.
4
I'm sitting near a <u>table</u>.
I saw a <u>deer</u> yesterday.
5
I cut my <u>hair</u>.
6
I had <u>dinner</u> yesterday night.
I like to watch the <u>sea</u>.
2
I saw a <u>bird</u> on my perch yesterday.
3
This is the first time I see the <u>snow</u>.
4
I have a <u>class</u> today.
My father has a <u>farm</u>.
I wrote a <u>letter</u> to my friend.
5
Today is my second <u>day</u> here.
My left <u>arm</u> is hurt
6
I have a <u>dog</u> and its name is Fizzy.
I have a very nice <u>room</u> here.

I hate butter.

Span: 21/60

Subject 8: LR

2
I had a piece of cake this morning.
I wash my hands every day.
3
I like eating toasted duck.
I usually make notes in pen.
In China we usually cook by gas.
4
I wear glasses.
5
There's no sun today.
I have a big mouth.
I have two keys, one for my apartment and one for the building.
6
There's no clock here.
2
This is my third week in the US now.
Yesterday we had ice rain.
3
I'm not a member of any club.
The spring will be coming soon.
I eat steak with a knife.
4
I'm sitting in front of a table.
It was surprise for me to see the deer in the highway.
5
My account belongs to TCF Bank.
I bought a white shirt for my husband.
6
I saw a cow.
The shoes always should be a pair.
There's a church very close to my apartment.
2
There are a lot of birds in the tree.
I drank a cup of tea.
3
There was a lot of snow this winter.
I wrote the number in a piece of paper.
4
Your blouse is beautiful.
There are lots of animals on that farm.
5
Today is not too cold day.
I usually drink a lot of water.
I bring a lunch box with me everyday.
6
I love dogs very much.
I have a one bedroom apartment.
I think does not have moon tonight.

Span: 23/60

Subject 9: I

2
This afternoon I'm going to bake a cake.
I hold a pen in my hand.
3
There's a duck on the pond.
This pen can write in blue.
I went out of gas this morning.
4
There's a desk in the writing room.
On the road there are many cars.
There's some glass on the street.
I have a very bad brain at the moment.
5
0
6
That is a big wave.
There's a clock on the wall.
There are 12 months in the year.
The sun is shining.
2
It rained a lot last week.
I saw the rain through my window.
3
I like going to the club.
I cut vegetables with a knife.
4
I saw a deer.
I like playing with the ball.
There's a letter on the table.
5
I go to the bank every week.
I usually wear a shirt and trousers.
6
I have a pair of shoes.
There's a cow in the field.
2
There's a bird in the sky.
There's a cup on the table.
3
Outside there's snow.
There's paper on the desk.
4
I have to write a letter.
5
This is a beautiful day.
I have a box at home.
I go to the mall every weekend.
6
There's a dog in the room.
There's a bed in the room.

Span: 25/60

Subject 10: PA

2
I like that <u>cake</u>.
Could you give me a <u>hand</u>?
3
The <u>duck</u> is black.
I have a new <u>pen</u>.
The car is out of <u>gas</u>.
4
This is not my <u>desk</u>.
My <u>brain</u> can not memorize all these words.
5
The <u>sun</u> is shining.
Her <u>mouth</u> is open.
The <u>bag</u> is heavy.
This <u>file</u> is full
6
The time in the <u>clock</u> is two o'clock.
2
The weather this <u>week</u> is pretty good.
3
I like to go to that <u>club</u>.
The <u>knife</u> is so sharp.
4
I have a <u>table</u> in my apartment.
There is a <u>ball</u> over there.
5
My <u>bank</u> is Northwest.
The sky is blue.
6
Yesterday I bought a new car.
The year is running very fast.
I enjoyed the <u>dinner</u> we had two months ago.
2
I saw a <u>bird</u> this afternoon.
I like drinking coffee in the car.
3
The <u>snow</u> in Minneapolis is very hard for us that come from another country.
I don't like <u>cheese</u>.
4
Your <u>blouse</u> is cute.
I had <u>class</u> before coming here.
I have to write a <u>letter</u>.
5
Tomorrow is the <u>day</u> when the assignment is due.
Let's go to the <u>mall</u>.
6
The <u>dog</u> is trying to reach the cat.
Yesterday I went to a dinner with my friend.

Span: 29/60

Subject 11: AL

2
I like chocolate <u>cake</u>.

I have one hand.
3
I like Peking duck.
I don't have a pen.
I stopped at the gas station.
4
The pen is on the desk.
5
The sun is shining.
6
The clock is on the table.
I have a tool set.
I have an ordinary map.
2
I'll see you next week.
Come and see the rain.
3
Let's go to the club.
We are looking forward to the spring.
I have one knife.
4
The book is on the table.
There's a deer in the zoo.
5
There's a TCF bank near here.
I'd like to have a sunny side up egg.
6
There's a cow on the farm.
2
There's a black bird.
Next year we'll have a world cup.
3
There's a lot of snow.
I don't have any paper to write soon.
4
That's your blouse.
It's a big farm.
5
Today is a bright day.
He has only one arm.
I'd like some water.
6
I have a dog.
The apartment has only one room.
Today we'll have a freezing night.

Span: 32/60

Subject 12: KB
2
This cake tastes good.
I usually use the right hand to eat.
3
Look at the duck over there.
I like pen, rather than pencil.
I need to go to gas station.

4
I don't have a desk in my house.
The road is really wide.
Please give me a glass of water.
5
The sun is very bright today.
Your mouth is really big.
I lost my car key.
6
This clock is out of order.
The sound of waves makes me feel good.
Look in the tool box.
I love this coat.
What is the map showing?
2
What are you gonna do this week?
I hate the rain.
3
I went to a nightclub last night.
I like the spring.
A knife is a very dangerous object.
4
We can't find a table in this restaurant.
5
I went to the bank to withdraw some money.
Your shirt is really nice.
I heard that egg has a lot of cholesterol.
What is the date for today?
6
Look at the cow!
I go to church on Sunday!
I feel very good by the sea.
I have never taken a bus in Minneapolis.
2
I like to hear a bird sing.
A cup of coffee sometimes relaxes me.
3
There's a lot of snow in Minneapolis.
I always prefer white paper to colored paper.
I like cheese but I don't like cheesy food.
4
Your blouse is really nice.
I had class today.
Is there any farm around here?
5
You work hard every day.
My arm is hurt.
Are you afraid of the water?
Is there any letter in the mailbox?
I went to the mall two days ago.
6
I'm afraid of dogs.
I like to stay up until late at night.
I need a spoon.
This is the woman I met in the mall.
I don't like butter.

Span: 43/60

Subject 13: AK

2
The cake is delicious.
My hands are beautiful.
3
The duck is beautiful.
The pen is not blue.
The gas station is not open.
4
The desk is empty.
The road is slippery.
The glass is broken.
5
The sun is hot.
My mouth is big.
I forgot my key.
My bag is heavy.
The file was lost.
6
The wave is big.
I have my tool.
The coat is beautiful.
2
I was busy last week.
The rain is cold.
3
I went to the club yesterday.
The spring is a beautiful time.
That knife is very dangerous.
4
The table is big.
The sky is blue.
The deer is so cute.
5
I have to go to the bank.
My shirt is beautiful.
I like eating egg.
My hair is beautiful.
Tomorrow is the day I celebrate my wedding anniversary.
6
Look at the cow.
I like going to the church.
The sea is beautiful.
The bus is too crowd.
2
The bird is singing.
I had a cup of coffee.
3
The snow is white.
I need a sheet of paper.
I like cheese.
4
My blouse is blue.

I have <u>class</u> tomorrow.
I love <u>farm</u>.
I have to write a <u>letter</u>.
5
The <u>day</u> is beautiful.
My <u>arm</u> is hurt.
The <u>water</u> is so hot.
The <u>box</u> is empty.
6
I have a <u>dog</u>.
My <u>room</u> is a mess.
I love the <u>night</u>.
The <u>spoon</u> is made of silver.
The <u>woman</u> is beautiful.
I like <u>butter</u>.

Strict score: 47/60

APPENDIX B
OPERATION-WORD SPAN TEST

I. Stimuli:

(8/4) + 2 = 4	OUT
(9x3) +2 = 29	DANCE
(9/1) + 1 = 10	WIRE
(10X2) + 3 = 23	LOCK
(8/1) –5 = 15	HARD
(7/1) + 6 = 12	FILE
(10/1) +1 =11	MEAL
(9/1) – 8 = 18	SCORE
(5/5) + 1 = 2	BOMB
(10/1) –1 = 11	KNEE
(4/2) –2 = 2	FORM
(10/2) + 6 = 10	TOWN
(6X4) –1 = 25	KNIFE
(7X7) – 1 = 49	EAST
(7/1) + 2 = 7	NOSE
(7/1) + 2 = 7	DUST
(9X1) – 9 = 1	CUT
(6X2) + 2 = 14	HALL
(3/1) + 1 = 4	ARM
(6/3) + 2 = 4	DOOR
94X2) –1 =9	GAS
(9/3) + 1 = 4	FAR
(9X7) – 1 = 49	CAUSE
(7X2) =3 = 17	TRUCK
(9/1) + 5 = 14	JUMP
(3/1) –1 = 4	BIRD
(10/10) + 9 = 19	RAIN
(3X2) – 1 = 6	HEAD
(10/2) – 4 = 3	CARE
(3/1) + 3 = 6	KEY
(6X1) – 6 = 1	BREAK
(9/1) + 8 = 18	MOON
(5/1) + 4 = 9	TREE
(9/3) + 3 = 6	ROCK
(8/4) – 2 = 2	SNAKE
(7X2) – 1 =14	FACT
(4/1) – 4 = 2	GUY
(6X3) –2 = 17	STAY
(4X2) + 1 = 9	TYPE
(7X1) – 6 = 2	BUY
(10/1) + 3 = 13	GREEN
(4/1) – 1 = 5	TASTE
(8/1) – 6 = 4	WAIT
(10/1) – 3 =13	DREAM
(10X2) + 3 = 23	TALK
(8X1) +8 = 16	HELP
(10X6) + 1= 61	NEAR
(4/2) – 1 = 3	BACK
(4X4) + 1 = 17	BRAIN
(8X4) + 2= 34	SKILL

(10/1) –5 = 7	SEND
(9/1) – 7 = 4	SCENE
(5X1) + 1 = 6	CLOSE
(6X4) +1 = 25	TOOL
(7X7) + 1 = 50	PAIR
(10/2) + 4 = 9	BEACH
(3X1) – 2 = 2	ADD
(3X1) – 3 = 1	GUEST
(2/1) – 2 = 2	SEA
(4X2) – 2 = 7	TRADE

II. Subjects organized from lower to higher working memory as measured by the Speaking Span Test

Subject 1: S
Span: 59/60
Math: 53/60
Subject 2: MJ
Span: 60/60
Math: 40/60
Subject 3: Ay
Span: 57/60
Math: 41/60
Subject 4: Ben
Span: 54/60
Math: 43/60
Subject 5: FL
Span: 58/60
Math: 55/60
Subject 6: Ve
Span: 56/60
Math: 43/60
Subject 7: MG
Span: 57/60
Math: 36/60
Subject 8: LR
Span: 59/60
Math: 40/60
Subject 9: I
Span: 59/60
Math: 53/60
Subject 10: PA
Span: 58/60
Math: 52/60
Subject 11: AL
Span: 60/60
Math: 44/60
Subject 12: KB
Span: 60/60
Math: 47/60
Subject 13: AK
Span: 58/60
Math: 58/60

APPENDIX C
DESCRIPTIONS AND NARRATIVES

Subjects' speech samples organized from lower to higher working memory as measured by the Speaking Span Test

Subject 1: S

I. Picture description

Uhm: **(0.6)** the left one (.) it is a (.) in Egypt . **(0.7)** **uh:** **(1.1)** it has a big pyramid **(0.9)** and (.) **uh** **(1.2)** *ma-* maybe inside **(1.5)** it has a mummy: . **(0.8)** and **uh** **(1.3)** I think that **(1.8)** **uhm** **(1.4)** many people **(1.2)** take some trips to visit mummy . **(0.8)** in these days **(1.6)** **uhm** **(0.6)** I (.) left side means a **(1.8)** **uhm** **(1.6)** people **(1.8)** **uhm** **(1.6)** astronaut **(1.3)** it means a: **(1.4)** **uh** **(1.9)** how people **(0.72)** people's power **(2.3)** its' growing . **(0.7)** and a **(1.)** and **(4.6)** last year **uh** last year ? No **(0.7)** *s-* (0.67)maybe couple of years ago (.) people already **(0.5)** **uh** reached the: (.) Mars . **(1.7)** so (.) **uhm** also then (.) for messages they want to **(1.1)** send us . **(1.4)** but **uh** **(2.7)** people is always **uh:** **(2.2)** look for the new thing and **uh:** **(1.)** how people **(1.3)** people's power (.) is growing (.) *growing* up **(1.7)** and **(1.5)** **uhm** **(2.4)** *and* **uh** **(0.6)** maybe long standing *ci- civirili-* civilization **(1.6)** I think .

Total time: 1m58s08 **(118s)**
No. of words: 108/103

II. Narrative:

It was a **(1.5)** **uhm** **(3.5)** man which: **(1.5)** try to: (.) save the: **(1.5)** earth . **(1.6)** he **uhm** **(2.5)** **uhm** **(0.8)** some **(1.2)** I'm not sure the **(0.8)** correct *b-* (.) word **but uh** **(1.)** meteorite (.) *meteorite* **(1.3)** almost **uh** (.) hits **uh** us . (.) **uh** (.) **uhm** **(0.6)** so **(0.7)** **uhm** **(2.5)** American army **(1.2)** send a (.) *a* special person to: **(0.8)** **uhm** **(0.5)** send a **(1.3)** *send a* special person **(2.1)** save the (.) earth . **(1.7)** and **uh** **(2.0)** they **(0.8)** *they* success **(0.6)** *success* but **uh** **(1.7)** they are proud (.) but **uh:** **(1.5)** **uh** (.) successful **(1.3)** or but **uh** **(0.6)** **uh** **(0.6)** someone had to: **(1.0)** leave the **(1.7)** meteorite (.) *meteorite* (.) *meteorite* so **uhm** **(4.1)** do you know the (.) actor ? Bruce Willis **uhm** **(2.9)** was a **(1.)** *a* **(3.7)** *a* **(1.4)** Bruce Willis **(2.3)** **uh** **(1.2)** decided to: **(1.3)** live the **(1.7)** live there and **uh** **(3.5)** stay there.

Total time: 1m59s98 **(120s)**
N. of words: 86/77

Subject 2: MJ

I. Picture description:

In Egypt **(0.5)** picture (.) there are *sh-* two **(0.5)** *s- st-* statue: **(0.87)** and **(4.1)** it's and it seems very important place . **(1.5)** and **(1.4)** **uhm** regarding astronome **(0.6)** there's one person in the: **(2.3)** in this kind of sky (.) but **(2.8)** **uhm** (.) I *I* don' t know the name . (laugh) **(2.7)** place above sky (laugh) **(0.9)** and **(2.2)** there (.) *there* only one **(2.7)** person and **(1.3)** it looks very cold (.) *cold* . **(1.2)** **uhm** (.) and dark **(1.2)** and maybe **(1.9)** this picture **(0.9)** want to say **(1.3)** the: **(1.1)** **uhm** **(1.6)** *cultu-* **(1.7)** the cultural progress ? **(2.3)** in the **(2.6)** Egypt . **(4.6)** **uhm** **(2.3)** before Christian there are **(1.6)** **uhm** **(0.7)** big **(1.)** cultural society in Egypt ? as they made big statue: **(2.3)** and **(0.5)** now **(1.)** we **(0.6)** **uh** **(0.8)** *we* **uh:** (.) *we* are creating more **(1.)** and we can go to the: **(1.5)** sky .

Total time: 1m40s99 **(101s)**
No. of words: 96/90

II. Narrative

I like the movie **(0.7)** *th-* **(.)** I'm not sure of the **(.)** *Eng-* **(0.8)** English title **(1.)** **uhm** **(1.5)** I like actress Audrey **(1.2)** Hepburn ? **(.)** *Audrey Hepburn* **(0.9)** and romance holiday **(.)** this kind of thing . **(1.7)** **uhm** **(0.5)** the reason why I like the movie ? **(.)** is that **(1.2)** I like Audrey Hepburn because she's very beautiful ? **(.)** and **(1.2)** her character in the movie is very cute so: **(.)** she's very attractive for me ? **(.)** and **(1.4)** I *I* also I like Rome ? **(0.5)** it's very **(.)** beautiful city ? an:d **(3.7)** *s-* and **(.)** there is love story . **(0.8)** the: **(0.6)** men **(.)** who are working **(.)** in this particular company ? **(.)** and Audrey Hepburn **(.)** so just **(0.6)** their relationship **(1.2)** is very: **uhm** **(0.6)** how can I say **(2.3)** very: good **(1.)** and so **(1.2)** I like that movie.

Total time: Im13s22 **(73,22s)**
Number of words: 105/101

Subject 3: Ay

I. Picture description

This **uhm** this picture **(.)** *this picture* **uh** **(0.6)** from old times . **(0.8)** from ancient times probably . **(1.2)** **uh** an old time picture . **uhm:** **(0.7)** the left hand side **(.)** is old picture . **(1.2)** **uh:** **(0.8)** it must **uh:** **(.)** *it must* belong to: **(0.8)** an **(.)** Egypt culture **(.)** maybe . **(1.6)** the other part is **uh:** **(3.4)** it includes an astronaut man . **(1.2)** I don't know what to call this **(1.4)** it **(.)** in this **uh:** *this* picture **(2.8)** **uh** **(0.8)** show **(1.2)** **uh** **(0.7)** **uh** beginning of **(.)** human being **(.)** and **(2.4)** **uh** **(.)** now where's **(.)** the **(1.4)** *where's the* human being **(1.)** and **(.)** **uh** **(0.7)** it's **(0.64)** this picture try **(.)** to describe **(.)** that the: **(1.6)** development of the life of human being **(1.2)** or the development of technology ? **(2.8)** and at the same time **(1.3)** the picture **(.)** is rather strange too . **(3.9)** maybe **uh** **(0.7)** saying **(.)** **uh** **(0.7)** astronaut work hard and **(0.9)** it's a hard job **(1.2)** in the space **(1.6)** so the: **uh** **(0.7)** hard work **(0.6)** these guys have done a hard work **(2.4)** **uhm** **(0.5)** two men are moving here **(2.4)** **uhm** **(.)** that's it.

Time: 1m39s07 **(98)**
No. of words: 130/122

II. Narrative

I saw **uh:** a movie **(.)** about **uh** **(.)** young people . **(1.5)** and **uh** **(0.7)** **uh** **(.)** the **uh** **(1.2)** American football **(0.6)** Varsity Blues **(0.8)** and **uh** **(0.6)** American **uh** football **(0.6)** **uh** I'm very fond of **uh** **(0.8)** American football . and so: sometimes **uh** **(0.9)** I watch it . **uh** **(1.2)** Americans like to play football **(.)** with their children **(1.1)** and **uh** **(0.6)** there was a coach **(0.6)** team coach **(0.8)** and **uh** **(0.7)** he had every time *on-* at this his own **uh** **(2.)** he was sad also **(0.6)** and **uh** **(0.9)** he only talks how to do things correctly **(0.8)** and this **(1.2)** a strong relationship between these **uhm** **(0.6)** **uh** **(1.)** young people: **(0.6)** and the coach of the team **(2.3)** and **(0.8)** I liked it very much.

Total time: Im08s06 **(68s)**
No. of words: 86/85

Subject 4: Ben

I. Picture description:

I:t is **(0.6)** a scene to **uh** looks like *Egyp-* **(0.5)** Egypt . **(0.5)** and there's a asphynx (sphinx) **(0.9)** a landscape and **(1.4)** **uh** **(1.)** the work have to **(1.)** about **(1.)** many image and many things . **(.)** and there is a lot of **(1.)** **uhm** **(3.4)** **uh:** *there is a lot of* **uhm** **(2.)** I'm sorry **(2.)** columns ? **(0.5)** a lot of columns an:d there is two people ? **(1.5)** **uh-uh** **(1.3)** **uhm** **(3.5)** it is day time . **(1.1)** **uh** **(1.2)** there is a small **(0.5)** *small* **(.)** aisle **(1.)** on the right **(.)** aisle **(5.4)** I think that's right **(1.1)** and the second one is a: **(5.1)** *is a* **(1.5)** he: **uh** operating the: some machine . **(.)** **uhm** the machine **has** **(.)** *the machine has* **uh** **(.)** a camera and **(.)** that **uh** **(.)** *th-* the person ? wear the helmet **(1.16)** that protect **(1.1)** for protecting the head and **uh** **(.)** he has a very special **(1.3)** **uh** wear **(1.)** that protecting **(0.9)** *hi-* his body . **(0.6)** an: d **(0.9)** *l- loo-* **uh** he *loo-* **(0.8)** he looks like fat **(0.6)** **uh** look fat . **(0.5)** and **uhm** the *l-* the *the* clothes was printed by the Nasa letter

. **(0.6) uh** letter color is red **(0.7)** and **(1.9)** *th-* **uh** the person is in the **(0.6)** out of the **(1.2)** atmosphere **(.)** the space . **(0.8) 2m // uh-uh (.)** he's using gloves **(> 0.5)** an:d that's all.

Total time: 2m07s07 **(120s)**
No. of words (up to 2m): 149/132
II. Narrative:

I watched the: **(.)** Attack Troopers **(.)** it's **uh** about the: uh **(.)** invaders **(0.5)** from **(.)** the outer space . **uhm** the enemy ? is the: bug **(0.8)** that is the cockroach . and a: **(.)** *a* very dangerous bug . **(0.7)** they also **uh (2.6)** the people **(.)** live in *in* the on the earth . **(0.6)** they want **(0.5)** *want* the truth . **uh** *figh-* fight with the bug **(0.6)** so they **(1.2)** *they* **(.)** went through the: **(0.8)** *ou-* outer space other sphere **(0.7)** so they: **(0.5)** fight with the bug **(0.6) uh** the first time they: **(.) uh** the human **(.)** are was defeated by the bug . **(0.5)** but **uh (.)** the: in the conclusion **(0.5)** *conclusion* of the **(0.7)** theater **(0.6) uh** movie **uh (0.9)** the human *figh-* **(1.3)** won the *the* fight . **(0.6)** I <u>don't</u> like this movie because the: **(0.9)** they are all over the earth so: **(0.73)** *so so* terrible **(1.1)** they are the **(0.5)** *the* **(2.7) uh:** there is a lot **uh: (.)** a human **uh (0.6)** was **uh (.)** killed by the bug **(1.1)** so terribly **(.)** and so cruel **(0.8)** so I don't like this movie .

Total time: 1m27s04 **(87s)**
Number of words: 138/127

Subject 5: FL

I. Picture description

Ok Mailce **(.) uhm: (0.8)** it is very difficult ? but **(0.8)** I will try . **(.)** this **(.)** figure (laugh) **(0.8)** sho:ws **(0.5) uh:** I guess **(.)** they are talking about **(0.5) uhm: (1.2)** the evolution of the history **(.)** or **(.)** the evolution of the humanity . **(0.5)** right ? **(1.1)** 'Cause **(0.8) uh:** ha:lf **(0.8)** of the figure **(1,6) uh (.)** shows **(1,1)** like an Egypt **(0.5)** symbol **(0.7)** you know ? **(0.6) uh** perhaps of **(.)** one thousand years ago **(0.6)** you know ? and the *predom-* predominant color is re: d . **(1. 5)** an:d the other si: de **(1.)** of the f: igure ? **(0.7)** shows a man in space **(1.6)** ok ? **(0.7)** and the predominant color is blue ? **(1.2)** an:d it's written the official network of every millennium **(1.3)** so it shows a: **(0.7)** the evolution ? **(0.7)** you know ? of humanity . **(1.2) uh: (1.)** what else **(.)** can I say ? I don't know **(0.6) uh: (1.5)** I guess that's all.

Total time: Im18s65 **(78,65s)**
No. of words: 108/107

II. Narrative

Uh: (1.) I'm going to: **(0.7)** talk about the **(0.7)** movie called city of angels . **(.)** ok **(.)** it's really beautiful it's really romantic **(1.1)** the story **(0.7) uh:** is about **(0.6)** an angel **(.)** that falled in love ? **(1.1)** with a woman ? **(0.6)** and she fall in love with him ? **(0.8)** an:d **(0.7)** *an:d* **(1.1)** he was trying to: **(0.7)** begin human **(1.1)** to **(.)** marry him **(0.6)** he her ok ? **(.)** an:d to become a human ? **(0.7) uh** all *e-* **(0.7) uh:** he has to do is **(1.1)** to jump **(0.6)** and **(0.5)** fall **(0.6) uh:** on the floor . **(.)** ok ? **(1.3)** and so **(.)** he jump **(1.4)** by **(0.7)** the building ? **(0.6)** ok ? **(.)** and fall the floor ? **(1.)** and he get it he got it ! yeah **(.)** he becomes **(.)** and so: **(0.7) uh: (1.2)** he can stay with her but after this unfortunately **(0.5)** she died ! **(0.7)** and so she stayed alone in this **(.)** terrible world **(0.6)** yeah he: stays alone **(.)** yeah **(.)** I liked it **(.)** very much **(.)** it's very beautiful **(.)** the movie.

Total time: Im37s13 **(89,88s)**
No. of words: 130/128

Subject 6: Ve

I. Description

Uhm (0.7) this is a **(.)** pi: cture **uhm** like **(1.)** sometimes like a statue . **(0.6)** and **(0.8)** is **(0.5)** only **(0.7)** in one country . **(2.8)** like **(0.8)** this side **(1.1) uhm (0.7)** there's lot of **(0.8)** pillars there . **(2.5)** and **(3.3)** they have **(1.1)** the country ha:d **(0.6)** has **(1.1)** a pyramid . **(1.1)** an: d **uhm (1.8)** *and* it's some kind like a: **(0.9)** sculpture **(1.)** I mean **(1.6)** the **(1.8)** old treasure of the country . **(3.7)** and **(0.7)** the other side is **(1.1)** an astronaut **(2.5)** he's fly the **(.)** base **(1.7)** like in the movie Armageddon **(4.7)** the:y **(0.8)** he: **(4.)** sit **(1.3)** on some kind like a: **(1.7)** robot **(3.4)** they *wea:* he wearing a **(2.5)** white **(0.6)** clothes **(2.5)** and **(0.5)** a helmet . **(3.5)** I think the picture is very interesting **(0.7)** especially **(1.7)** the statue one . because it's something like a **(3.5)** it's known in the world **(0.7)** and it's very famous.

Total time: 1m55s50 **(115,50s)**
No. of words: 108/106

II. Narrative

The **(1.)** I like **(0.5)** Patch Adams very much . **(0.5)** it's **uh: (.)** it is a **(.)** true story **(0.9)** and **(1.8)** *and* I never thought that **(1.2)** there was a man like that **(0.7)** who cared for **(2.3)** somebody else who was **(.)** not as **(1.6)** a related **(.)** to him . **(1.7)** and it 's very touchful I guess **(2.2)** and (cough) **(1.7)** he *jus:* he tried to do anything **(2.)** for make **(2.5)** others happy **(3.9) uhm (3.6)** it's about **(2.6)** a man who **(0.9)** *who* likes to help people ? **(0.7)** and study in a **(1.2)** medical . **(1.3)** in the medical ? he's the oldest from all the school ? **(1.)** so **(3.2)** he likes **(1.8)** to visit the hospital ? **(.)** and cheer up the patients there ? **(1.4)** he never studies **(1.)** but he always **(.)** has **(.)** a good result . **(.)** the in the test **(.)** only his *cl-* classmate was jealous with him and said bad things about him **(0.83)** that he's cheating **(1.3)** and he always **(0.9)** dreamed **(1.6)** about helping.

Total time: 1m48s99 **(109s)**
No. of words: 130/127

Subject 7: MG

I. Picture Description

Well **(.)** when I first **(0.5)** looked at the picture: **(0.7)** I divided it **(0.6)** because **(0.6)** the: left side is more **(1.1)** looks like a picture from Egypt .. **(0.9)** and **(0.8)** Egypt is very close to Israel . **(0.6)** so I saw this kind of stuff but **. (0.7)** when I look at this whole picture ? **(.)** I think that **(1.3)** this **(0.8)** the **(0.8)** they try to **(0.6) uhm (2.)** connect **(1.)** the past **(0.9)** and the: future . **(2.)** an:d **(1.1)** we **(1.2)** I can see here **(.)** like **(.)** **uhm**: **(3.)** the difference in technology **(0.6)** this is very *sophisti-* **(.)** the *the* right side is very sophisticated **(0.6)** an: d there is very: **(0.6)** in **(.)** in the future ? **(0.8)** and the left side is very: **(0.6)** old ? **(0.8)** and: not sophisticated **(2.5)** an:d **(2.8)** *and* **(0.9)** I don't really know what they're trying to say here ? **(0.8)** bu:t **(5.6)** **uhm (.)** well I don't have anything to say but *but* I really **(.)** want to know what it means.

Total time: 1m29s88 **(89,88s)**
No. of words: 129/125

II. Narrative

Ok **(.)** it's about a man that **uhm : (.)** went to prison because **uhm (1.) uhm** the judge *judge* believed that he murdered **(0.7)** his wife and he: her **uhm** mistress . **(.)** an:d **(0.7)** so he's been in the jail like twenty years ? **(.)** and during this time he was **(1.2)** digging a hole from his **(0.6) uhm (1.5)** room ? it's not really room ? but I don't know any other word **(0.7)** so: **(1.4)** digged he *th-* he dug this room ? an: **(0.5)** a ho:le **(0.8)** a very long one ? **(.)** and after twenty two years he: **(0.7)** *he*: **(0.6)** went out **(0.9)** an:d **(.)** ran away from the prison . **(0.7)** so: **(1.1)** that's it now he's free . **(.)** I liked it very much because when: **(1.)** when I saw the **movie (0.8)** I *l-* I tried to see myself instead of him ? **(0.7)** and I think I'd do the same.

Total time: 1m05s13 **(65,13s)**
No. of words: 119/117

Subject 8: LR

I. Picture description

I have a picture here (.) an:d (.) the: (.) in the: (.) left part of the picture (0.7) there is a: (0.5) there are *s-* (.) **uhm** (.) several (0.7) a lots of columns (.) *th-* large (0.6) tall or high columns ? (1.) an:d (.) it seems (.) like in Egypt . (0.6) an:d (0.6) the: (1.) in the *mid-* (1.1) middle have (.) **uh** (.) in the middle have other picture on the left (0.5) there is half (0.7) *st-* statue ? (1.1) like a (0.5) *s-* (.) you can imagine like something like a (.) a *ei-* age (.) how do (.) *how do* (.) a pyramid (.) yeah (.) there's a man sit there (.) and it's very huge (0.6) and you can see: (.) there're two: (1.4) s:ome people like who (.) come from *ei-* (.) Eygypt and here (.) and it looks very small (.) so the pic.ture and the statue looks very huge . (.) and on the: (.) left side of the picture (.) it's a blue one ? and there is a: man ? who is the (0.7) astro.naut (0.5) looks like (.) he is in the space ? (0.6) and *th-* this like some in the: (0.9) like something you can see: in (0.7) the: science fiction (0.6) *f-* movies ? and I think this means (0.9) **uh** you can compare with the (.) old part of time (.) and the: (0.6) *the the* science developing very quickly now you can go to the (.) space (0.7) an:d (0.5) and find some (1.1) *find some* (0.7) **uhm** (.) surprise there.

Total time: Im52s67 **(112,67s)**
No. of words: 194/186

II. Narrative

Uh (0.5) several weeks ago (.) I saw: a movie called Mighty Joe Young ? (0.5) I loved this movie very much (0.7) because (0.6) somehow I think I *l-* love animals . (0.7) and also (.) this story talked about a (.) gloria (gorilla) (0.6) a lots of people wanted to protect him ? (0.5) because (.) the nature was destroyed by some people and also (0.7) some (.) bad people wanted to make money ? from the animals ? they *k-* killed the animals or they (.) just **uh** (0.7) **uh:** how to: chased them ? (.) and they locked them ? (1.1) made money from them . (0.6) so: (.) this story talks about a (.) a lady ? I forgot her name ? (0.5) and she wants (0.5) to protect this gloria ? (.) and this is a very huge (.) *huge* gloria (0.9) like thousand pounds . (.) really (.) very huge (.) I think it's not real (.) it's not true (0.6) how do you say that the word made by machine ? (0.6) mechanical (.) yeah (.) and (.) it's looks real so it's lovely! (.) wow I loved that movie ! (.) and also you can (0.7) see his (.) *his* (0.6) face (0.9) **uh:** emotion (.) something like that he can (0.5) **uh** response for *yo-* what you say to him ? (0.7) oh! I can't find the right word for this . (.) but it's so cute you know? (.) And he (.) *he* (0.6) *he* knows who (.) treated him well ? who is (.) his animal ? oh no enemy! Enemy right? (.) who wants to kill him and (0.7) oh! I *I* (0.7) how to say that ? (0.7) and also he (.) *he* tried to: protect the (.) nice people (0.5) and because he's huge sometimes **2m** // (> 0.5) he can (.) he can do that (.) he can (.) oh (.) finally (.) finally he caught a: (> 0.5) very nice place (> 0.5) and they got some money from the: (> 0.5) from the some organization (.) and they they (.) how do you say ? (> 0.5) they build they build a nice (.) garden for him (> 0.5) yeah (.) like in the nature (.) he can live there (.) freedom and nobody: like **uh** want (.) can nobody can (> 0.5) atta- attack ! attacked him (> 0.5) It's it's it's very good movie! I love that!

Total time: 2m45s52 **(120s)**
Number of words (up to 2m): 226/220

Subject 9: I

I. Picture Description

I see **uh:** a picture: (.) that is **uh** divided into two halves . (0.7) **uh** (.) the left hand (.) **uh** side (1.1) **uh** shows an (.) Egyptian **uhm** temple: (.) with a: (2.) very big **uh** (0.9) symbol of an Egyptian god in the: very center of the picture . (0.7) and uh the right handed side **uh** shows (.) **uh** a flying astronaut in the sky . (1.1) **uh** (.) the connection of the two: halves (.) **uh** is given **uh** through that the two: (0.7) **uh** figures the Egyptian god and **uh** the astronaut (.) are in the: **uh** (1.1) very front and splitted in two halves . (.) as *as* if each (.) as if they were one person **(0.8)** but still the two different sides give a very different impression and a different mood . (.) **uh** and **uh** give the *the:* **uhm** **(1.8)** the impression of really two very different pictures that are just (.) thrown together **(1.)** **uh** the **(0.5)** *the* left hand side ? of the Egyptian **uh** temple (.) **uh** (.) sh:ows this very big **uh** god *sh-* **uh** from (.) the top to: the: **(0.8)** **uh** basic of **uh** the picture: (.) and **uhm** **(0.6)** on the left hand side ? shows **uh** a row o:f **uhm** (.) columns **uh** of **uh** six columns **(0.6)** that **uh** lead to the end (.) **uh** there's no move on *on* **uh** the temple ? you can look into the: **uh** (.) kind of **uh:** very bright ? sounding sky **(0.6)** right hand side oh is it finished ?

Total time: 2m **(120s)**
No. of words: 184/180

II. Narrative

Yesterday night **uh** (.) I saw **uh** (.) the film Gold Rush by **uh** Charles Chaplin . **(0.5)** and **uh** I saw that a long time ago (.) **uh** but I couldn't I could hardly (.) *hardly* remember it anymore . **(0.8)** you don't know the film ? so **uh** (.) **uh** (.) the story of the film (.) is very easy (.) it's **uh** basically is about this *li-* fellow (.) **uh** played by Charles Chaplin who goes out into the world (.) to **uh** make his money to make a fortune and **uh** to: **uh:** (.) get **uh** the woman **uh** that he loves . **(0.7)** so **uh** the story is very simple and there is a *a* happy end (.) in that story as well . **(0.8)** but **uhm** **(1.9)** yeah **(0.7)** first of all in that (.) **uh** whole **uh** story the language is not important at all . (.) like in <u>all</u> Chaplin films **(0.6)** **uh** it's more this kind of basic (.) **uh** playing he's doing . **(2.3)** and **uhm** **(0.5)** I liked it especially (.) **uh** because of **uh** the (.) I liked it especially because of **uh** **(0.7)** the: **uh** possibility of Chaplin to express (.) really basic skills (.) and basic feelings **(0.6)** of all people in a very inter-national (.) language . **(0.9)** so **uh** I always feel very touched when I see him **(0.8)** **uhm** emotionally: **(0.9)** **uh** and **uh** he touches me (.) especially that (.) he can be so: **uh** show so opposite so different things in on in *in in* one scene . (.) so he can be **uh** poor and rich (.) at the same time and **uh** **uh** especially use a lot to be weak **(2m)** and clever both in one (.) and so that I find really **uh** amazing **(0.5)** the way how he can express his **uhm** (.) normal and usual daily life **uh** **(0.5)** contro-versial **uh** things but he can put it **uh** into one situation into **uh:** one person and I find that amazing **(0.5)** **uhm**: **(0.5)** also I think (.) although the film **uh** emotionally **uh** touches you straight away (.) it gives you the: **uh** opportunity **uh** in **uh** lots of situation **(0.5)** to think about **(0.5)** **uh** your own life and about **uh** the own kind of basic **uh** situations you are in (.) at the moment and the kind of handling and feeling **uh** that you are doing for those (.)so he gives you a good opportunity to: think about yourself and about your own life (.) for example what **uh** makes **uh:** **uh:** it **uh:** (.) what makes it **uh:** **(0.5)** strong or weak **(0.5)** and **uh** what kind of which kind of strength do you yourself need to keep your own life in a balance and **uhm** (.) yeah (.) what is really important for you personally and **uh** whom do you really love and this kind of basic really basic questions in your own life (.) so: **(0.5)** I think **uh** there are different kinds of **uhm** levels that are touched by this film (.) it's not only the emotional **uh:** **uhm:** (.) skills (.) and **uh** ways you can see the film you can also **uh** take it very intellectually: **(0.5)** and **uh** think about different things in **uh** a way and **uh** (.) yeah use it as an opportunity to **uh** go forward yourself .

Total time: 3m5429 **(120s)**
No. of words (up to 2 m): 211/206

Subject 10: PA

I. Picture description

Well (.) the picture (.) is divided (.) into two half . **(0.6)** one of them shows (.) **uh:** half of an old sculpture (.) of a king of a pharaoh **(0.6)** in: Egypt . **(1.1)** the second complementary half **(0.6)** shows an astronaut (.) in: the space . **(2.8)** in my: opinion (.) the idea (.) that **uh** the picture *tri-* tries to *to* show **(0.6)** **uh:** the message that the picture tries to give **(0.6)** us . **(0.5)** is of the evolution ? of **uh:** mankind (.) since the ancient times to our days **(4.1)** and **uh:** (.) trying to: **(0.6)** *to* show **uh:** the technology: **(1.)** **uh:** employed (.) in our days and the relation (.) with **uh:** (.) *with* the old times (.) when the: those this technology were **uh** was not available at all . **(1.2)** and even though (.) the man **uh**: kind (.) was able to: **(0.5)** *to* evoluted *s-* **(0.7)** from that time until now . (.) that's ok.

Total time: Im08s99 **(69s)**
No of words: 118/114

II.Narrative

Well (.) I *I* will talk about a *a* movie (.) actually **uh:** (.) one that I saw last week **(0.7) uh:** on the tv . **(0.7)** it was a series **(0.6)** and **uh:** (.) it **uh: (0.7)** *it* was about the war (.) between Israel (.) and the Arabs **(1.1)** and **uh:** it was divided into (.) two sessions . (.) the first one was on Sunday ? **(1.58)** and **uh:** three hours **uhm** in the evening ? (.) an:d the second part (.) was on Monday (.) also three hours (.) without any break . **(1.1)** for both of them . **(1.3)** an:d **uh:** (.) what was interesting ? **(0.5)** was that they *they* show **(0.6) uh:** the history (.) of **uh:** the relationship **(0.8) uh:** of Israel ? (.) with all the surrounding countries ? **(0.9)** an:d since the (.) *the* war in nineteen sixty seven **(0.5)** with the *the* war between **uh** Israel (.) and the Palestinians ? **(0.7)** when the: **(0.6)** Israel won and **uh: (0.6)** *and* took some **uh: (0.5)** lands of- from Egypt ? **(0.7)** and the Sinai ? (.) and also from: the: *the:* **(0.7)** *the:* wills (hills) of **uh:** Jordan ? **(1.2)** and **uh:** it *it* showed (.) images of the war at that time ? of **uh:** (.) *of* all this period of time different wars not wonly (only) one but different wars . **(0.5)** and also ? we: *we* heard the testimony (.) of different people (.) involved (.) in different ways (.) in that war **(0.7)** in those wars (.) and **uh:** we had opportunity (.) to: *to* hear **(0.7) uh:** and to see the images and *and* hear the deposition (.) of people ? that -*e* were involved (.) **uh:** like **uh** presidents (.) of the United States giving their own deposition ? (.) about what **uh** how was the: (.) how were the negotiations ? at that time **uh**: in the different periods *o-* of the war ? how do they (.) *did* they act ? (.) in order to: *to* solve the conflicts ? (.) and **uh: 2m** also people from different nationalities (.) that were involved in different countries ? (.) and how do they: try to solve the: the the questions the probems **uh:** of **uh:** mainly related to to Isra- Israel **(0.5)** and **uh: (0.5) uh:** also **uh:** about **uh: uh th-** k- **uh**: the king Hussein ? **uh:** from from Jordan **(0.5)** he: he: **(0.5)** he had a very strategic **uh: (0.5)** participation in those events (.) and he talked about **uh:** (.) what what **uh:** (.) what was exactly his participation how how did Jordan saw (.) **uh:** the conflict at that time (.) and **uh:** also **uh**: we had image from Egypt ? (.) and Iraq (.) and **uh:** (.) another part that called my attention was the: **(0.5) uh**: **(0.5)** the creation ? of **uh** the: (.) PRO **(0.5)** how u: the organization was **uh:** organized and created **(0.5)** uh: at the beginning of this period of time. That's ok.

3m22s74 **(120s)**
Total number of words (up to 2m): 245/230

Subject 11: AL

I. Picture Description:

Uh this picture **(0.9)** is from **uh** history channel **(0.5)** and **(1.5)** it represents (.) **uh** (.) two steps of the human **(1.4)** race . **(0.9) uh** (.) its represents **(2.)** by the: **(2.8) uhm (2.1)** *the* placement **(0.5)** of **(0.5)** a *egi-* Egyptian (egypchian) egyptian **(0.6)** egyptian sculpture . **(1.1) uh** m- I think it's a **(0.9)** kind of a god **(0.8)** *an*-**uh** half of it and on the other half (.) we have **uhm** astronaut on a (.) astronaut suit **(0.6) uhm (2.5)** in space . **(0.9)** and **uhm (1.3)** as I said it's the other part (.) of the sculpture (.) so we have the representation of **(0.6) uhm (2.5)** *of* the human race (.) I *I* mean four *hundr-* four thousand years **uh (1.8)** ago . and **uh:** the human race kind of representation nowadays . **(0.5) uh: (0.9)** and it seems to *to to* say that (.) in the past (.) the world was **uh (0.8)** represented by **(1.8)** the **(1.1) uh: (1.7)** Egypt? **(0.9)** Of Egypt ? **(2.4)** and nowadays (.) it's **uh** represented the human race is represented by the (.) *the* domination of the United States . **(0.8)** because you'll see that the astronaut's from the United States (.) and the TV channel is from the United States **(1.5) uhm (2.1)** ok that's it .

Total time: 1m48s17 **(108,17s)**
No. of words: 151/142

II. Narrative:

OK (.) the film (.) I will describe is called (.) the thi:n red line . (.) and (.) unfortunately I *d-* I did not like it **(0.5)** at all because **uh** I couldn't (.) *couldn't* understand what **(1.3)** the actors **uh** were saying . **(1.)** and **(0.5) uh** (.) so: I *I I* just tried to figure (.) out what was happening (.) and what were they saying **(0.7)** and what was it all about . **(0.9)** I think **uh** well **(1.5)** it's about ? the film is about (.) **uh: (1.1)** the American **(0.5) uh** occupation (.) during the second **(1.5)** world (.) war **(0.9)** and **uh** in the south Pacific (.) and it's about the **(0.9)** *th-* the reasons **(1.)** of *of* this I guess so: the reasons **uh: (0.5)** for young *sol-* soldiers (.) to be fighting that war **(0.5) uh** away from home **(0.5)** away from the persons *o- o-* of the people they *the-* they loved . (.) and **(2.7)** *and* the reasons (.) to be (.) fighting that war **(0.9)** so the:y **uh:** *th-* the film? **Uh:** describes **(1.)** this kind of thinking **uhm** (.) what did they **uh** think ? and **uh** what did they feel ? (.) their fears and **(0.8) uh: (2.)** what was on their mind (.) and (.) **uhm** (.) the *the* film **uh (1.3)** also shows **uhm (0.5)** the problems concerning the: **uhm (1.8)** *the* orders (.) that **(0.7)** the soldiers in the front **uhm (0.8)** had to follow . **(1.1)** but these orders **uh:** were from **(1.8) uh** superiors which were not (.) in the right in the front as they were **2m** so: **(0.5)** they ordered things that (.) they were not able to: **(0.5) uh** accomplish (.) this kind of drama **(0.5)** and situation .

Total time: 2m10s23 **(120s)**
No. of words (up to 2 m): 189 /180

Subject 12: KB

I. Description

Uhm (.) at the: right side of the picture (.) on right side of the picture (.) there is a: **(0.7)** spaceman ? **(1.5)** who is flying **(0.6)** in the space (.) in **(0.6)** space clothes . **(1.2)** and **(0.6)** in left side **(0.7)** there is a **(0.6)** huge **(1.3) uh:** picture **(0.6)** of **(0.8)** Asian king ? maybe **(1.1) uh** (.) and yeah (.) just a half **(1.1)** *half* of their body **(0.6)** and (.) yeah the: **(1.3)** their bodies are (.) matching **(0.7)** as a one person (.) **uh** like one person . **(1.0)** and **(1.5) uhm (0.9)** yeah **(2.3) uh** in *in* this picture: **(2.) uh** (.) the **(1.3)** maybe the person who: (.) took the picture **uh** make the picture wanted to: (.) made the picture wanted to: make wan- wanted to: **(0.6)** give some message about **(0.8) uh** (.) message that there was a **(0.5)** high-ly (.) developed **(1.)** culture **(1.)** on the earth in ancient time **(1.2)** maybe before **(1.8) uh** maybe **(0.8)** three (.) *three* thousand years ago ? in my knowledge maybe and **(1.) uhm** so: **(2.1)** yeah I guess that's it . I don't have any (.) other ideas with that (.) but **(0.7)** yeah **(1.2)** I think this *this* (.) picture is (.) kind of sterile for me (.) **uh** (.) and (.) Ok.

Total time: 1m43s36 **(103,36s)**
No. of words: 150/145

II.Narrative

I like the movie (.) Love Affair **(2.0)** yeah **(1.8)** which is starred (.) **uh** starring **uh** starred by ? I don't know this grammar . (.) starred by Warren Beatty and Annete Benning **uh: (0.7)** and maybe there was a classic movie which **(1.)** had the sa: me title but **(1.)** yeah (.) the same story **uh: (0.6)** but I like (.) this movie Love Affair very much . (.) so (.) I watched it over and over maybe over (.) ten times ? **uh (0.6)** I love the movie ! **uh (0.9)** and is *is* a love story (.) **uh (0.7) uhm** yeah two: *two* persons (.) yeah a man and woman **(0.7)** *and a woman me-* met on the plane **(0.6) uh** meet on the plane (.) an:d **(1.0)** the *boo-* the man is a playboy actually and **(1.1)** the woman is (.) just ordinary person but **(0.7)** each of them has *ha-* has **(1.3)** a fiancé **(0.5)** and (.) but they felt **(1.1)** love to each other (.) felt in love with to each other (.) and **(2.1)** so maybe the title is affair but (.) I don't like the title but (.) the story is really beautiful uhm and **(0.6)** but as they have (.) fiance **(1.2) uhm** (.) so they just **(1.1)** come to meet **(1.2)** on top of the empire state building (.) with three months later . **(0.7)** after they: **(1.)** yeah (.) make some decisions during the period . **(1.6) uh** and (.) yeah they (.) of course the man the *g-* the *wo* the man **(0.7)** goes to empire state building on the day **(0.6)** and a woman (.) and the woman **(1.1)** yeah goes to empire state building too . but **(0.8)** suddenly **(1.2) uh** (.) she got she get traffic accident so: **(1.5) 2m** she couldn't go (.) she can't go she: (.) cannot go to: (.) yeah (.) to him (.) on the empire state building **(0.5)** an:d but **(0.5)** he didn't know why (.)

she (0.5) doesn't come (.) so: ju- she he was waiting (.) but (0.5) she did- she doesn't show up (.) so: (0.5) he give up gives up and (0.5) go back to: (0.5) his: (.) yeah (.) life (0.5) but (.) in Christmas season (.) in (0.5) yeah in Christmas season (0.5) the two of (.) **uh** (0.5) each of them (.) goes to: (0.5) a kind of Christmas concert ? (0.5) of Ray Charles (.) Ra- Ray Charles the singer (0.5) I like him very much (0.5) **uh** and (0.5) suddenly (.) he (.) he sees the (.) woman (.) who is (.) sitting on (0.5) on chair with (.) yeah (.) former fiancé (0.5) her former fiancé (.) and (0.5) of course he was he is surprised (0.5) and but (.) just (.) go (.) go over (.) goes over (.) without (.) saying anything and (0.5) on Christmas day (0.5) he: .**uh** (.) but you know ? **uh** yes I didn't tell this one (0.5) **uh** (.) because of the accident she: she became (0.5) she become (0.5) **uh** a handicapped person ? with leg so (.) she can't walk and (0.5) so that's why she can not meet him (.) and but he doesn't know that (.) anyway (.) on Christmas day she was (.) sitting (.) on the sofa (0.5) alone (.) reading books (0.5) but she suddenly (0.5) hears the sounds of knocking and (0.5) there was a man **uh** (0.5) there was the man (.) yeah (.) at the first time they (.) they met (0.5) they visit (.) visited (0.5) yeah (.) his aunt in Tahiti (.) island (0.5) yeah (.) that's why she (.) fell in love with him (0.5) yeah (.) I cannot explain well about this but I think you'd better. watch it (.) and (0.5) just (.) he: painted a picture for her (.) **uh** (0.5) an- and the present (0.5) was a scarf (.) re- red beautiful scarf (0.5) and **uh uh** she was moved (.) by the present (0.5) and she knows that his aunt was there (.) **uh** (.) and (0.5) and yeah (0.5) he painted a picture for her but (.) (0.5) when the **uh** and but he: (.) throw it away when **uh** in the empire state building (0.5) but (.) she (.) by accident ? she: yeah (.) found the picture (0.5) was in **uh** in the hotel **uh** in in a hotel and (0.5) he **uh** she (0.5) yeah (0.5) bought it from hotel (0.5) and (.) an: d (.) h- sh **uh** she: (0.5) she : han- hang it on the wall (0.5) in the room (.) but he didn't see it yet he doesn't see it yet because they're in living room (.) an: d (0.5) of course he heard he: heard that (.) some woman (0.5) some hand- handicapped woman (.) buy bought a picture from hotel and (0.5) just sh- **uh** he he felt like it was just a handicapped person who: (0.5) has not (.) ha- who has no relationship with him so (0.5) he doesn't (.) doubt anything (.) about it (.) but (0.5) suddenly he he knew that he knows that she (0.5) doesn't move at all (0.5) on the sofa (.) so (.) **uhm** (0.5) so he thou – he (.) thought (.) maybe and he try to (.) find (.) a picture (.) of in the house and (.) he (.) yeah (0.5) when he (.) open the door (0.5) there was a (.) picture of his own so (0.5) (.) he knew that (.) it was her (.) who bought the picture and (0.5) an:d realize that (.) he was he'd became handicapped person (0.5) so: (.) yeah (.) this is the end (.) they kiss in the end it's a happy end .

Total time: 6m47s76 **(120s)**
No. of words: 226/217

Subject 13: AK

I. Picture description

Well (.) what I see in this picture (.) **uhm** is about the: (.) history channel . (.) **uhm** (1.1) it's write in here (.) **uh** *th-* that the history channel is the official network of every millenium . (0.6) and for *e-* me ? it means that (.) it covers everything that (1.50 happen (.) is happen (0.8) is (0.7) *is* happenning since the: (1.5) *since the:* (0.8) the past until nowadays (1.0) and this *pic-* an:d the picture (0.7) we can see the: (1.3) some things (0.5) I think it's (.) from Egypt ? (0.8) it looks like the esfinge but it's not the esfinge I know that . (0.5) maybe it's the (.) coliseum in *At-* Athens I don't know (.) and *th-* the page is divided (0.6) with this (0.6) *with this* picture (.) and the other part of the page we have the an astronaut in the space . (0.9) which means the: how men are progressing (.) in technology during this millenium . (0.5) and this channel use to talk about <u>everything</u> (0.9). they *they* gave us the: (1.0) *the* email address (0.9) so you *y-* you can see: (1.2) **uhm:** (.) many: news about this channel in the internet if you: (.) *if you* prefer (0.8) use the internet . (.) and I think it's very very interesting ? because I (.) personally (.) I love history (.) and we can learn a lot about the men (.) the past and also what we the men are planning for our future (1.3) like go to the Mars (1.8) and the: (1.1) I (0.7) *I* really enjoyed the picture I think it's a good ad.

I saw a movie ? last week that I (0.5) and I really enjoyed the movie (0.6) the: *the* name is (.) was (0.9) **uhm** at first sight (1.1) and it's a story of a man (0.7) first it's *it's* a true story . (.) and it's a story of a man (.) that worked at a: spa (.) I don't know (0.6) the word in English for this one as a: (0.6) **uhm** (0.9) his job was massage (.) making massage (.) and the: (.) so he met a girl a woman (0.8) and they fall (.) in love (.) with each other (.) and she tried to help him to: (0.9) *to:* (0.5) *to* see again (1.2) **uhm** (1.0) so he: **uhm** (.) they move to New York (.) to try to find a: (0.8) *a* cure for (.) his problem . (1.) and: (1.7) I thought it was very: **uhm** (0.8) good movie because (.) they (1.1) *they* (0.8) it shows us **uh** (1.1) how the woman (1.3) helped him (.) because at first (0.5) **uhm** (1.5) **uh** it wasn't **uh** (.) it was very difficult for him because he (.) *he* lost her *sigh-* his sight (.) **uh** when he was a child . (0.5) so when he started to see again (.) he didn't know what *what* was a *a*: can of coke . (0.9) he had to learn everything again (0.8) and *sh:* (.) she *help-* she helped him a lot (0.5) but then (0.5) **uhm** he lost his sight again (0.8) and the (0.8) it was very sad (.) part of the story (0.5) and the end I'm not tell you (.) because you should see this movie (.) because I really enjoyed and I think you will enjoy it too .

Total time: Im53s03 **(113s)**
No. of words: 227/216

APPENDIX D
INSTRUCTIONS TO SUBJECTS

I. Speaking Span Test

Dear student,

Read these instructions carefully. They are about the task you will be asked to perform now. Ask as many questions as necessary!

The task you will perform now is a memory test. You will see sets of words on the screen. Each word in a set will appear alone on the middle line of the screen. The word will remain on the screen for **one** second only and then it will disappear from the screen. Ten milliseconds later, the next word of the set will appear. The sets begin with 2 words. Then a blank screen will be shown and you will have to produce two sentences: one containing the first word you saw and the other containing the second word you saw on the screen. You have to produce the sentences orally. When you finish producing the sentences, I will press "enter" and then a set with **three** words will be shown. Again, it will be one word at a time. For example, in a set with "book, child, cat", you will first see "book", then "child", then "cat". Then you will see a blank screen. Again, you will have to produce, orally, a sentence containing the word "book", another sentence containing the word "child", and, finally, a sentence containing the word "cat". When you are done, a set with **four** words will appear. Then, a set with **five** words will be shown. The last set has **six** words and you will have to produce six sentences, one sentence for each word in the set. Then, we will go back to two words again, then three, four, five, and six words. After that, we will go through the last sets of two, three, four, five, and six words. It is important to concentrate on and pay attention to these words because they will be shown for one second, only. The words are not related, so do not try to combine the sentences. It is not necessary. You must use the word in its exact order of presentation and in its form of presentation. Thus, in the example above, your first sentence has to contain the word "book" because that was the first word in the set, and it must be used in exactly the same way it appears. Your last sentence must contain the word "cat" since that was the last word in the set. You can produce both long and short sentences, but they must make sense on their own and should contain at least a subject and a verb and, of course, the word you are making the sentence for. You will first go through a practice trial. You can rehearse as many times as necessary. We will begin the test only when you decide you are comfortable enough.

II. Operation-word span test

The second task you will perform is another memory test. In this test, you will see sets of a math operation followed by a word, in English, on the middle line of the screen. You have to analyze the operation, decide if the result shown on the screen is True or False and memorize the word next to the operation. If the result is true, then press the letter "T" on the keyboard. If the result is false, then press the letter "F" on the keyboard. As soon as you press any of these letters, the operation and the word will disappear and another pair of operation-word will be immediately shown on the screen. Unlike the first task, there is no time pressure on this one, but you must remember that the operation and the word will be gone as soon as you press "T" or "F". Again, the sets will begin with 2 pairs of operation-word, one pair shown at a time. When you hear a beep and see a blank screen, you have to recall **the words** accompanying the operations. For example, in the following set of two pairs, you will first see:

6 X 3 = 18 BOOK

Then, you will decide if the answer is true or false, memorize the word "book", and press "T" or "F". As soon as you press any of these letters, the next pair will be shown:

2 / 2 = 1 CHILD

After deciding if the result of the operation is true or false and memorizing the word, you will hear a beep and see a blank screen. This will signal that you have to recall the words "book" and "child", in their exact order and form of presentation. Sets will increase from two to 6 pairs of operation-word, but this time, we can not predict the order in which they will be presented. For example, you may see a set of two pairs, then a set of 6 pairs, then one of 3 pairs, and so on. Like in the first task, you will be presented with three sets of two pairs, three of three pairs, three of four, three of five, and three of six. However, their order will be randomized. Again, you will go through a practice trial. You can practice as much as necessary and we will begin the test only when you decide you are comfortable enough.

III. Picture description task

The two tasks you will perform now aim at eliciting your oral production in English. The first task is a picture description task. You will be shown a picture and will have to describe the picture and make comments about it. Try to describe the picture as if I had not seen it. Besides describing the picture, you also have to give your opinion about the message the picture carries, if any. Try to speak as much as possible about the picture. You have to say as much as possible in **two minutes**. Before speaking, you can plan what to say and check any vocabulary difficulties you have. You can also write down any words and sentences you might want to use. You can check your notes quickly while speaking, but you can not speak as if you were reading them. The task will begin when you signal you are ready. Please, remember that you have to say as much as possible in two minutes. A stopwatch will be used to signal the beginning and end of the time allotted.

IV. Narrative task

This is the second task aimed at eliciting your oral production in English. Now, I want you to retell the story of a movie you have seen which you have liked or disliked. Besides retelling the story, you also have to state why you liked (or not) the movie. There are no time constraints on this task, but you should try to give as much information as possible. Again, you can plan what to say and check any vocabulary difficulties you have. You can also write down any words and sentences you might want to use. You can check your notes quickly while speaking, but you can not speak as if you were reading them. The task will begin when you signal you are ready.

APPENDIX E
RECRUITMENT SCRIPT

Dear___________,

This is an invitation to participate in my research! My name is Mailce Mota and I am from Brazil. I am writing a Ph.D. dissertation about oral fluency in English and would like to invite you to participate in my study. Your teacher, Patricia Maguire, who teaches speaking to you at the Minnesota English Center, gave me your name. I am inviting every student in your class, since all of you have the profile of the students I need – advanced students!

If you accept to participate in my study, we will need to meet 2 TIMES, for approximately 1 HOUR each time, during the quarter: the first time will be next week and the second and last time will be before your final exams. In each meeting, I will ask you to perform speaking tasks --such as describing a picture and telling a story-- and memory tasks.

Unfortunately, I can not pay you to participate in my study. But if you accept to help me, you will learn why speaking fluently in English is sometimes so hard! You will also help me to find out how to help other students learning English!

If you accept to participate, we'll need to meet next week for the first session. I will meet each of you individually, at 128 Klaeber Court. We can meet from Monday to Friday, from 1 p.m. to 4 p.m. As I said, you'll need to stay with me for 1 hour in this first session. If you can't come at any of these times, please let me know what good time for you would be. I'll be glad to arrange our meeting at you convenience.

Please email me and tell me:

a) If you accept to participate in the study,
b) The day of the week and the time you can meet me.
c) A phone number where I can contact you to remind you of our meeting

If you have ANY questions, please email me or call me at 612-617-0424.

I look forward to hearing from you. Thank you for your time and attention.

Mailce Mota
motax001@tc.umn.edu

Appendix F
Consent Form

CONSENT FORM

Title of study: Working memory capacity and L2 speech production

You are invited to be in a research study of the development of L2 speech production and working memory capacity. You were selected as a possible participant because you are an ESL student at the intermediate level and are in the process of developing your speech production in English. We ask you to read this form and ask any questions you may have before agreeing to be in the study.

This study is being conducted by **Mailce Borges Mota** (doctoral candidate in linguistics) under the supervision of Dr. Andrew Cohen, and the **Federal University of Santa Catarina, Brazil**, in conjunction with the **University of Minnesota**.

Background information:

The purpose of this study is to examine the relationship between working memory capacity and L2 speech production across time.

Procedures:

If you agree to be in this study, we would ask you to do the following things: (1) Perform memory tasks on a computer, (2) describe a picture and tell a story, and (3) answer 3 questionnaires. You will have to perform these tasks on two occasions: at the beginning and the end of the term. All three tasks will be performed in one session of approximately 1 hour, on both occasions.

Risks and benefits of being in the study:

There are no risks of being in the study. The benefits to participation, on the other hand, include your knowing and understanding better how your speech production and memory performance increase.

Confidentiality:

The records of this study will be kept private. In any sort of report we might publish, we will not include any information that will make it possible to identify a subject. Research records and tape recordings will be kept in a locked file. Only the principal investigator and the advisor of this study will have access to the records. The records will be erased in August 1999, when this study will be finished.

Voluntary nature of the study:

Your decision whether or not to participate will not affect your current or future relations with the University. If you decide to participate, you are free to withdraw at any time without affecting those relationships.

Contacts and questions:

The researchers conducting this study are Mailce Borges Mota (student researcher) and Dr. Andrew D. Cohen (advisor). You may ask any questions you have now. If you have questions later, you may contact Mailce B. Mota at 612-378-1761 or Dr. Andrew Cohen at 130 Klaeber Court, 320 16th Ave SE, Minneapolis, MN 55455. Phone: 612-624-3806.

You will be given a copy of this form to keep for your records.

Statement of Consent:

I have read the above information. I have asked questions and have received answers. I consent to participate in the study.

Signature ____ Date _________

Signature of Investigator ________________________________ Date _________

Date: ________

Code: ________

Appendix G
Biographical Information/Personal data

BIOGRAPHICAL INFORMATION/PERSONAL DATA

CONFIDENTIAL

1. Name: ____
   (Last) (First)

2. Email: ____________________________ Tele: __________

3. Dominant language: ______________ Country of origin: ____________

4. Other languages: ________________________________

5. Age: ________ 6.Gender: Female ( ) Male ( )

7. Occupation: __________

8. Education (check highest level):
( )Secondary ( )Jr. College ( )BA/BS ( )MA/MS ( )Law ( )Doctorate

9. How long have you been in the US (in months and years)?
____

10. How long will you stay after the Winter Quarter is over?
____

11. How long have you been studying English (in years)?
____

12. Besides this course in speaking skills, what other English courses are you taking or have taken?
____

Thank you for your interest in this research project!


THE HISTORY CHANNEL
THE OFFICIAL NETWORK OF EVERY MILLENNIUM
www.HistoryChannel.com

APPENDIX I - STRATEGY USE QUESTIONNAIRE

Subject: ____

Directions: Below you have a number of questions assessing the strategies you use to improve your speaking skills in English. Please answer these questions, checking **ONE** answer for each.

PART I - General

1-Do you speak your dominant language out of class?

() Very often () Often () Sometimes () Rarely () Never

2-Do you speak English with other nonnative students out of class?

() Very often () Often () Sometimes () Rarely () Never

3-Do you speak English with native speakers out of class?

() Very often () Often () Sometimes () Rarely () Never

4-Do you look for opportunities to speak English out of class?

() Very often () Often () Sometimes () Rarely () Never

5-Do you read (newspapers, magazines, books, brochures, pamphlets) with the primary objective of learning new words and structures in English?

() Very often () Often () Sometimes () Rarely () Never

6-Do you make an effort to use these new words and structures when you are speaking in English?

() Very often () Often () Sometimes () Rarely () Never

7-When listening to native speakers, do you observe the ways they express themselves (e.g. how they pause and hesitate, the vocabulary they use, their pronunciation and intonation)?

() Very often () Often () Sometimes () Rarely () Never

8-Do you make an effort to use these ways?

() Very often () Often () Sometimes () Rarely () Never

9-Do you watch TV or listen to the radio with the primary purpose of learning new words and structures to use when speaking?

() Very often () Often () Sometimes () Rarely () Never

10-Do you make an effort to use these words and structures learned from TV or radio when you are speaking?

() Very often () Often () Sometimes () Rarely () Never

11-Do you practice reading aloud to improve your pronunciation?

() Very often () Often () Sometimes () Rarely () Never

12-Do you repeat words or phrases to yourself to improve your pronunciation?

() Very often () Often () Sometimes () Rarely () Never

13-Do you watch TV/movies or listen to the radio/recordings with the primary purpose of improving your pronunciation and intonation?

() Very often () Often () Sometimes () Rarely () Never

14-Do you think in English while working and/or studying?

() Very often () Often () Sometimes () Rarely () Never

15-Do you speak to yourself in English, either silently or aloud?

() Very often () Often () Sometimes () Rarely () Never

Part II-Planning

1-Before starting a conversation, do you have an overall mental plan of what to say?

() Very often () Often () Sometimes () Rarely () Never

2-If you plan what to say do you translate all from your native language into English?

() Very often () Often () Sometimes () Rarely () Never

3-If you plan what to say do you translate specific words from your native language into English?

() Very often () Often () Sometimes () Rarely () Never

4-Do you write down full sentences or key words before speaking?

() Very often () Often () Sometimes () Rarely () Never

5-Do you practice silently what you will say before speaking?

() Very often () Often () Sometimes () Rarely () Never

6-Do you practice pronunciation of specific words before speaking?

() Very often () Often () Sometimes () Rarely () Never

7-Do you try to predict any difficulties you might have?

() Very often () Often () Sometimes () Rarely () Never

8-Do you try to predict what the other person will say?

() Very often () Often () Sometimes () Rarely () Never

Part III-Monitoring

1-When you are speaking in English do you pay attention to your grammar?

() Very often () Often () Sometimes () Rarely () Never

2-Do you try to correct your grammar as you speak?

() Very often () Often () Sometimes () Rarely () Never

3-When you are speaking in English do you pay attention to your pronunciation?

() Very often () Often () Sometimes () Rarely () Never

4-Do you try to correct your pronunciation as you speak?

() Very often () Often () Sometimes () Rarely () Never

5-When you are speaking and can't remember a word, do you make it up?

() Very often () Often () Sometimes () Rarely () Never

6-When you are speaking and can't remember a word, do you substitute it with another word or phrase in English you know?

() Very often () Often () Sometimes () Rarely () Never

7-When you are speaking and can't remember a word, do you use gestures?

() Very often () Often () Sometimes () Rarely () Never

Thank you!

Strategy Use Questionnaire – Speech production task 1: description of picture

Subject: ____

Directions: The objective of this questionnaire is to assess the strategies you used when performing the picture description task. Rate on a scale of 1 to 5 the extent to which you used each strategy below.

1= not at all 2=a little bit 3=part of the time 4=a lot 5=extensively

Planning

1-Before starting to describe and comment on the pictures, I had an overall plan of what I would say.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

2-When planning, I translated all from my native language.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

3-When planning, I translated only some specific words from my native language.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

4-I planned what to say in English, only.

1= not at all ( ) | 2=a little bit ( ) | 3=part of the time ( ) | 4=a lot ( ) | 5=extensively ( )

5-I wrote what I would say in full sentences before describing the pictures.

1= not at all ( ) | 2=a little bit ( ) | 3=part of the time ( ) | 4=a lot ( ) | 5=extensively ( )

6-I wrote key vocabulary words I wanted/needed to use.

1= not at all ( ) | 2=a little bit ( ) | 3=part of the time ( ) | 4=a lot ( ) | 5=extensively ( )

7-I practiced everything I would say silently before I started describing and commenting on the pictures.

1= not at all ( ) | 2=a little bit ( ) | 3=part of the time ( ) | 4=a lot ( ) | 5=extensively ( )

8-I practiced the pronunciation of specific words before I began speaking.

1= not at all ( ) | 2=a little bit ( ) | 3=part of the time ( ) | 4=a lot ( ) | 5=extensively ( )

9-I tried to predict any difficulties I might have.

1= not at all ( ) | 2=a little bit ( ) | 3=part of the time ( ) | 4=a lot ( ) | 5=extensively ( )

10-I didn't do any special preparation.

1= not at all ( ) | 2=a little bit ( ) | 3=part of the time ( ) | 4=a lot ( ) | 5=extensively ( )

Part II – Monitoring

1-I paid attention to my grammar as I was speaking.

1= not at all ( ) | 2=a little bit ( ) | 3=part of the time ( ) | 4=a lot ( ) | 5=extensively ( )

2-I tried to correct my grammar as I was speaking.

1= not at all ( ) | 2=a little bit ( ) | 3=part of the time ( ) | 4=a lot ( ) | 5=extensively ( )

3-I paid attention to my pronunciation.

1= not at all ( ) | 2=a little bit ( ) | 3=part of the time ( ) | 4=a lot ( ) | 5=extensively ( )

4-I tried to correct my pronunciation as I was speaking.

1= not at all ( ) | 2=a little bit ( ) | 3=part of the time ( ) | 4=a lot ( ) | 5=extensively ( )

5-When I couldn't remember a word, I substituted it with another word or phrase I know.

1= not at all | 2=a little bit | 3=part of the time | 4=a lot | 5=extensively

( )	( )	( )	( )	( )

6-When I couldn't remember a word, I made up a word.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

7-When I couldn't remember a word, I just skipped that part of the description.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

8-I tried to stick to my initial plan of what to say.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

9-I altered my plan as I was speaking.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

10-I didn't pay much attention to how I was speaking. I concentrated on what I was saying.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

Thank you!

Strategy questionnaire – Speech production task 2: narrative

Subject: ____

Directions: The objective of this questionnaire is to assess the strategies you used when performing the narrative task. Rate on a scale of I to 5 the extent to which you used each strategy below.

1= not at all 2=a little bit 3=part of the time 4=a lot 5=extensively

Planning

1-Before starting to narrate, I had an overall plan of what I would say.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

2-When planning, I translated all from my native language.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

3-When planning, I translated only some specific words from my dominant language.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

4-I planned what to say in English, only.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

5-I wrote what I would say in full sentences before start narrating.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

6-I wrote key vocabulary words I wanted/needed to use.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

7-I practiced everything I would say silently before I started narrating.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

8-I practiced the pronunciation of specific words before I began speaking.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

9-I tried to predict any difficulties I might have.

I= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

10-I didn't do any special preparation.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

Part II – Monitoring

1-I paid attention to my grammar as I was speaking.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

2-I tried to correct my grammar as I was speaking.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

3-I paid attention to my pronunciation.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

4-I tried to correct my pronunciation as I was speaking.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

5-When I couldn't remember a word, I substituted it with another word or phrase I know.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

6-When I couldn't remember a word, I made up a word.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

7-When I couldn't remember a word, I just skipped that part of the narration.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

8-I tried to stick to my initial plan of what to say.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

9-I altered my plan as I was speaking.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

10-I didn't pay much attention to how I was speaking. I concentrated on what I was saying.

1= not at all	2=a little bit	3=part of the time	4=a lot	5=extensively
( )	( )	( )	( )	( )

Thank you!

APPENDIX J – COMPLEXITY SCORES (raw)

Participant	Description	Narrative
1	4	6
2	1	5
3	4	2
4	3	3
5	2	9
6	3	13
7	10	8
8	7	15
9	7	10
10	9	16
11	7	18
12	10	9
13	16	17

APPENDIX K – LEXICAL DENSITY SCORES (raw) - Description

Participant	Grammatical Items	Lexical Items	Total	Percentag e
1	26.5	38	64.5	58.91
2	30	35.5	65	54.61
3	31	49	80	61.25
4	38.5	48	86.5	55.49
5	41.5	33.5	75	44.66
6	36	38	74	51.35
7	39	37.5	76.50	49.01
8	54	51	105	48.57
9	51	58.5	109.5	53.42
10	38	36.5	74.5	48.99
11	42	39	81	48.14
12	47	47	94	50
13	72	62.5	134.5	46.46

Narrative

Participant	Grammatical Items	Lexical Items	Total	Percentage
1	25	28	53	52.83
2	29	50	79	63.29
3	25	33.5	58.5	57.26
4	40.05	44	84.5	52.07
5	37.5	36.5	74	49.32
6	40	42.5	82.5	51.51
7	45.5	36	81.5	44.17
8	67	67.5	134.5	50.18
9	60	71	131	54.19
10	91	66	157	42.03
11	58	51	109	47.78
12	54.5	64.5	119	54.20
13	68.5	62.5	131	47.70

APPENDIX L

DESCRIPTIVE AND INFERENTIAL STATISTICS

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
SSTS	13	11.00	47.00	24.5385	10.8214
OWST	13	54.00	60.00	58.0769	1.7541
SRU	13	54.91	117.32	82.7600	19.0602
SRP	13	52.46	111.42	79.0969	18.6067
SPPM	13	7.99	22.12	18.0500	4.5132
HPM	13	4.15	30.49	18.2438	8.4238
MLR	13	2.45	5.09	3.2331	.6982
ACC100	13	1.55	19.44	9.6577	5.7554
SRU2	13	42.99	122.49	95.2177	22.5715
SRP2	13	38.49	114.99	91.4600	21.9250
SPPM2	13	6.99	22.05	15.5923	4.0864
HPM2	13	8.25	36.99	21.7969	7.6097
MLR2	13	1.70	4.76	3.5208	.8080
ACC1002	13	1.68	15.11	6.6200	3.8445
REVLD1	13	44.66	61.25	51.6046	4.8814
REVLD2	13	42.03	63.29	51.1946	5.5843
COMPPM	13	.59	8.57	3.8131	2.6160
COMPPM2	13	1.76	9.02	5.7246	2.5156
Valid N (listwise)	13				

Correlations

		SSTS	SRU	SRP	MLR	SPPM	HPM
SSTS	Pearson Correlation	1.000	.732**	.724**	.706**	-.432	.203
	Sig. (1-tailed)	.	.002	.003	.003	.070	.253
	N	13	13	13	13	13	13
SRU	Pearson Correlation	.732**	1.000	.994**	.839**	-.687**	.539*
	Sig. (1-tailed)	.002	.	.000	.000	.005	.029
	N	13	13	13	13	13	13
SRP	Pearson Correlation	.724**	.994**	1.000	.824**	-.693**	.523*
	Sig. (1-tailed)	.003	.000	.	.000	.004	.033
	N	13	13	13	13	13	13
MLR	Pearson Correlation	.706**	.839**	.824**	1.000	-.573*	.213
	Sig. (1-tailed)	.003	.000	.000	.	.020	.243
	N	13	13	13	13	13	13
SPPM	Pearson Correlation	-.432	-.687**	-.693**	-.573*	1.000	-.681**
	Sig. (1-tailed)	.070	.005	.004	.020	.	.005
	N	13	13	13	13	13	13
HPM	Pearson Correlation	.203	.539*	.523*	.213	-.681**	1.000
	Sig. (1-tailed)	.253	.029	.033	.243	.005	.
	N	13	13	13	13	13	13
ACC100	Pearson Correlation	-.536*	-.610*	-.663**	-.519*	.320	.080
	Sig. (1-tailed)	.029	.013	.007	.035	.143	.398
	N	13	13	13	13	13	13
COMPPM	Pearson Correlation	.766**	.795**	.799**	.719**	-.509*	.290
	Sig. (1-tailed)	.001	.001	.001	.003	.038	.168
	N	13	13	13	13	13	13
REVLD1	Pearson Correlation	-.578*	-.570*	-.602*	-.604*	.253	.148
	Sig. (1-tailed)	.019	.021	.015	.014	.202	.315
	N	13	13	13	13	13	13
SRU2	Pearson Correlation	.698**	.852**	.843**	.679**	-.602*	.463
	Sig. (1-tailed)	.004	.000	.000	.005	.015	.056
	N	13	13	13	13	13	13
SRP2	Pearson Correlation	.681**	.848**	.846**	.671**	-.586*	.436
	Sig. (1-tailed)	.005	.000	.000	.006	.018	.068
	N	13	13	13	13	13	13
MLR2	Pearson Correlation	.621*	.680**	.686**	.616*	-.434	.193
	Sig. (1-tailed)	.012	.005	.005	.013	.069	.264
	N	13	13	13	13	13	13
SPPM2	Pearson Correlation	-.224	-.297	-.319	-.218	.767**	-.390
	Sig. (1-tailed)	.231	.162	.144	.237	.001	.094
	N	13	13	13	13	13	13
HPM2	Pearson Correlation	.429	.589*	.554*	.413	-.757**	.673**
	Sig. (1-tailed)	.072	.017	.025	.080	.001	.006
	N	13	13	13	13	13	13
COMPPM2	Pearson Correlation	.542*	.552*	.579*	.640**	-.395	-.236
	Sig. (1-tailed)	.028	.025	.019	.009	.091	.219
	N	13	13	13	13	13	13
REVLD2	Pearson Correlation	-.395	-.553*	-.563*	-.524*	.308	.035
	Sig. (1-tailed)	.091	.025	.023	.033	.153	.455
	N	13	13	13	13	13	13

Correlations

		ACC100	COMPPM	REVLD1	SRU2	SRP2	MLR2
SSTS	Pearson Correlation	-.536*	.766**	-.578*	.698**	.681**	.621*
	Sig. (1-tailed)	.029	.001	.019	.004	.005	.012
	N	13	13	13	13	13	13
SRU	Pearson Correlation	-.610*	.795**	-.570*	.852**	.848**	.680**
	Sig. (1-tailed)	.013	.001	.021	.000	.000	.005
	N	13	13	13	13	13	13
SRP	Pearson Correlation	-.663**	.799**	-.602*	.843**	.846**	.686**
	Sig. (1-tailed)	.007	.001	.015	.000	.000	.005
	N	13	13	13	13	13	13
MLR	Pearson Correlation	-.519*	.719**	-.604*	.679**	.671**	.616*
	Sig. (1-tailed)	.035	.003	.014	.005	.006	.013
	N	13	13	13	13	13	13
SPPM	Pearson Correlation	.320	-.509*	.253	-.602*	-.586*	-.434
	Sig. (1-tailed)	.143	.038	.202	.015	.018	.069
	N	13	13	13	13	13	13
HPM	Pearson Correlation	.080	.290	.148	.463	.436	.193
	Sig. (1-tailed)	.398	.168	.315	.056	.068	.264
	N	13	13	13	13	13	13
ACC100	Pearson Correlation	1.000	-.637**	.760**	-.626*	-.662**	-.660**
	Sig. (1-tailed)	.	.010	.001	.011	.007	.007
	N	13	13	13	13	13	13
COMPPM	Pearson Correlation	-.637**	1.000	-.487*	.737**	.720**	.669**
	Sig. (1-tailed)	.010	.	.046	.002	.003	.006
	N	13	13	13	13	13	13
REVLD1	Pearson Correlation	.760**	-.487*	1.000	-.623*	-.632*	-.560*
	Sig. (1-tailed)	.001	.046	.	.012	.010	.023
	N	13	13	13	13	13	13
SRU2	Pearson Correlation	-.626*	.737**	-.623*	1.000	.995**	.863**
	Sig. (1-tailed)	.011	.002	.012	.	.000	.000
	N	13	13	13	13	13	13
SRP2	Pearson Correlation	-.662**	.720**	-.632*	.995**	1.000	.887**
	Sig. (1-tailed)	.007	.003	.010	.000	.	.000
	N	13	13	13	13	13	13
MLR2	Pearson Correlation	-.660**	.669**	-.560*	.863**	.887**	1.000
	Sig. (1-tailed)	.007	.006	.023	.000	.000	.
	N	13	13	13	13	13	13
SPPM2	Pearson Correlation	.224	-.244	.257	-.515*	-.512*	-.447
	Sig. (1-tailed)	.231	.211	.198	.036	.037	.063
	N	13	13	13	13	13	13
HPM2	Pearson Correlation	-.152	.460	-.160	.610*	.559*	.215
	Sig. (1-tailed)	.310	.057	.301	.013	.024	.240
	N	13	13	13	13	13	13
COMPPM2	Pearson Correlation	-.728**	.602*	-.833**	.533*	.540*	.546*
	Sig. (1-tailed)	.002	.015	.000	.030	.028	.027
	N	13	13	13	13	13	13
REVLD2	Pearson Correlation	.721**	-.669**	.618*	-.443	-.424	-.348
	Sig. (1-tailed)	.003	.006	.012	.065	.074	.122
	N	13	13	13	13	13	13

Correlations

		SPPM2	HPM2	COMPPM2	REVLD2
SSTS	Pearson Correlation	-.224	.429	.542*	-.395
	Sig. (1-tailed)	.231	.072	.028	.091
	N	13	13	13	13
SRU	Pearson Correlation	-.297	.589*	.552*	-.553*
	Sig. (1-tailed)	.162	.017	.025	.025
	N	13	13	13	13
SRP	Pearson Correlation	-.319	.554*	.579*	-.563*
	Sig. (1-tailed)	.144	.025	.019	.023
	N	13	13	13	13
MLR	Pearson Correlation	-.218	.413	.640**	-.524*
	Sig. (1-tailed)	.237	.080	.009	.033
	N	13	13	13	13
SPPM	Pearson Correlation	.767**	-.757**	-.395	.308
	Sig. (1-tailed)	.001	.001	.091	.153
	N	13	13	13	13
HPM	Pearson Correlation	-.390	.673**	-.236	.035
	Sig. (1-tailed)	.094	.006	.219	.455
	N	13	13	13	13
ACC100	Pearson Correlation	.224	-.152	-.728**	.721**
	Sig. (1-tailed)	.231	.310	.002	.003
	N	13	13	13	13
COMPPM	Pearson Correlation	-.244	.460	.602*	-.669**
	Sig. (1-tailed)	.211	.057	.015	.006
	N	13	13	13	13
REVLD1	Pearson Correlation	.257	-.160	-.833**	.618*
	Sig. (1-tailed)	.198	.301	.000	.012
	N	13	13	13	13
SRU2	Pearson Correlation	-.515*	.610*	.533*	-.443
	Sig. (1-tailed)	.036	.013	.030	.065
	N	13	13	13	13
SRP2	Pearson Correlation	-.512*	.559*	.540*	-.424
	Sig. (1-tailed)	.037	.024	.028	.074
	N	13	13	13	13
MLR2	Pearson Correlation	-.447	.215	.546*	-.348
	Sig. (1-tailed)	.063	.240	.027	.122
	N	13	13	13	13
SPPM2	Pearson Correlation	1.000	-.580*	-.364	.016
	Sig. (1-tailed)	.	.019	.110	.480
	N	13	13	13	13
HPM2	Pearson Correlation	-.580*	1.000	.187	-.288
	Sig. (1-tailed)	.019	.	.270	.170
	N	13	13	13	13
COMPPM2	Pearson Correlation	-.364	.187	1.000	-.697**
	Sig. (1-tailed)	.110	.270	.	.004
	N	13	13	13	13
REVLD2	Pearson Correlation	.016	-.288	-.697**	1.000
	Sig. (1-tailed)	.480	.170	.004	.
	N	13	13	13	13

**. Correlation is significant at the 0.01 level (1-tailed).

*. Correlation is significant at the 0.05 level (1-tailed).

Correlations

		SSTS	SRU	SRP	MLR	SPPM	HPM
SSTS	Pearson Correlation	1.000	.732**	.721**	.800**	-.604*	.118
	Sig. (1-tailed)	.	.005	.006	.002	.024	.365
	N	11	11	11	11	11	11
SRU	Pearson Correlation	.732**	1.000	.992**	.836**	-.688**	.576*
	Sig. (1-tailed)	.005	.	.000	.001	.010	.032
	N	11	11	11	11	11	11
SRP	Pearson Correlation	.721**	.992**	1.000	.831**	-.694**	.548*
	Sig. (1-tailed)	.006	.000	.	.001	.009	.041
	N	11	11	11	11	11	11
MLR	Pearson Correlation	.800**	.836**	.831**	1.000	-.607*	.255
	Sig. (1-tailed)	.002	.001	.001	.	.024	.225
	N	11	11	11	11	11	11
SPPM	Pearson Correlation	-.604*	-.688**	-.694**	-.607*	1.000	-.708**
	Sig. (1-tailed)	.024	.010	.009	.024	.	.007
	N	11	11	11	11	11	11
HPM	Pearson Correlation	.118	.576*	.548*	.255	-.708**	1.000
	Sig. (1-tailed)	.365	.032	.041	.225	.007	.
	N	11	11	11	11	11	11
ACC100	Pearson Correlation	-.726**	-.646*	-.702**	-.738**	.316	.114
	Sig. (1-tailed)	.006	.016	.008	.005	.172	.370
	N	11	11	11	11	11	11
COMPPM	Pearson Correlation	.633*	.709**	.718**	.650*	-.506	.255
	Sig. (1-tailed)	.018	.007	.006	.015	.056	.225
	N	11	11	11	11	11	11
REVLD1	Pearson Correlation	-.610*	-.490	-.531*	-.646*	.197	.212
	Sig. (1-tailed)	.023	.063	.046	.016	.281	.266
	N	11	11	11	11	11	11
SRU2	Pearson Correlation	.694**	.851**	.832**	.804**	-.617*	.442
	Sig. (1-tailed)	.009	.000	.001	.001	.022	.087
	N	11	11	11	11	11	11
SRP2	Pearson Correlation	.682*	.852**	.841**	.809**	-.598*	.410
	Sig. (1-tailed)	.010	.000	.001	.001	.026	.105
	N	11	11	11	11	11	11
MLR2	Pearson Correlation	.517	.618*	.623*	.680*	-.424	.137
	Sig. (1-tailed)	.052	.021	.020	.011	.097	.344
	N	11	11	11	11	11	11
SPPM2	Pearson Correlation	-.451	-.396	-.414	-.459	.825**	-.398
	Sig. (1-tailed)	.082	.114	.103	.078	.001	.113
	N	11	11	11	11	11	11
HPM2	Pearson Correlation	.622*	.637*	.588*	.567*	-.772**	.670*
	Sig. (1-tailed)	.021	.017	.029	.034	.003	.012
	N	11	11	11	11	11	11
COMPPM2	Pearson Correlation	.734**	.463	.505	.581*	-.316	-.286
	Sig. (1-tailed)	.005	.076	.057	.030	.172	.197
	N	11	11	11	11	11	11
REVLD2	Pearson Correlation	-.689**	-.579*	-.590*	-.626*	.262	.033
	Sig. (1-tailed)	.009	.031	.028	.020	.218	.461
	N	11	11	11	11	11	11

Correlations

		ACC100	COMPPM	REVLD1	SRU2	SRP2	MLR2
SSTS	Pearson Correlation	-.726**	.633*	-.610*	.694**	.682*	.517
	Sig. (1-tailed)	.006	.018	.023	.009	.010	.052
	N	11	11	11	11	11	11
SRU	Pearson Correlation	-.646*	.709**	-.490	.851**	.852**	.618*
	Sig. (1-tailed)	.016	.007	.063	.000	.000	.021
	N	11	11	11	11	11	11
SRP	Pearson Correlation	-.702**	.718**	-.531*	.832**	.841**	.623*
	Sig. (1-tailed)	.008	.006	.046	.001	.001	.020
	N	11	11	11	11	11	11
MLR	Pearson Correlation	-.738**	.650*	-.646*	.804**	.809**	.680*
	Sig. (1-tailed)	.005	.015	.016	.001	.001	.011
	N	11	11	11	11	11	11
SPPM	Pearson Correlation	.316	-.506	.197	-.617*	-.598*	-.424
	Sig. (1-tailed)	.172	.056	.281	.022	.026	.097
	N	11	11	11	11	11	11
HPM	Pearson Correlation	.114	.255	.212	.442	.410	.137
	Sig. (1-tailed)	.370	.225	.266	.087	.105	.344
	N	11	11	11	11	11	11
ACC100	Pearson Correlation	1.000	-.686**	.762**	-.619*	-.656*	-.658*
	Sig. (1-tailed)	.	.010	.003	.021	.014	.014
	N	11	11	11	11	11	11
COMPPM	Pearson Correlation	-.686**	1.000	-.376	.672*	.657*	.570*
	Sig. (1-tailed)	.010	.	.127	.012	.014	.034
	N	11	11	11	11	11	11
REVLD1	Pearson Correlation	.762**	-.376	1.000	-.568*	-.582*	-.493
	Sig. (1-tailed)	.003	.127	.	.034	.030	.062
	N	11	11	11	11	11	11
SRU2	Pearson Correlation	-.619*	.672*	-.568*	1.000	.995**	.831**
	Sig. (1-tailed)	.021	.012	.034	.	.000	.001
	N	11	11	11	11	11	11
SRP2	Pearson Correlation	-.656*	.657*	-.582*	.995**	1.000	.863**
	Sig. (1-tailed)	.014	.014	.030	.000	.	.000
	N	11	11	11	11	11	11
MLR2	Pearson Correlation	-.658*	.570*	-.493	.831**	.863**	1.000
	Sig. (1-tailed)	.014	.034	.062	.001	.000	.
	N	11	11	11	11	11	11
SPPM2	Pearson Correlation	.231	-.340	.291	-.584*	-.575*	-.513
	Sig. (1-tailed)	.247	.153	.193	.030	.032	.053
	N	11	11	11	11	11	11
HPM2	Pearson Correlation	-.135	.494	-.129	.624*	.565*	.184
	Sig. (1-tailed)	.346	.061	.353	.020	.035	.294
	N	11	11	11	11	11	11
COMPPM2	Pearson Correlation	-.774**	.579*	-.843**	.523*	.535*	.539*
	Sig. (1-tailed)	.003	.031	.001	.049	.045	.043
	N	11	11	11	11	11	11
REVLD2	Pearson Correlation	.751**	-.790**	.632*	-.477	-.455	-.371
	Sig. (1-tailed)	.004	.002	.018	.069	.080	.130
	N	11	11	11	11	11	11

Correlations

		SPPM2	HPM2	COMPPM2	REVLD2
SSTS	Pearson Correlation	-.451	.622*	.734**	-.689**
	Sig. (1-tailed)	.082	.021	.005	.009
	N	11	11	11	11
SRU	Pearson Correlation	-.396	.637*	.463	-.579*
	Sig. (1-tailed)	.114	.017	.076	.031
	N	11	11	11	11
SRP	Pearson Correlation	-.414	.588*	.505	-.590*
	Sig. (1-tailed)	.103	.029	.057	.028
	N	11	11	11	11
MLR	Pearson Correlation	-.459	.567*	.581*	-.626*
	Sig. (1-tailed)	.078	.034	.030	.020
	N	11	11	11	11
SPPM	Pearson Correlation	.825**	-.772**	-.316	.262
	Sig. (1-tailed)	.001	.003	.172	.218
	N	11	11	11	11
HPM	Pearson Correlation	-.398	.670*	-.286	.033
	Sig. (1-tailed)	.113	.012	.197	.461
	N	11	11	11	11
ACC100	Pearson Correlation	.231	-.135	-.774**	.751**
	Sig. (1-tailed)	.247	.346	.003	.004
	N	11	11	11	11
COMPPM	Pearson Correlation	-.340	.494	.579*	-.790**
	Sig. (1-tailed)	.153	.061	.031	.002
	N	11	11	11	11
REVLD1	Pearson Correlation	.291	-.129	-.843**	.632*
	Sig. (1-tailed)	.193	.353	.001	.018
	N	11	11	11	11
SRU2	Pearson Correlation	-.584*	.624*	.523*	-.477
	Sig. (1-tailed)	.030	.020	.049	.069
	N	11	11	11	11
SRP2	Pearson Correlation	-.575*	.565*	.535*	-.455
	Sig. (1-tailed)	.032	.035	.045	.080
	N	11	11	11	11
MLR2	Pearson Correlation	-.513	.184	.539*	-.371
	Sig. (1-tailed)	.053	.294	.043	.130
	N	11	11	11	11
SPPM2	Pearson Correlation	1.000	-.590*	-.433	.035
	Sig. (1-tailed)	.	.028	.092	.460
	N	11	11	11	11
HPM2	Pearson Correlation	-.590*	1.000	.169	-.287
	Sig. (1-tailed)	.028	.	.310	.196
	N	11	11	11	11
COMPPM2	Pearson Correlation	-.433	.169	1.000	-.684*
	Sig. (1-tailed)	.092	.310	.	.010
	N	11	11	11	11
REVLD2	Pearson Correlation	.035	-.287	-.684*	1.000
	Sig. (1-tailed)	.460	.196	.010	.
	N	11	11	11	11

**. Correlation is significant at the 0.01 level (1-tailed).

*. Correlation is significant at the 0.05 level (1-tailed).

Correlations

			SSTS	SRU	SRP	MLR
Spearman's rho	SSTS	Correlation Coefficient	1.000	.831**	.831**	.757**
		Sig. (1-tailed)	.	.000	.000	.001
		N	13	13	13	13
	SRU	Correlation Coefficient	.831**	1.000	.984**	.819**
		Sig. (1-tailed)	.000	.	.000	.000
		N	13	13	13	13
	SRP	Correlation Coefficient	.831**	.984**	1.000	.797**
		Sig. (1-tailed)	.000	.000	.	.001
		N	13	13	13	13
	MLR	Correlation Coefficient	.757**	.819**	.797**	1.000
		Sig. (1-tailed)	.001	.000	.001	.
		N	13	13	13	13
	SPPM	Correlation Coefficient	-.710**	-.758**	-.720**	-.709**
		Sig. (1-tailed)	.003	.001	.003	.003
		N	13	13	13	13
	HPM	Correlation Coefficient	.261	.560*	.555*	.302
		Sig. (1-tailed)	.194	.023	.025	.158
		N	13	13	13	13
	ACC100	Correlation Coefficient	-.674**	-.577*	-.643**	-.670**
		Sig. (1-tailed)	.006	.019	.009	.006
		N	13	13	13	13
	COMPPM	Correlation Coefficient	.806**	.786**	.791**	.659**
		Sig. (1-tailed)	.000	.001	.001	.007
		N	13	13	13	13
	REVLD1	Correlation Coefficient	-.685**	-.621*	-.637**	-.632*
		Sig. (1-tailed)	.005	.012	.010	.010
		N	13	13	13	13
	SRU2	Correlation Coefficient	.796**	.916**	.913**	.718**
		Sig. (1-tailed)	.001	.000	.000	.003
		N	13	13	13	13
	SRP2	Correlation Coefficient	.823**	.940**	.929**	.753**
		Sig. (1-tailed)	.000	.000	.000	.001
		N	13	13	13	13
	MLR2	Correlation Coefficient	.741**	.757**	.737**	.630*
		Sig. (1-tailed)	.002	.001	.002	.011
		N	13	13	13	13
	SPPM2	Correlation Coefficient	-.490*	-.505*	-.451	-.423
		Sig. (1-tailed)	.045	.039	.061	.075
		N	13	13	13	13
	HPM2	Correlation Coefficient	.564*	.571*	.538*	.599*
		Sig. (1-tailed)	.022	.021	.029	.015
		N	13	13	13	13
	COMPPM2	Correlation Coefficient	.746**	.659**	.659**	.692**
		Sig. (1-tailed)	.002	.007	.007	.004
		N	13	13	13	13
	REVLD2	Correlation Coefficient	-.490*	-.451	-.500*	-.588*
		Sig. (1-tailed)	.045	.061	.041	.017
		N	13	13	13	13

Correlations

			SPPM	HPM	ACC100	COMPPM
Spearman's rho	SSTS	Correlation Coefficient	-.710**	.261	-.674**	.806**
		Sig. (1-tailed)	.003	.194	.006	.000
		N	13	13	13	13
	SRU	Correlation Coefficient	-.758**	.560*	-.577*	.786**
		Sig. (1-tailed)	.001	.023	.019	.001
		N	13	13	13	13
	SRP	Correlation Coefficient	-.720**	.555*	-.643**	.791**
		Sig. (1-tailed)	.003	.025	.009	.001
		N	13	13	13	13
	MLR	Correlation Coefficient	-.709**	.302	-.670**	.659**
		Sig. (1-tailed)	.003	.158	.006	.007
		N	13	13	13	13
	SPPM	Correlation Coefficient	1.000	-.676**	.275	-.648**
		Sig. (1-tailed)	.	.006	.182	.008
		N	13	13	13	13
	HPM	Correlation Coefficient	-.676**	1.000	.110	.242
		Sig. (1-tailed)	.006	.	.360	.213
		N	13	13	13	13
	ACC100	Correlation Coefficient	.275	.110	1.000	-.659**
		Sig. (1-tailed)	.182	.360	.	.007
		N	13	13	13	13
	COMPPM	Correlation Coefficient	-.648**	.242	-.659**	1.000
		Sig. (1-tailed)	.008	.213	.007	.
		N	13	13	13	13
	REVLD1	Correlation Coefficient	.280	.121	.747**	-.484*
		Sig. (1-tailed)	.177	.347	.002	.047
		N	13	13	13	13
	SRU2	Correlation Coefficient	-.765**	.589*	-.580*	.754**
		Sig. (1-tailed)	.001	.017	.019	.001
		N	13	13	13	13
	SRP2	Correlation Coefficient	-.780**	.549*	-.610*	.786**
		Sig. (1-tailed)	.001	.026	.013	.001
		N	13	13	13	13
	MLR2	Correlation Coefficient	-.627*	.248	-.525*	.743**
		Sig. (1-tailed)	.011	.207	.033	.002
		N	13	13	13	13
	SPPM2	Correlation Coefficient	.753**	-.445	.159	-.302
		Sig. (1-tailed)	.001	.064	.302	.158
		N	13	13	13	13
	HPM2	Correlation Coefficient	-.797**	.681**	-.187	.352
		Sig. (1-tailed)	.001	.005	.271	.119
		N	13	13	13	13
	COMPPM2	Correlation Coefficient	-.527*	-.088	-.698**	.687**
		Sig. (1-tailed)	.032	.388	.004	.005
		N	13	13	13	13
	REVLD2	Correlation Coefficient	.335	.132	.775**	-.588*
		Sig. (1-tailed)	.131	.334	.001	.017
		N	13	13	13	13

Correlations

			REVLD1	SRU2	SRP2	MLR2
Spearman's rho	SSTS	Correlation Coefficient	-.685**	.796**	.823**	.741**
		Sig. (1-tailed)	.005	.001	.000	.002
		N	13	13	13	13
	SRU	Correlation Coefficient	-.621*	.916**	.940**	.757**
		Sig. (1-tailed)	.012	.000	.000	.001
		N	13	13	13	13
	SRP	Correlation Coefficient	-.637**	.913**	.929**	.737**
		Sig. (1-tailed)	.010	.000	.000	.002
		N	13	13	13	13
	MLR	Correlation Coefficient	-.632*	.718**	.753**	.630*
		Sig. (1-tailed)	.010	.003	.001	.011
		N	13	13	13	13
	SPPM	Correlation Coefficient	.280	-.765**	-.780**	-.627*
		Sig. (1-tailed)	.177	.001	.001	.011
		N	13	13	13	13
	HPM	Correlation Coefficient	.121	.589*	.549*	.248
		Sig. (1-tailed)	.347	.017	.026	.207
		N	13	13	13	13
	ACC100	Correlation Coefficient	.747**	-.580*	-.610*	-.525*
		Sig. (1-tailed)	.002	.019	.013	.033
		N	13	13	13	13
	COMPPM	Correlation Coefficient	-.484*	.754**	.786**	.743**
		Sig. (1-tailed)	.047	.001	.001	.002
		N	13	13	13	13
	REVLD1	Correlation Coefficient	1.000	-.547*	-.599*	-.421
		Sig. (1-tailed)	.	.026	.015	.076
		N	13	13	13	13
	SRU2	Correlation Coefficient	-.547*	1.000	.993**	.780**
		Sig. (1-tailed)	.026	.	.000	.001
		N	13	13	13	13
	SRP2	Correlation Coefficient	-.599*	.993**	1.000	.784**
		Sig. (1-tailed)	.015	.000	.	.001
		N	13	13	13	13
	MLR2	Correlation Coefficient	-.421	.780**	.784**	1.000
		Sig. (1-tailed)	.076	.001	.001	.
		N	13	13	13	13
	SPPM2	Correlation Coefficient	.220	-.523*	-.555*	-.462
		Sig. (1-tailed)	.235	.033	.025	.056
		N	13	13	13	13
	HPM2	Correlation Coefficient	-.187	.613*	.615*	.217
		Sig. (1-tailed)	.271	.013	.013	.238
		N	13	13	13	13
	COMPPM2	Correlation Coefficient	-.852**	.591*	.659**	.602*
		Sig. (1-tailed)	.000	.017	.007	.015
		N	13	13	13	13
	REVLD2	Correlation Coefficient	.720**	-.492*	-.533*	-.399
		Sig. (1-tailed)	.003	.044	.030	.088
		N	13	13	13	13

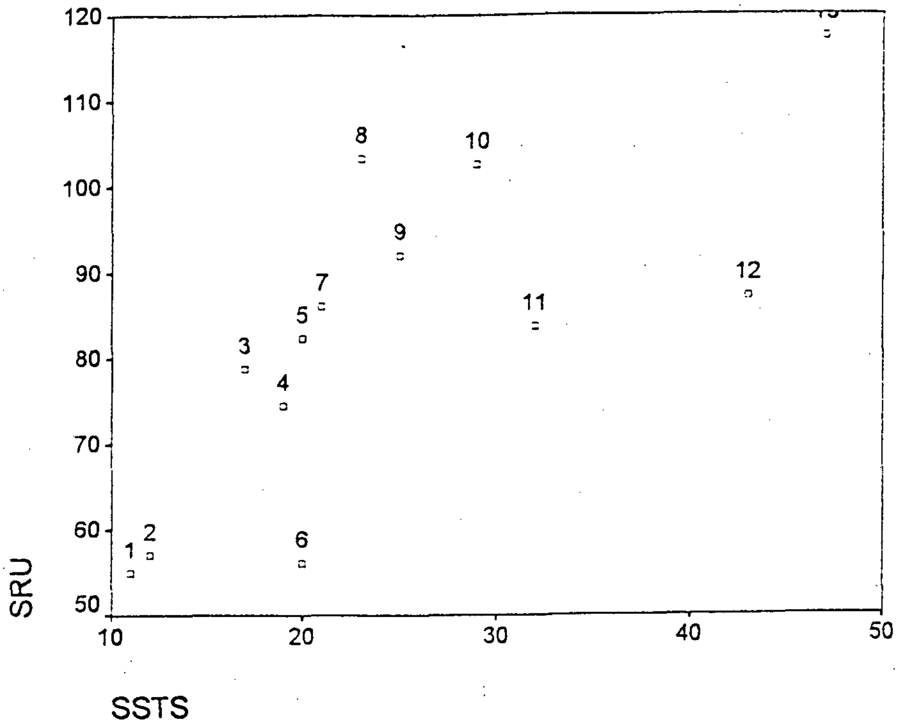
Correlations

			SPPM2	HPM2	COMPPM2	REVLD2
Spearman's rho	SSTS	Correlation Coefficient	-.490*	.564*	.746**	-.490*
		Sig. (1-tailed)	.045	.022	.002	.045
		N	13	13	13	13
	SRU	Correlation Coefficient	-.505*	.571*	.659**	-.451
		Sig. (1-tailed)	.039	.021	.007	.061
		N	13	13	13	13
	SRP	Correlation Coefficient	-.451	.538*	.659**	-.500*
		Sig. (1-tailed)	.061	.029	.007	.041
		N	13	13	13	13
	MLR	Correlation Coefficient	-.423	.599*	.692**	-.588*
		Sig. (1-tailed)	.075	.015	.004	.017
		N	13	13	13	13
	SPPM	Correlation Coefficient	.753**	-.797**	-.527*	.335
		Sig. (1-tailed)	.001	.001	.032	.131
		N	13	13	13	13
	HPM	Correlation Coefficient	-.445	.681**	-.088	.132
		Sig. (1-tailed)	.064	.005	.388	.334
		N	13	13	13	13
	ACC100	Correlation Coefficient	.159	-.187	-.698**	.775**
		Sig. (1-tailed)	.302	.271	.004	.001
		N	13	13	13	13
	COMPPM	Correlation Coefficient	-.302	.352	.687**	-.588*
		Sig. (1-tailed)	.158	.119	.005	.017
		N	13	13	13	13
	REVLD1	Correlation Coefficient	.220	-.187	-.852**	.720**
		Sig. (1-tailed)	.235	.271	.000	.003
		N	13	13	13	13
	SRU2	Correlation Coefficient	-.523*	.613*	.591*	-.492*
		Sig. (1-tailed)	.033	.013	.017	.044
		N	13	13	13	13
	SRP2	Correlation Coefficient	-.555*	.615*	.659**	-.533*
		Sig. (1-tailed)	.025	.013	.007	.030
		N	13	13	13	13
	MLR2	Correlation Coefficient	-.462	.217	.602*	-.399
		Sig. (1-tailed)	.056	.238	.015	.088
		N	13	13	13	13
	SPPM2	Correlation Coefficient	1.000	-.632*	-.445	.088
		Sig. (1-tailed)	.	.010	.064	.388
		N	13	13	13	13
	HPM2	Correlation Coefficient	-.632*	1.000	.269	-.159
		Sig. (1-tailed)	.010	.	.187	.302
		N	13	13	13	13
	COMPPM2	Correlation Coefficient	-.445	.269	1.000	-.808**
		Sig. (1-tailed)	.064	.187	.	.000
		N	13	13	13	13
	REVLD2	Correlation Coefficient	.088	-.159	-.808**	1.000
		Sig. (1-tailed)	.388	.302	.000	.
		N	13	13	13	13

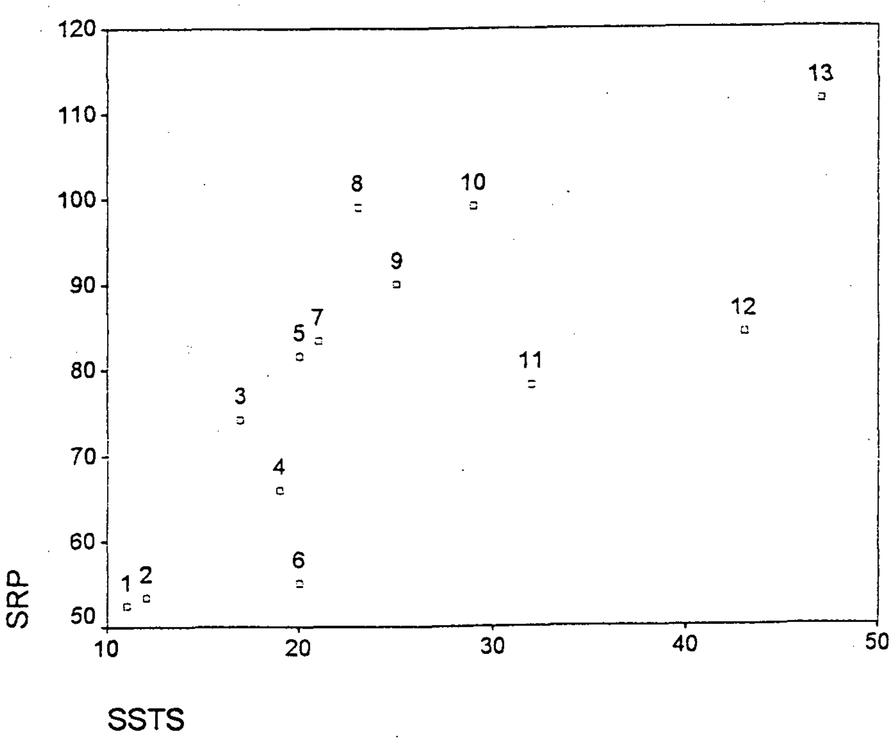
**. Correlation is significant at the .01 level (1-tailed).

*. Correlation is significant at the .05 level (1-tailed).

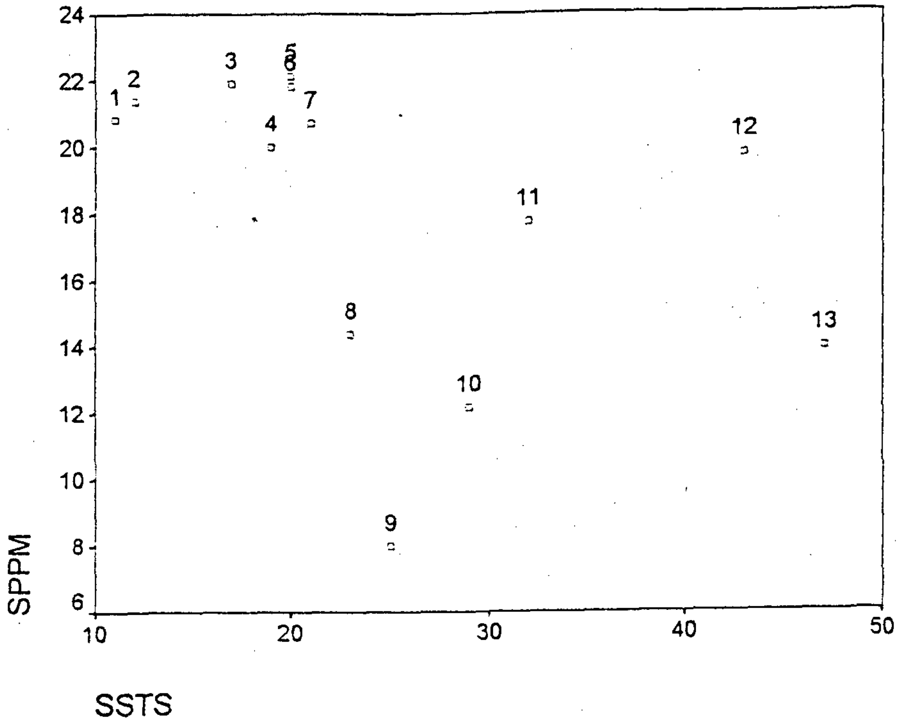
Graph



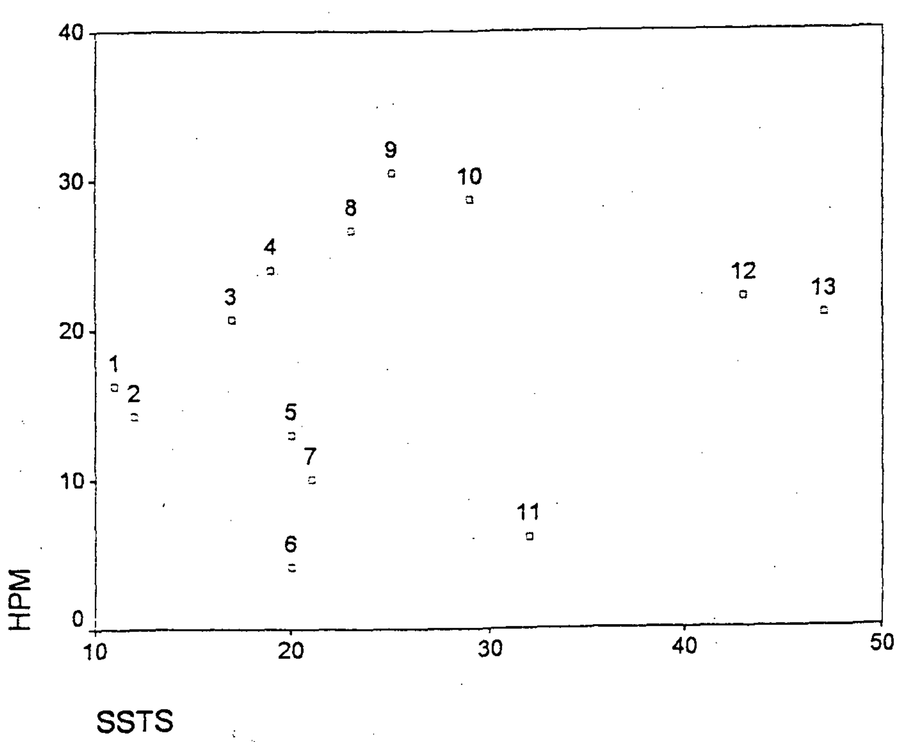
Graph



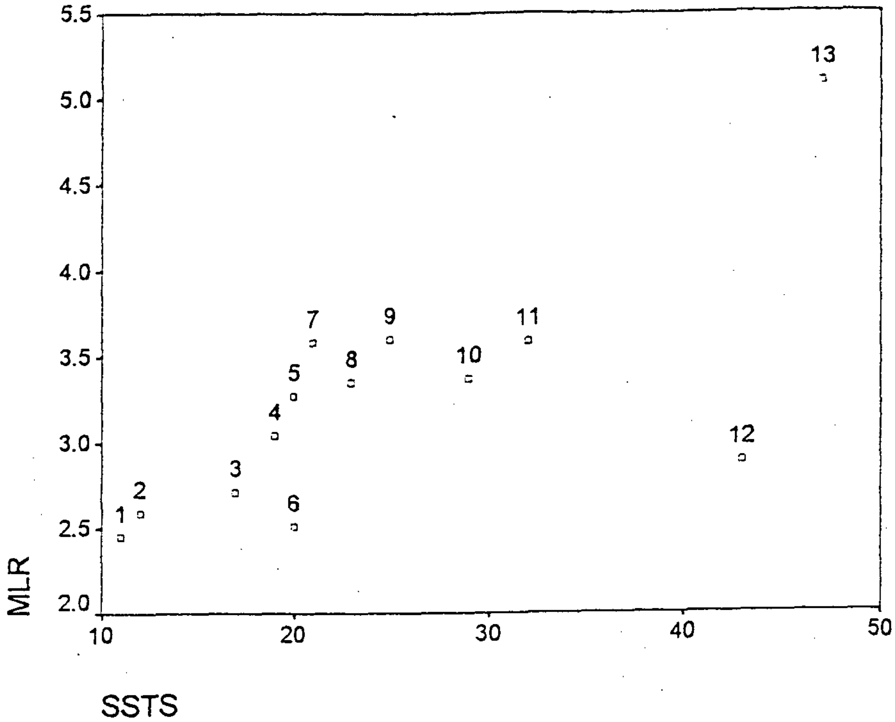
Graph


SPPM
24
22
20
18
16
14
12
10
8
6
10
20
30
40
50
1
2
3
4
5
6
7
8
9
10
11
12
13
SSTS

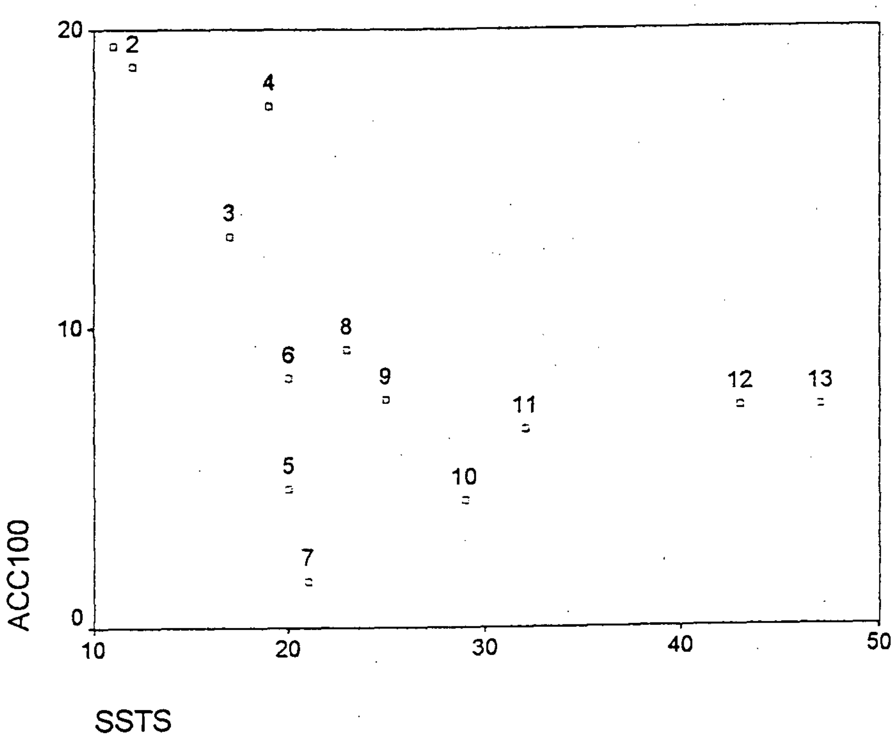
Graph


HPM
40
30
20
10
0
10
20
30
40
50
1
2
3
4
5
6
7
8
9
10
11
12
13
SSTS

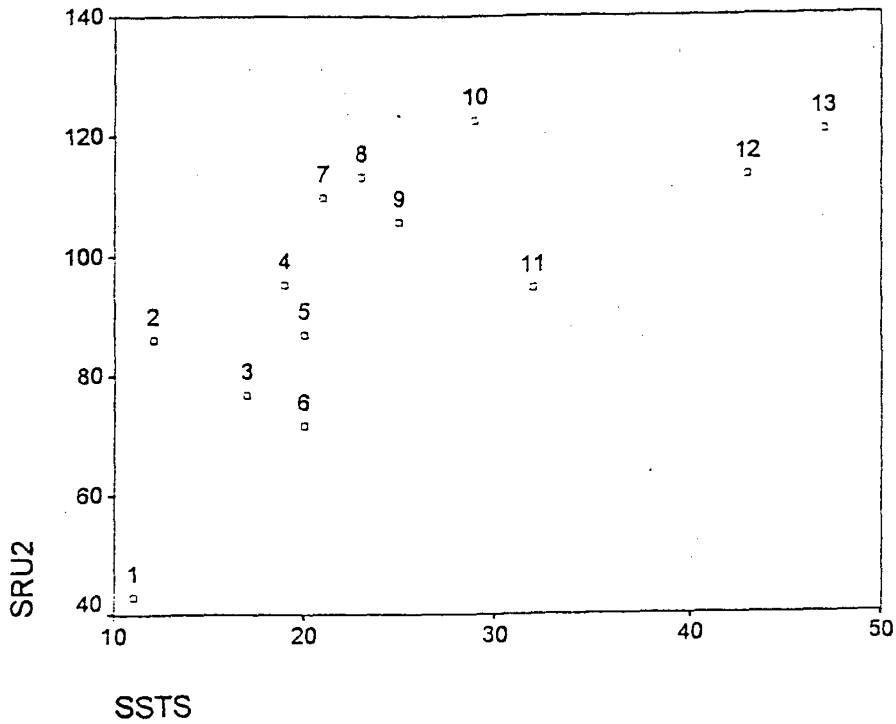
Graph


MLR
SSTS
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
10
20
30
40
50
1
2
3
4
5
6
7
8
9
10
11
12
13

Graph

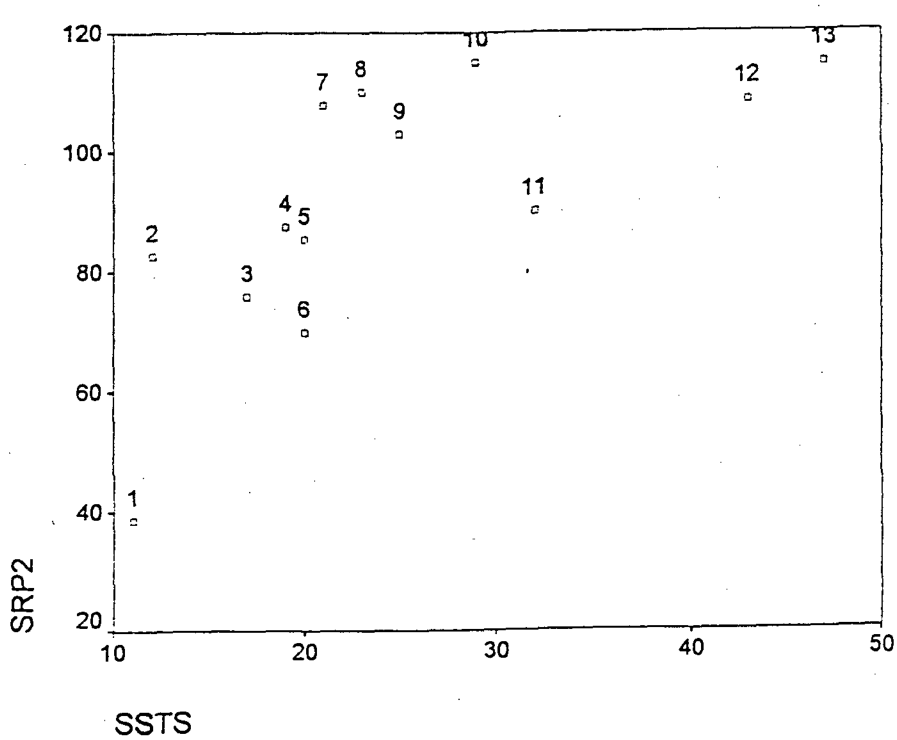

ACC100
SSTS
20
10
0
10
20
30
40
50
2
3
4
5
6
7
8
9
10
11
12
13

Graph



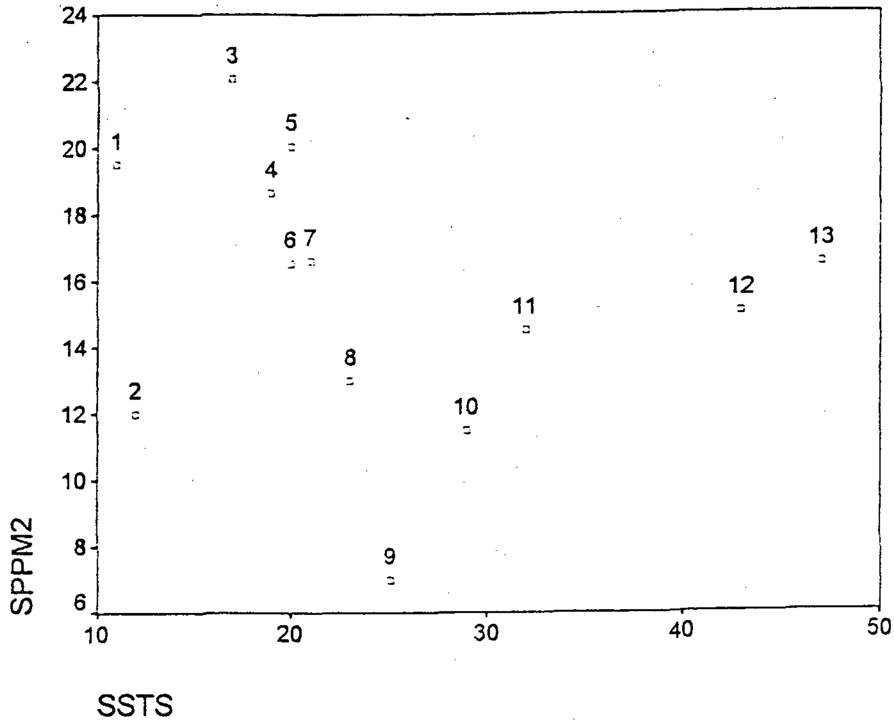
140
120
100
80
60
40
10
20
30
40
50
SRU2
SSTS
1
2
3
4
5
6
7
8
9
10
11
12
13

Graph

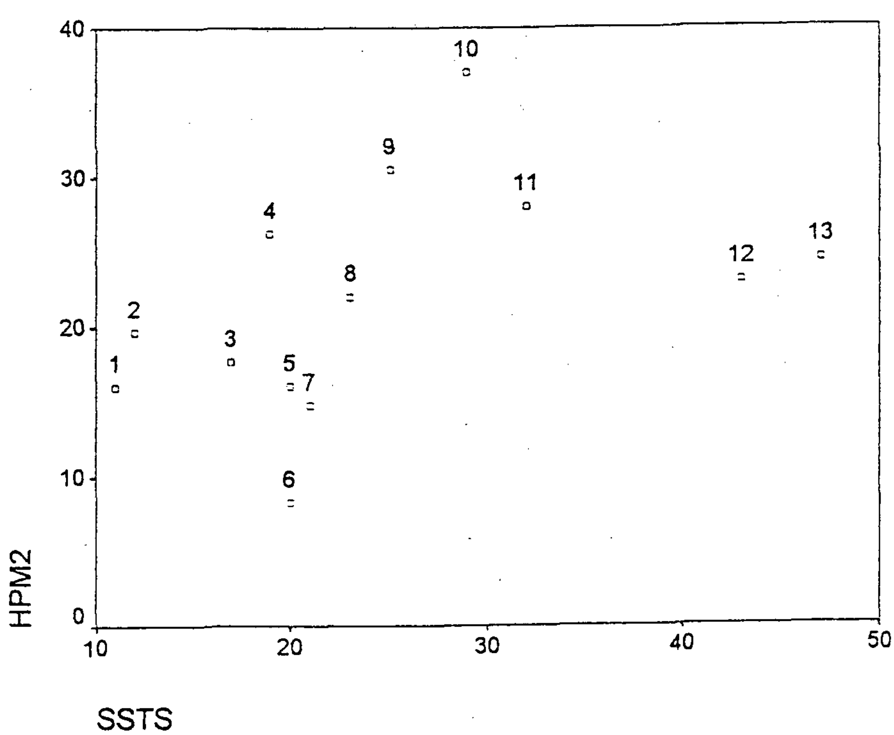


120
100
80
60
40
20
10
20
30
40
50
SRP2
SSTS
1
2
3
4
5
6
7
8
9
10
11
12
13

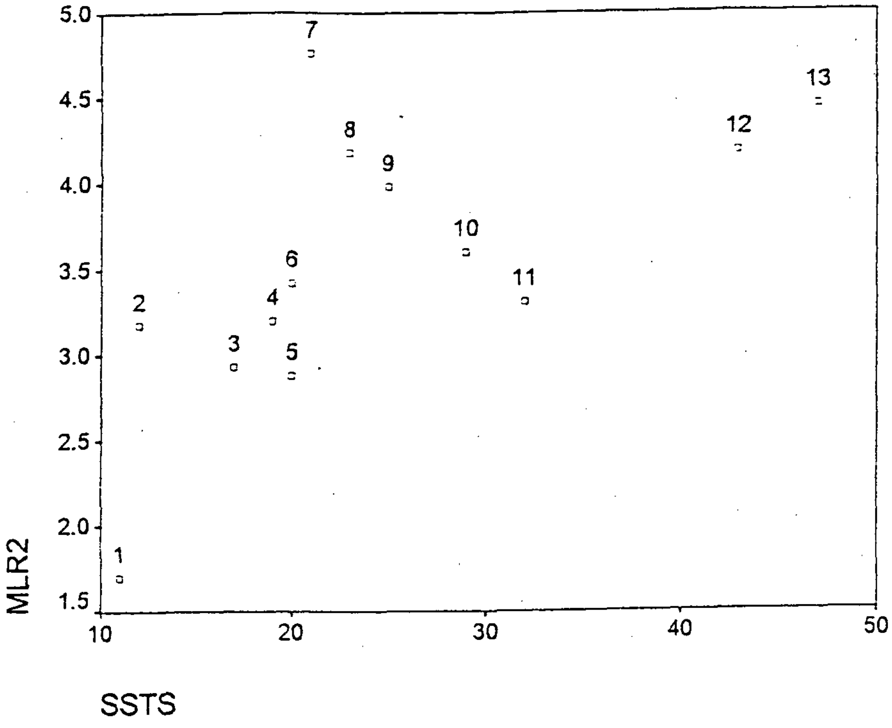
Graph


SPPM2
24
22
20
18
16
14
12
10
8
6
SSTS
10
20
30
40
50
1
2
3
4
5
6
7
8
9
10
11
12
13

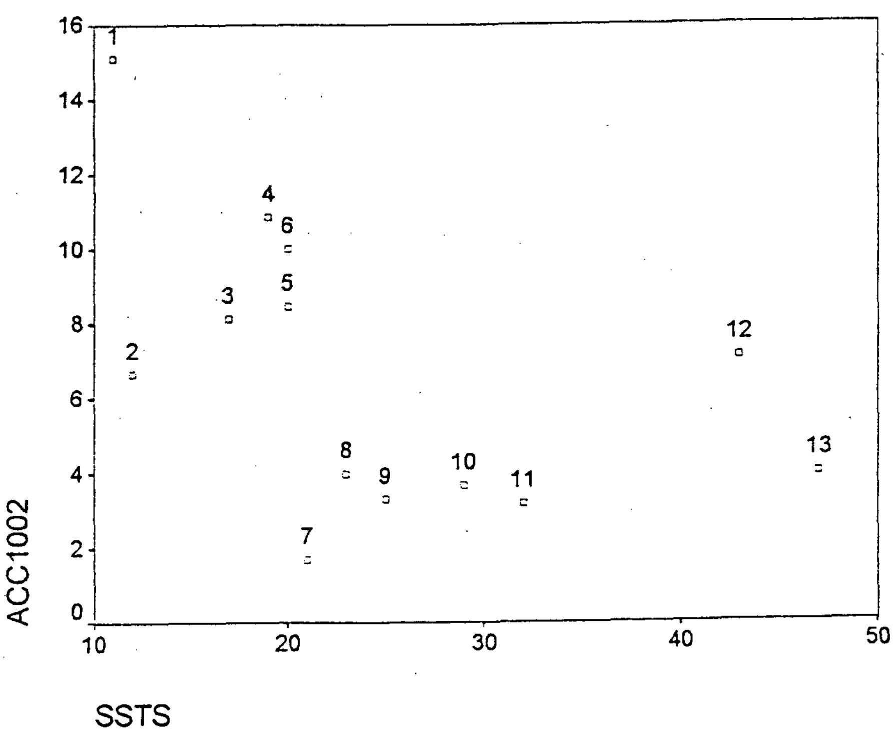
Graph


HPM2
40
30
20
10
0
SSTS
10
20
30
40
50
1
2
3
4
5
6
7
8
9
10
11
12
13

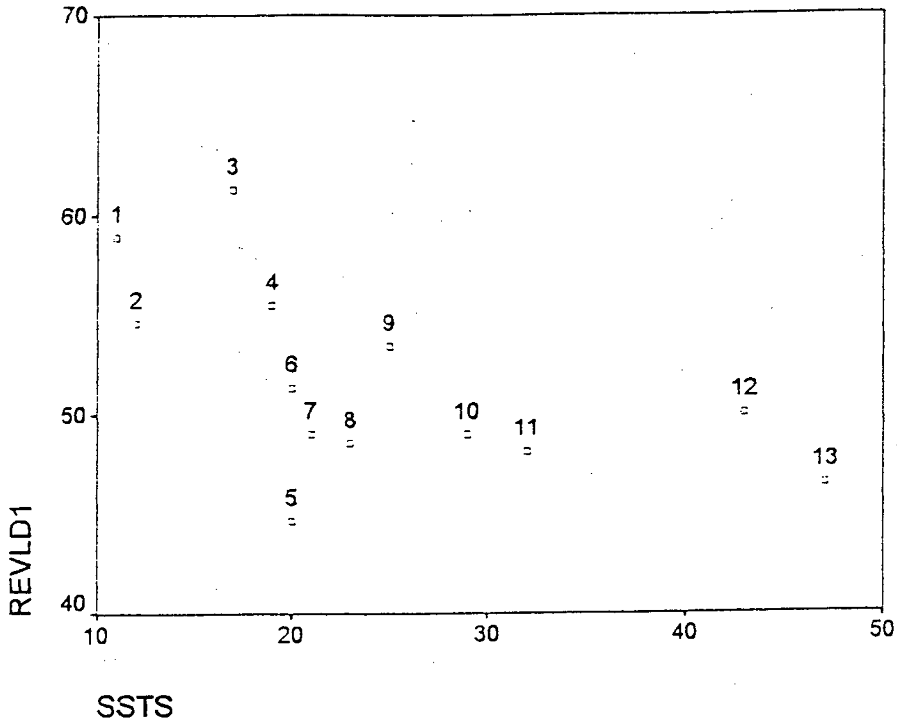
Graph


MLR2
SSTS
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
10
20
30
40
50
1
2
3
4
5
6
7
8
9
10
11
12
13

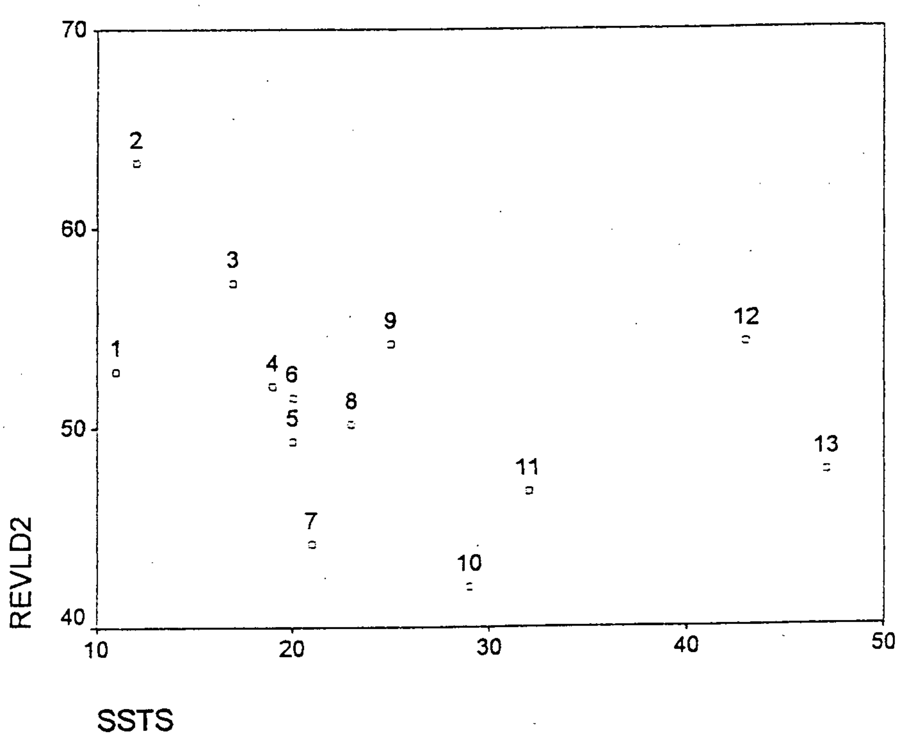
Graph


ACC1002
SSTS
16
14
12
10
8
6
4
2
0
10
20
30
40
50
1
2
3
4
5
6
7
8
9
10
11
12
13

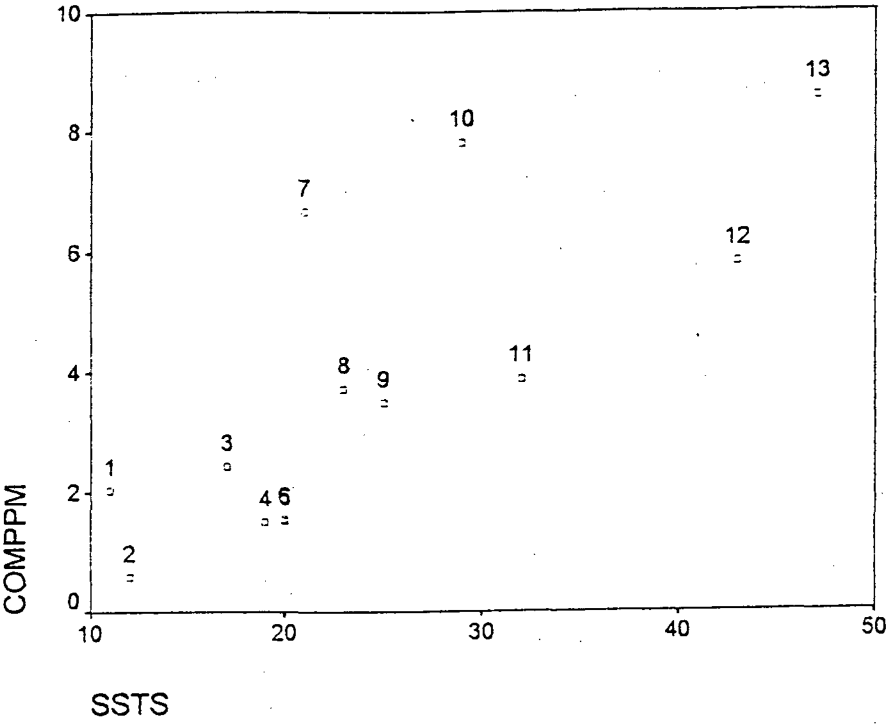
Graph


70
60
50
40
10
20
30
40
50
REVLD1
SSTS
1
2
3
4
5
6
7
8
9
10
11
12
13

Graph

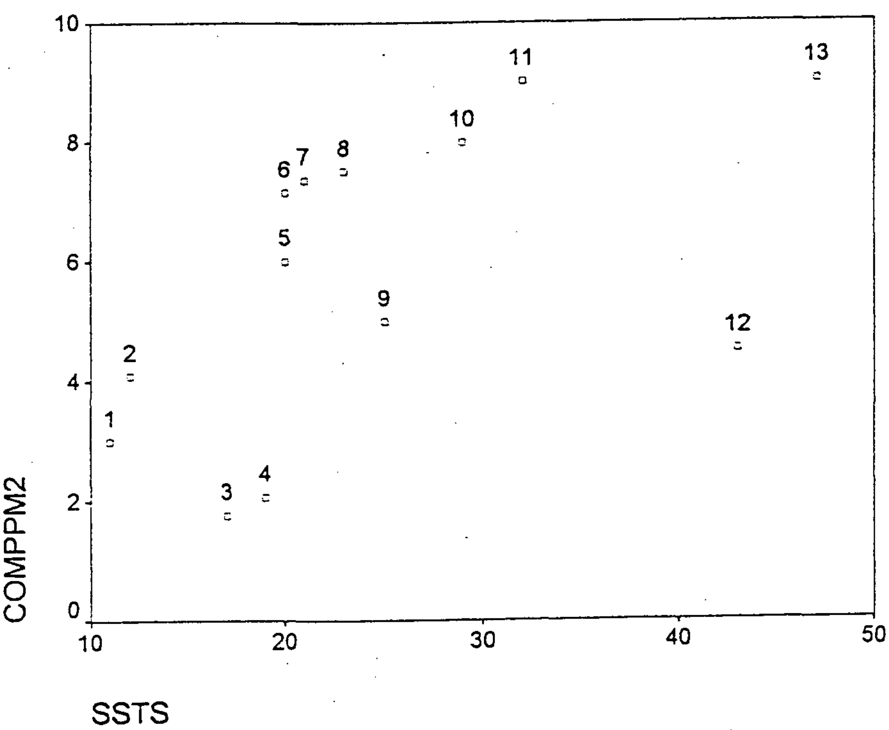

70
60
50
40
10
20
30
40
50
REVLD2
SSTS
1
2
3
4
5
6
7
8
9
10
11
12
13

Graph



10
8
6
4
2
0
COMPPM
10
20
30
40
50
SSTS
1
2
3
4 6
7
8
9
10
11
12
13

Graph



10
8
6
4
2
0
COMPPM2
10
20
30
40
50
SSTS
1
2
3
4
5
6
7
8
9
10
11
12
13

	ssts	owst	sru	srp	sppm	hpm	mlr
1	11.00	59.00	54.91	52.46	20.84	16.27	2.45
2	12.00	60.00	57.03	53.46	21.38	14.25	2.59
3	17.00	57.00	78.78	74.24	21.90	20.69	2.71
4	19.00	54.00	74.50	66.00	19.99	24.00	3.04
5	20.00	58.00	82.39	81.62	22.12	12.96	3.27
6	20.00	56.00	56.10	55.06	21.81	4.15	2.51
7	21.00	57.00	86.11	83.44	20.69	10.01	3.58
8	23.00	59.00	103.31	99.05	14.37	26.62	3.35
9	25.00	59.00	92.00	90.00	7.99	30.49	3.60
10	29.00	58.00	102.61	99.13	12.17	28.69	3.37
11	32.00	60.00	83.75	78.21	17.74	6.10	3.59
12	43.00	60.00	87.07	84.17	19.73	22.05	2.88
13	47.00	58.00	117.32	111.42	13.92	20.89	5.09

C:\My Documents\SPSS\variables.sav

	acc100	sru2	srp2	sppm2	hpm2	mlr2	acc1002
1	19.44	42.99	38.49	19.50	15.99	1.70	15.11
2	18.75	86.04	82.76	12.00	19.66	3.18	6.66
3	13.07	76.76	75.88	22.05	17.64	2.93	8.13
4	17.44	95.10	87.58	18.62	26.20	3.20	10.86
5	4.62	86.78	85.44	20.02	16.02	2.88	8.46
6	8.33	71.55	69.90	16.51	8.25	3.42	10.00
7	1.55	109.62	107.78	16.58	14.73	4.76	1.68
8	9.27	112.99	109.99	12.99	21.99	4.18	3.98
9	7.60	105.49	102.99	6.99	30.49	3.98	3.31
10	4.23	122.49	114.99	11.49	36.99	3.60	3.67
11	6.62	94.50	90.00	14.49	27.99	3.31	3.17
12	7.33	112.99	108.49	15.00	22.99	4.18	7.07
13	7.30	120.53	114.69	16.46	24.42	4.45	3.96

	ind	filter_$	revld1	revld2	comppm	comppm2
1	1	1	58.91	52.83	2.03	3.00
2	2	1	54.61	63.29	.59	4.09
3	3	1	61.25	57.26	2.43	1.76
4	4	1	55.49	52.07	1.50	2.06
5	5	1	44.66	49.32	1.52	6.00
6	6	1	51.35	51.51	1.55	7.15
7	7	1	49.01	44.17	6.67	7.36
8	8	1	48.57	50.18	3.72	7.50
9	9	1	53.42	54.19	3.49	4.99
10	10	1	48.99	42.03	7.82	7.99
11	11	1	48.14	46.78	3.88	9.00
12	12	0	50.00	54.20	5.80	4.50
13	13	0	46.46	47.70	8.57	9.02

Kuder-Richardson Formula 20 (Cronbach's Alpha)

$r_{tt} = n/n\text{-}1\ (s^2t - \Sigma s^2 i / s^2 t)$

r_{tt} = the KR 20 reliability estimate

n = the number of items in the test
s^2t = the variance of the test scores
$\Sigma s^2 i$ = the sum of the variances of all items

From Henning, G. (1987). *A guide to language testing*. New York: Newbury House.