# MEASURES OF WORKING MEMORY CAPACITY AND L2 ORAL FLUENCY

**Mailce B. M. Fortkamp**

## Introduction

A considerable amount of research is now available which addresses the nature of nonnative oral fluency. The research carried out over the past two decades has shown with some consistency that (1)nonnative speech production tends to reproduce first language (L1) speech organization, with a greater number of pauses, greater pause time, increased hesitation phenomena, and decreased speech rate (Deschamps, 1980; Raupach, 1980 cited in Olynyk, Sankoff, and d'Aglejan, 1983, p. 232), (2)nonnative oral fluency is sensitive to context (Riggenbach, 1991) and task structure (Ejzenberg, 1992), (3)nonnative fluent speakers share a great number of fluency features whereas nonnative nonfluent speakers will be nonfluent in idiosyncratic ways. (Riggenbach, 1989; Olynyk, D'Anglejan, & Sakoff, 1990; Ejzenberg, 1992; Freed, 1995), and (4)frequency of hesitation phenomena is related to the production of new utterances and to the level of cognitive difficulty of the task whereas an increase in speech rate is observed when the speaker is being repetitive (Goldman-Eisler, 1968).

| Ilha do Desterro | Florianópolis | nº 35 | p.201-238 | jul./dez. 1998 |
|---|---|---|---|---|

Most of the research on L2 oral fluency has investigated the phenomenon from sociolinguistic or psycholinguistic perspectives, focusing on the profiles of fluent and nonfluent nonnative speakers or on the temporal and linguistic aspects of the speech produced. The present study examined L2 oral fluency under the assumptions of information processing theory, the theory informing current cognitive science research. More specifically, the present study investigated whether there is a correlation between measures of working memory capacity and measures of L2 oral fluency.

*Working Memory*

It is now widely accepted that memory is composed of two major systems - long-term and short-term memory. With respect to short-term memory, this notion is to date generally referred to as working memory[1] (Baddeley and Hitch 1974; Hitch and Baddeley, 1976) and is conceptualized as the system responsible for the simultaneous storage and processing of information (Baddeley, 1992a, 1992b; Daneman, 1991) during the performance of cognitive tasks (Baddeley, 1990).

As Baddeley (1992) suggests, research on working memory has been developed along two different but complementary approaches. The first one, the dual-task neuropsychological approach, focuses on the analysis of the structure of the system as proposed by Baddeley and Hitch (1974) and Hitch and Baddeley (1976). Its methodology consists of the application of dual tasks, e.g., remembering a list of digits while reasoning (Baddeley, 1990, p. 68), and of the study of neuropsychological evidence to explain, mainly, the slave subsystems.

The second approach, the psychometric correlational, is concerned with the correlations existing between working memory capacity and the performance of complex cognitive tasks. Within this approach the two functions of working memory - storage and processing - compete for its capacity during the performance of complex cognitive tasks (Daneman and Carpenter, 1980, 1983). The methodology generally consists of devising laboratory tasks in which both storage and processing of information are necessary, and subsequently using the individual's results of performance on these tasks to predict his/her skills in demanding cognitive tasks, as for instance, reading comprehension. The present study was carried out within the psychometric correlational approach and focuses on L2 speech production.

*The Psychometric Correlational Approach to Working Memory Capacity*

Under the assumption that working memory has the dual function of storing and processing information and that traditional digit or word span tasks do not reflect the processing function efficiently, Daneman and Carpenter (1980) devised a complex measure of working memory span which they termed the Reading Span Test. In their view, there is a trade-off between storage and processing in working memory, which is likely to be a source of individual differences in reading comprehension. They propose (Daneman & Carpenter, 1980, 1983), then, that the processing and storage functions of working memory compete for its limited capacity.

The Reading Span Test, as it was first devised by Daneman and Carpenter (1980), requires subjects to use both functions of working memory: 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 subject's reading span is the maximum number of sentence-final words recalled (in the order they were presented) and is taken as an index of his/her working memory capacity.

The Reading Span Test has been the basis for most of the research on individual differences in working memory and reading comprehension, and has been extensively used as a predictor of performance on various other aspects of reading: (1) the ability to detect inconsistencies in sentences with homonyms (Daneman & Carpenter, 1983); (2) the ability to make inferences of ideas not explicitly mentioned in the text (Masson & Miller, 1983); (3) the ability to make use of contextual cues to infer the meaning of new words in the text (Daneman & Green, 1986); (4) the resolution of lexical ambiguity in reading (Miyake, Just & Carpenter, 1994); and (5) the perception of text structure (Tomitch, 1995). Various researchers (e.g. Daneman & Carpenter, 1980; Masson & Miller, 1983; Turner & Engle, 1989) have also found strong correlations between the Reading Span Test and standardized measures of reading ability such as the Verbal Scholastic Aptitude Test and the Nelson-Denny reading test.[2]

Claiming that individuals differ considerably in the fluency with which they speak, Daneman (1991)—on which the present study draws heavily - verified whether differences in working memory capacity could account for this variation in L1 speech production. Considering that speaking is a complex cognitive task which requires coordination of storage and processing of information in the various stages of the speech production process, Daneman hypothesized that individuals with larger working memory capacities would perform better on tasks measuring fluency.

Subjects' working memory capacity was assessed by means of the Speaking Span Test (Daneman & Green, 1986), aimed at measuring working memory capacity during speech production. The test consisted of presenting subjects with increasingly longer sets of unrelated words, which they had to read silently. At the end of a set, subjects were required to produce aloud a sentence for each individual word presented, in their original order and form of presentation. A subject's speaking span was operationalized in terms of his/her total capacity -the total number of words for which he/she was able to produce a grammatical sentence. 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 the speaking span lenient, counting also sentences containing the word in a different form.

Subjects' oral fluency was assessed by means of the Speech Generation Task, the Oral Reading Task and the Oral Slip Task. The

Speech Generation Task aimed at eliciting speech at the discourse level and consisted of the description of a picture for 1 minute and 30 seconds. Measures of fluency in this task were number of words completely articulated - the main measure - and richness and originality of context. The Oral Reading Task and the Oral Slip Task both aimed at measuring fluency in terms of speed and accuracy in the articulation of words. In the Oral Reading Task subjects were required to read a passage aloud and the main measure of fluency was reading time. In the Oral Slip Task, which aimed at eliciting spoonerisms, subjects were required to say cued pairs of words shown on a computer screen. The measures of fluency used were number and types of errors made. In addition, Daneman applied a Reading Span Test, which she hypothesized would correlate with the reading related task.

The study was carried out with 29 English L1 university students and results show that the Speaking Span Test correlated significantly with the Speech Generation Task, the Oral Reading Task, and the Oral Slip Task; i.e., subjects with larger working memory capacity performed better on the picture description task, took less time reading the passage aloud, and were less prone to producing spoonerisms. Also, as predicted, the Reading Span Test correlated significantly only with the reading-related task - the Oral Reading Task. The Speaking Span Test yielded two types of scores - one strict and one lenient - and, as hypothesized, the two speaking span scores differed in the aspects of fluency they predicted. Speaking span strict correlated better with the Oral Reading Task and the Oral Slip Task, while speaking span lenient correlated more significantly with the Speech Generation Task.

These results are explained by the claim that the Speaking Span Test is a complex measure of working memory span for language production which taxes both the storage and processing functions of this limited system during the production of speech. 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 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. It is important to note that the Speaking Span and Reading Span tests are recall tests which were devised to measure working memory span under language production or comprehension processing demands. The 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. Thus, as claimed by Daneman and colleagues, 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 reading comprehension processing is more efficient and thus they use less of their capacity. Accordingly (Daneman, 1991), more fluent speakers have a larger working memory capacity, as measured by the Speaking Span Test, because they are more efficient in executing the processes required during speech production, leaving greater resources available for the storage of the intermediate products of this processing.

In an attempt to discover whether this correlation would be also true for L2 oral fluency, Mota (1995) adapted Daneman's (1991) methodology and applied a set of seven experiments to 16 advanced speakers of English as an L2. Working memory was assessed by means of the Speaking Span Test and the Reading Span Test, both of them in Portuguese and in English. L2 fluency was assessed by means of the Speech Generation Task, the Oral Reading Task, and the Oral Slip Task. Working memory capacity, as measured by the Speaking Span Test in English, correlated significantly only with the Speech Generation Task, which was aimed at assessing fluency at the discourse level. Working memory capacity, as measured by the Reading Span Test, both in

Portuguese and English, correlated significantly only with the reading-related task, the Oral Reading Task, aimed at assessing fluency at the articulatory level. The Oral Slip Task suffered a methodological flaw in terms of the way subjects' results could be entered for statistical computations and thus no statistical analysis could be carried out.[3] The results obtained in Mota (1995) corroborate the task-specific view of working memory capacity (e.g. Cantor and Engle, 1993).

*Three Views of Working Memory Capacity*

There are currently three main theories that account for individual differences in working memory capacity: the task-specific view, the processing efficiency view and the activation view (Cantor & Engle, 1993).

The task-specific view poses that the greater an individual's efficiency in processing information, the greater the capacity left available for storage of the products of this processing and of material retrieved from long-term memory (Cantor & Engle, 1993). The singularity of this view is that this more efficient processing is highly task-specific (Daneman & Green, 1986); i.e., an individual's working memory capacity will vary according to his/her efficiency in the processes specific to the task with which working memory capacity is being correlated. Thus, for instance, good readers will have a functionally larger working memory capacity in reading-related tasks, but not necessarily in language production tasks. Within this view, the processing component of the span test must require the same processes present in the cognitive task whose performance is being predicted.

Further elaborations of the task-specific view have led to the processing efficiency view, which claims that there are general skills which are employed in any task demanding the manipulation of language. For instance, Daneman and Tardiff (1987) argue that individual differences in working memory capacity can be measured through processing efficiency alone, without including a simultaneous storage component in the task.

Daneman and Tardiff (1987) 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. However, to show that the crucial variable in individual differences in working memory is processing efficiency, 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 conclude that it is individual differences in processing that explain differences in verbal abilities. Thus, the emphasis is on the efficient processing skills individuals have when performing language-related tasks. The difference between the task-specific and the processing efficiency views is that in the latter the processing component of the span task need not specifically require the same processes of the task being predicted.

The activation view defines working memory as information in long-term memory that is temporarily activated to a level that makes it available for cognitive activity (Cantor & Engle, 1993; Engle, Cantor & Carullo, 1992). The capacity of this system is the total amount of activation an individual has available to retrieve information from long-term memory in order to carry out a cognitive task. Individuals with higher or lower spans, as measured by the span test, differ in their limits of activation. This limited capacity, in the activation view, is independent of the nature of the task, being, thus, a single unitary resource.

### *L2 Oral Fluency*

We all have an intuitive concept of what it is to be fluent and upon hearing someone talk, we immediately judge him as more (+) or less (-) fluent, although we might not be aware of what makes us consider the speaker as such. Fillmore (1979), one of the first researchers to point out individual differences in fluency, suggests we may judge speakers to be fluent in their L1 in four main ways. In his view, fluency is related to

(1) the ability to fill time with fast talk, (2) the ability to produce semantically dense speech, (3) the ability to perform in several pragmatic aspects of language, and (4) the ability to speak with creativity and imagination, building metaphors, punning and making jokes with the meanings and sounds of words, on line.

Highlighting the multidimensional nature of the phenomenon in each of these aspects of fluency, Fillmore proposes that different types of knowledge and skills are involved in the production of fluent speech, mentioning that speakers vary in their vocabulary size, in their knowledge of linguistic forms and formulaic expressions, in their ability to create new expressions as well as to access and use syntactic constructions of their L1 in the various conversational settings and discourse patterns, varying also in their knowledge of appropriateness of language.

While these four kinds of fluency might well be considered true also for second languages, L2 fluency has been judged and defined in a rather different fashion. As Riggenbach (1991) points out, the notion of fluency has played a much more central role in L2 research than it has in L1, since fluency has been considered an important factor in assessing L2 proficiency.

Most studies dealing with L2 fluency have described the phenomenon at isolated levels of occurrence, from the utterance to the discourse level (Ejzenberg, 1995). As a result, fluency has been defined in a number of different ways. Traditional definitions of fluent speech are "speech that lacks unnatural pauses" or "speech that exhibits smoothness, continuity, and naturalness" (Riggenbach, 1991: 423-24).

In an attempt to organize the ways in which L2 fluency has been understood, Lennon (1990) concludes that the term fluency is generally used in two senses. In its broader sense, it is equated to oral proficiency: a fluent speaker would be the one whose oral production is native-like in all aspects—vocabulary range, grammatical correctness, pronunciation, idiomaticness, appropriateness, and relevance[4] . In its narrower sense, Lennon argues, fluency in an L2 is one component of oral proficiency and is basically related to speech rapidity, to the flow of speech without this being impeded by hesitations. In this narrower sense, fluency is opposed to other components of oral proficiency such as lexical range, grammatical correctness, pronunciation, idiomaticness, appropriateness, and relevance.

This narrower sense is related to the definition Lennon (ibid.) gives for fluency as the perception we have, when hearing someone talk, that the speaker's psycholinguistic processes involved in speech planning and production are working easily and efficiently (p. 391). In line with this view, Schmidt (1992) defines fluency as an *automatic procedural skill (*cf. Carlson, Sullivan, & Schneider, 1989). For him, "fluent speech is automatic, not requiring much attention or effort" (p. 358), in contrast to nonfluent speech, which is effortful and which demands focused attention on a number of processes involved in the various stages of speech production.

Early studies of L2 oral fluency emphasized mainly the temporal variables of speech production. Möhle (1984) compared speech samples of advanced L2 learners of German and French performing a description task and a free discourse task. She was able to identify a number of measures of fluency, among them, speech rate, length and position of unfilled pauses, number and distribution of filled pauses, and length of speech runs between pauses.

Rehbein (1987) analyzed the pauses produced by learners of German as an L2 and developed a set of hypotheses concerning L2 fluency. He posits that fluency is dependent on the activity of planning, which requires the L2 speaker to create a global scheme for his/her utterance. Planning and uttering take place in part simultaneously causing the speaker to pause. Rehbein also points out that fluent speech depends on the type of task the speaker is required to perform, the type of event he/she is involved in, the type of discourse being carried out, and the expectations of the hearer.

Lennon (1990) attempted to quantify the components of fluency by analyzing speech samples of four adult German university students of English as a second language on two occasions—before and after subjects' study visits to England. Based on subjects' narration of a sequence of pictures, Lennon devised a wide range of measures of fluency encompassing both temporal variables and disfluency markers, many of them in the tradition of Goldman-Eisler (1968). By comparing each subjects' first and last narratives, Lennon found that there had been improvements in their fluency mainly in terms of speech rate and number of filled pauses. He reports that subjects' speech was faster, with fewer repetitions and filled pauses per T-units, less time occupied by unfilled pauses, longer fluent runs between pauses and T-units, and a reduction of pause time at T-unit boundaries.

Riggenbach (1991) was one of the first studies to use conversational data and to include interactive features of speech production in the evaluation of L2 oral fluency. Riggenbach (1991) analyzed the speech of 6 Chinese students of English as an L2, 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 the ones of those considered to be highly nonfluent. Riggenbach asked her subjects to record a dialogue and the quantitative analysis of the speech samples included specific fluency-related items such as hesitation phenomena, repair phenomena, rate and quantity of speech, interactive phenomena, and turn change types. Each of these categories contained a set of sub-items, summing up 19 variables. The results obtained showed few significant differences in features between fluent and nonfluent subjects. However, Riggenbach was able to verify that fluent and nonfluent subjects differed in terms of speech rate and number of filled pauses, supporting Lennon's (1990) findings. Subjects judged as very fluent speakers also showed more ability to make appropriate topic changes and to anticipate end of turns.

Ejzenberg (1995) investigated the effect of task structure - dialogue vs. Monologue - on the display of L2 oral fluency of 50 subjects. In addition, she verified whether there were quantitative and qualitative differences in the speech produced by the very fluent and the very disfluent subjects. By manipulating the structure of the tasks used in the study, Ejzenberg was able to show that "interactivity" is an important variable affecting speakers' display of fluency (1995, p. 17). Thus, her subjects appeared to be more fluent in dialogues than in monologues, with subjects' fluency varying according to the degree of interactivity present in the context of speech production. The qualitative analysis of four features of speech of three high- and three low-fluency subjects across tasks showed that high-fluency speakers tend to speak more and faster than their low-fluency counterparts. High-fluency speakers also produce longer talk units and longer fluent units (Ejzenberg, 1995, pp. 34, 36); Postma, Kolk, & Pole, 1990; Pawley & Syder, 1983), displaying, in addition, a number of discourse strategies during speech production in order to maintain an "air of fluency" (Ejzenberg, 1995, p. 38).

Freed (1995) investigated whether native speaker judges' global perceptions of fluency would distinguish between two groups of L2 learners - one with experience in studying in the country of the target language and the other with formal classroom instruction only. Freed also attempted to identify features of fluency that distinguished the two groups. The speech samples of 30 subjects were first subjectively analyzed by a group of 6 native speaker judges on a 7-point scale. Subsequently, linguistic analyses of 8 subjects' speech samples were performed in order to identify attributes of fluency that would help determine those subjects who had been abroad from those who had not. For this linguistic analysis, Freed chose mainly temporal variables and a number of disfluency markers. The analyses performed by native speaker judges' revealed a small difference in the perceived global fluency between the two groups, with a modest increase for the less advanced students (Freed, 1995, p. 134). The linguistic analyses, however, showed that subjects who had lived abroad tended to speak more and faster, with fewer silent and non-lexical pauses, longer speech runs, and a greater number of reformulations and false starts.

The studies reviewed above all focus on L2 fluency as a product of the speech process and attempt to identify the features of L2 fluent speech production. The present study focuses on fluency from a cognitive perspective, thus being primarily concerned with the cognitive processes involved in the production of L2 speech - more precisely, with the ability the speaker has to coordinate the various mental processes involved in this production. For the purposes of the present study, and following Lennon (1990) and Ejzenberg (1995), fluency is here restricted to the oral mode and is considered a component of language proficiency, being operationalized as the impression the listener has that the speech being produced is smooth, continuous, with few hesitation phenomena, coherent and adequate to the context. Working memory is assumed to be involved in the coordination of fluent L2 speech production.

## Method

### *Research Questions*

The main objective of the present study was to investigate whether there is a correlation between individuals' working memory capacity and their oral fluency in English as an L2. A set of 4 experiments was applied in order to assess subjects' working memory capacity and L2 fluency: the operation-word span test (Cantor & Engle, 1993; Engle et al., 1992) in English, the speaking span test ( Daneman, 1991; Daneman & Green, 1986) in English (Mota, 1995), a picture description task and a narrative task. The present study pursued the following question: Is there a correlation between working memory capacity and temporal variables of fluency in L2?

### *Subjects*

Subjects for this study were 11 undergraduate students of English at the Federal University of Santa Catarina. There were 6 students at the intermediate level and 5 students at the advanced level; 7 were female and 4 were male. All subjects' first language was Portuguese.

*Materials and Tasks*

Measures of Working Memory Capacity: Subjects' working memory capacity was assessed by means of the operation-word span test and the speaking span test. The operation-word span was devised by Turner and Engle (1989), and its background processing task consists of the resolution of a mathematical operation string. The speaking span test version used in the present study was that applied by Daneman (1991) and Mota (1995), and its background task consists of the production of sentences. The tests are fully described as follows:

(1) Operation-Word Span Test: the operation-word span test was constructed with 60 operation strings and 60 English words presented on the middle line of a computer video screen. Each operation string was accompanied by a word. Subjects were required to pace themselves through the operation string, evaluate whether the result presented for the operation was true or false—by pressing the letters V or F on the keyboard—and retain the word accompanying the operation for subsequent recall. Each pair of operation-word was arranged in three sets each of two, three, four, five, and six operation-word pairs (see Appendix A, for the organization of pairs). After solving the operation, checking if the result presented was true or false, and retaining the word for subsequent recall, subjects would press "enter" and the next pair would appear, as shown in the example below:

| | | |
|---|---|---|
| (6x2) - 2 = 10 | | HOUSE |
| V | F | <enter> |
| (9/9) + 1 = 10 | | FINE |
| V | F | <enter> |

This procedure was followed until a blank screen accompanied by a beep signaled that a set had ended. Subjects were then required to recall the words accompanying each operation in the order they had appeared and in the exact form they were presented. The combination of an operation with a word was randomized for each subject in order to minimize test difficulty effects. The number of operation word sequences presented before recall was also randomized to prevent subjects from being able to predict the number of words they would have to recall. Practice trials were given to each subject and the actual span task would begin only when the subject felt comfortable enough with the test. The operation strings were taken from Cantor and Engle (1993) and the words from Cantor and Engle (1993) and La Pointe and Engle (1990). Following Engle et al. (1992), a subject's operation-word span was his/her total performance on the test, i.e., the total number of words recalled - in this case, the maximum being 60.

(2) Speaking Span Test in English (SSTE): The SSTE was constructed with 40 unrelated one-syllable words, arranged in two sets each of two, three, four, five, and six words (see Appendix B, for the organization of words in sets). Each word was presented on the middle line of a computer video screen for 1 second and was accompanied by a beep. Subjects were instructed to read the words 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, until a blank screen signaled that a set had ended. Subjects 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. Thus, for instance, after being presented with the set:

duck pen gas

a subject generated the sentences:

"The duck is in the pond."
"The pen is mine."
"I need some gas."

Subjects were told that there were no restrictions as to the length of the sentences, but they were required to make them grammatical as regards syntax and semantics. 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. The two-word sets were presented first, followed by the three-word sets, the four-word sets, and so on. Following Daneman (1991) and Daneman and Green (1986), the measure applied to a subject's speaking span in English was his/her total performance on the test, i.e., the total number of words for which a grammatical sentence was produced—in this case, the maximum being 40.

Measures of L2 fluency: Subjects' L2 oral fluency was assessed by means of a picture description task and a narrative task. Six temporal variables of fluency were used to measure each subject's fluency.

(1) Picture Description Task: Subjects were presented with a set of three colorful pictures taken from a magazine, all related to technology and technological means of communication. The first picture portrayed an Asian man in a desert holding a palm-top at the beginning of a line of different models of computers, the last being the bigger and older models. The second picture portrayed a number of people forming a circle in a distant region of Romania. The people, dressed in different ways, each held a telephone. The last picture portrayed a late 1950s family watching TV. On the screen as well as outside the window, Bill Gates' face appears. Subjects were required to describe the pictures and express their opinion on (1) the relationship, if any, among the three pictures, (2) how the three pictures related to technology, and (3) the impact/importance of technology on our society. Subjects were explicitly instructed to give as much information as possible and talk as much as they wanted. They were given time to analyze the pictures to decide the order in which they would describe them and to solve vocabulary problems before the beginning of the task. The task would begin when subjects' signaled they were ready.

(2) Narrative Task: subjects were required to tell the most important aspects of their school life, from when they first went to school to their present days at college. Again, subjects were explicitly instructed to give as much information as possible and to talk as much as they wanted.

(3) Variables of Fluency: The selection of the measures of fluency for the present study was based on the findings of previous studies on L2 oral fluency. The selected measures consisted of those temporal variables and disfluency markers that have been considered most salient in the display of L2 fluency as shown by Ejzenberg (1992), Freed (1995), Lennon (1990), Riggenbach (1989), and Temple (1992):

(a) rate of speech: number of semantic units (words or comprehensible parts thereof) per minute.
(b) amount of speech: total number, calculated as raw frequencies, of non-repeated words including filled lexical pauses and partial words. Two versions of amount of speech were considered: (i) total number of non-repeated words including filled lexical pauses and partial words in English only, and (ii) including non-repeated words, filled lexical pauses and partial words in Portuguese.
(c) micropauses: a silence of 2 seconds or less.
(d) hesitation: a silence of 2 to 4 seconds.
(e) unfilled pause: a silence of 4 seconds or greater
(f) filled pause: voiced fillers which do not add information to the speech being produced. These fillers were subdivided into 2 types:

(i) nonlexical: fillers which do not carry semantic information (e.g., "uhm" and "uh") and are not recognized as words.
(ii) lexical: fillers which are recognized as words but do not add information relevant to what is being said (e.g. "you know", "I mean"). Lexical fillers produced in Portuguese, such as "é ", were also considered.

All subjects' responses in all tasks were tape-recorded for later transcription. Following Ejzenberg (1992) and Riggenbach (1989), the transcription system used was adapted from that proposed by Jefferson (1979, in Maxwell and Heritage, 1984). An explanation of the symbols used in the transcriptions follows.

Unfilled pauses were timed with a stop-watch in seconds. Pauses of 2 seconds or less are considered micro-pauses and are indicated by a period inside a parentheses (.); pauses of greater length are indicated by the time period within parentheses (.8). Colons ( : ) are used to indicate sound stretches. A period indicates a fall in tone, as (but not necessarily) at the end of a sentence. A comma indicates continuing intonation. A question mark indicates rising intonation, such as (but not necessarily) in yes/no questions. An exclamation mark indicates an animated tone. A slash indicates a halting, abrupt cut off. Words or parts of words that are underlined indicate emphasis and louder volume than other words in the surrounding environment. Parentheses were used to indicate transcriber's doubts as to what the speaker said. A string of the letter X within parentheses indicates that the speech was incomprehensible.

*Procedures*

The data for this study were collected individually with each subject in a small room at CCE/UFSC during the month of November 1997, in one session which lasted about one hour. There were two computers, a printer, a tape recorder and a few tables and chairs in the room. First, subjects were given the memory span tests—the operation word span test followed by the speaking span test. Then, after a short break of five minutes, subjects were given the fluency tasks—the description task followed by the narrative task. Subjects were explicitly told that the span tests were memory tests and that it was necessary to focus their attention on the stimuli. Instructions were given orally and in Portuguese.

## Results and Discussion

Table 1 shows the Means (M) and Standard Deviations (SD) of the span tests. Two scores for both the operation-word span test and speaking span test were obtained. In the operation-word span test, a subjects' strict score refers to the number of words he/she was able to recall in their exact form and order of presentation in the set. A subject's lenient score refers to the number of words he/she was able to recall including those words in a slightly modified form and/or out of their original order of presentation in the set. In the speaking span test, following Daneman (1991) and Mota (1995), a speaking span strict score refers to all the grammatical sentences the subject produced containing the target word in its exact form and order of presentation. A speaking span lenient refers to the number of grammatical sentences a subject produced including those in which the target word was in a form and/or order other than that of presentation.

**Table 1.** Mean performance and standard deviations for measures of working memory capacity (descriptive statistics)

| | M | SD |
|---|---|---|
| Operation word span test (strict) | 45.8 | 12.1 |
| Operation word span test (lenient) | 44.1 | 12.8 |
| Speaking span test (strict) | 16.8 | 6.3 |
| Speaking span test (lenient) | 21.4 | 4.9 |

N=11

Tables 2 and 3 show the Means (M) and Standard Deviations (SD) for variables of fluency in the description and narrative tasks, respectively.

**Table 2.** Mean performance and standard deviations for variables of L2 fluency in the description task

| | M | SD |
|---|---|---|
| Amount of speech (1) | 370.2 | 224.0 |
| Amount of speech (2) | 381.1 | 224.7 |
| Speech rate | 71.9 | 22.2 |
| Filled nonlexical pauses | 14.2 | 11.3 |
| Filled lexical pauses | 8.7 | 8.3 |
| Unfilled pauses | 2.5 | 3.1 |
| Micropauses | 45.5 | 23 |
| Hesitations | 9.9 | 4.9 |

N=11

**Table 3.** Mean performance and standard deviations for variables of L2 fluency in the narrative task

| | M | SD |
|---|---|---|
| Amount of speech (1) | 534.5 | 368.4 |
| Amount of speech (2) | 545.5 | 368.6 |
| Speech rate | 88.2 | 18.0 |
| Filled nonlexical pauses | 12.6 | 10.1 |
| Filled lexical pauses | 7.8 | 8.6 |
| Unfilled pauses | 0.5 | 1.2 |
| Micropause | 52.2 | 30.7 |
| Hesitations | 6.4 | 8.2 |

N=11

As can been seen from Tables 2 and 3, the narrative task elicited a greater amount of speech production than the descriptive task. However, to answer the main question of the present study, no significant correlations were found between measures of working memory capacity and variables of fluency in either task. Tables 4 and 5 show the results of Pearson Product Moment Correlations.

**Table 4.** Correlations between the operation-word span test (strict and lenient) and variables of fluency

| | OWSTs | p-value | OWSTl | p-value |
|---|---|---|---|---|
| Speech rate D | 0.04 | 0.91 | -0.15 | 0.65 |
| Speech rate N | 0.08 | 0.82 | 0.09 | 0.70 |
| Amount of speech 1D | 0.30 | 0.36 | 0.40 | 0.22 |
| Amount of speech 1N | 0.20 | 0.5 | 0.34 | 0.31 |
| Amount of speech 2 D | 0.20 | 0.33 | 0.40 | 0.22 |
| Amount of speech 2 N | 0.21 | 0.54 | 0.34 | 0.29 |
| Filled nonlexical pauses D | -0.11 | 0.7 | 0.18 | 0.58 |
| Filled nonlexical pauses N | 0.03 | 0.9 | -0.00 | 0.98 |
| Filled lexical pauses D | 0.43 | 0.18 | -0.21 | 0.53 |
| Filled lexical pauses N | 0.34 | 0.30 | 0.20 | 0.55 |
| Unfilled pauses D | -0.16 | 0.62 | 0.04 | 0.89 |
| Unfilled pauses N | -0.08 | 0.81 | 0.03 | 0.93 |
| Micropauses D | 0.44 | 0.17 | 0.54 | 0.08 |
| Micropauses N | 0.23 | 0.4 | 0.37 | 0.26 |
| Hesitations D | -0.09 | 0.7 | 0.35 | 0.28 |
| Hesitations N | 0.11 | 0.73 | 0.29 | 0.39 |

N=11

D=description N=Narrative

**Table 5.** Correlations between the speaking span test (strict and lenient) and variables of fluency

| | SSTs | p-value | SSTl | p-value |
|---|---|---|---|---|
| Speech rate D | 0.23 | 0.50 | 0.20 | 0.55 |
| Speech rate N | 0.09 | 0.80 | -0.11 | 0.73 |
| Amount of speech 1D | 0.16 | 0.63 | 0.01 | 0.98 |
| Amount of speech 1N | -0.17 | 0.61 | -0.40 | 0.22 |
| Amount of speech 2 D | 0.16 | 0.64 | 0.00 | 0.99 |
| Amount of speech 2 N | -0.18 | 0.60 | -0.41 | 0.20 |
| Filled nonlexical pauses D | 0.47 | 0.14 | 0.23 | 0.50 |
| Filled nonlexical pauses N | -0.11 | 0.75 | -0.32 | 0.34 |
| Filled lexical pauses D | -0.28 | 0.41 | -0.28 | 0.40 |
| Filled lexical pauses N | -0.18 | 0.58 | -0.31 | 0.34 |
| Unfilled pauses D | -0.08 | 0.80 | -0.27 | 0.42 |
| Unfilled pauses N | -0.21 | 0.54 | -0.35 | 0.28 |
| Micropauses D | 0.14 | 0.68 | 0.02 | 0.95 |
| Micropauses N | -0.29 | 0.38 | -0.44 | 0.17 |
| Hesitations D | 0.25 | 0.45 | 0.33 | 0.31 |
| Hesitations N | -0.11 | 0.75 | -0.23 | 0.50 |

N=11
D=Description N=Narrative

Since these findings were not expected, two other ways of analysing the results were considered and new correlations were computed. First, subjects were divided into 2 groups according to level of proficiency: 6 in the intermediate group and 5 in the advanced group. Again, no significant correlations were found between measures of working memory capacity and variables of fluency in either group.

The analysis of subjects' speech samples shows no statistically significant differences among subjects when they are compared as intermediate and advanced learners (t-tests were performed comparing the means of the two groups in all variables of fluency). Previous research (e.g. Riggenbach, 1990; Ejzenberg, 1995; Freed, 1995) has shown that there are few statistically significant differences between more and less fluent speakers. Thus, if we consider the subjects in the advanced group to be more fluent than those in the intermediate group, finding no statistically significant differences among them in the quantitative measures of fluency analyzed is consistent with previous research.

Second, because the oral fluency tasks were not timed, there was a great variation in the amount of speech and speech rate of each subject, which obviously influenced the frequency of the other variables used in the present study (the various types of pauses and hesitations). Thus, a rate for frequency of occurrence of each type of pause and hesitations was computed for each subject and new correlations between these measures and memory spans were computed. As shown in Tables 6 and 7, no significant correlations were found between the rates and spans:

**Table 6.** Correlations between the speaking span test and rate of pauses and hesitations

| | SSTs | p-value | SSTl | p-value |
|---|---|---|---|---|
| Filled nonlexical pauses D | 0.07 | 0.84 | -0.02 | 0.95 |
| Filled nonlexical pauses N | -0.28 | 0.39 | -0.38 | 0.25 |
| Filled lexical pauses D | -0.29 | 0.37 | -0.23 | 0.49 |
| Filled lexical pauses N | -0.25 | 0.46 | -.036 | 0.27 |
| Unfilled pauses D | -0.24 | 0.47 | -0.33 | 0.32 |
| Unfilled pauses N | -0.25 | 0.46 | -0.37 | 0.26 |
| Micropauses D | -0.37 | 0.20 | -0.22 | 0.51 |
| Micropauses N | -0.39 | 0.23 | -0.44 | 0.18 |
| Hesitations D | -0.18 | 0.58 | 0 .05 | 0.89 |
| Hesitations N | -0.26 | 0.44 | -0.36 | 0.27 |

N=11
D=DescriptionN=Narrative

**Table 7**. Correlations between the operation-word span test and rate of pauses and hesitations

| | OWSTs | p-value | OWSTl | p-value |
|---|---|---|---|---|
| Filled nonlexical pauses D | 0.07 | 0.82 | 0.12 | 0.73 |
| Filled nonlexical pauses N | -0.21 | 0.53 | -0.23 | 0.49 |
| Filled lexical pauses D | -0.24 | 0.47 | -0.26 | 0.44 |
| Filled lexical pauses N | 0.07 | 0.82 | 0.07 | 0. 83 |
| Unfilled pauses D | -0.04 | 0.90 | -0.05 | 0.88 |
| Unfilled pauses N | -0.06 | 0.87 | -0.06 | 0.86 |
| Micropauses D | 0.23 | 0.49 | 0.21 | 0.53 |
| Micropauses N | 0.03 | 0.91 | 0.02 | 0.94 |
| Hesitations D | -0.02 | 0.96 | -0.03 | 0.94 |
| Hesitations N | 0.18 | 0.59 | 0.17 | 0.61 |

N=11
D=Description N=Narrative

Correlations performed among the variables of fluency revealed that those subjects with a faster speech rate also presented a greater amount of speech in both the description and narrative tasks; that is, those who speak faster also speak more, two salient features of fluent speech production ($r(11) = 0.70$, $p < 0.01$, for amount of speech 1 and 2 and speech rate in the descriptive task, and $r(11) = 0.67$, $p < 0.02$, for the narrative task). Other significant correlations were found among the variables of fluency, indicating the interdependency of these variables and a tendency for subjects with a faster speech rate to present not only a greater amount of speech but also a greater number of micropauses and filled nonlexical pauses. These correlations were consistent in both the description and narrative tasks, which might suggest that, at least for these subjects, there was no task effect on their display of fluency.

However, the main concern of this study is the relationship between working memory capacity and L2 fluency. Finding no significant correlations between these two variables contradicts a massive body of research developed within cognitive psychology that has given working memory a central role in human cognition. In trying to explain the results of the present study, I will argue that current theories of working memory capacity are not developed enough to account for the cognitive processes involved in speech production.

As pointed out by Engle (1996), working memory capacity has predicted performance in a number of real-world cognitive tasks, including (1) reading and listening comprehension (Daneman & Carpenter, 1980, 1983), (2) learning to spell (Ormrod & Cochran 1988), (3) following directions (Engle, Carullo, & Collins, 1991), (4) vocabulary learning (Daneman & Green, 1986), (5) notetaking (Kiwera & Benton, 1988), (6) writing (Benton, Kraft, Glover, & Plake (1984), (7) language comprehension (King & Just, 1991; McDonald, Just & Carpenter, 1992), (8) learning of complex information (Shute, 1991), (9) first language oral production (Danemann, 1991), and (10) foreign language reading comprehension (Harrington & Sawyer, 1993). The correlations between working memory capacity and all of the tasks mentioned above have been explained in different ways, but, basically it is the number of words or digits recalled,i.e., the span, that predicts subjects' performance.

The memory span test used in these studies is generally a variation of Daneman and Carpenter's (1980) reading span test, a complex memory span test that taxes both functions of working memory - processing and storage. It is in language comprehension, and more specifically, in reading comprehension, that this complex measure of working memory capacity has proven to be a useful methodology. The general finding has been that individuals with a larger working memory capacity, i.e., with a larger span, are better language comprehenders as assessed by a number of comprehension tasks. Researchers (e.g., Daneman & Carpenter, 1983) have explained this relationship by claiming that subjects' larger working memory capacity is a result of the speed with which they process information: since they have efficient processing, they use less of their total capacity for this processing, leaving it for the words that have to be recalled. Such behavior predicts reading performance: these individuals apply efficient comprehension processes when reading, using most of their capacity to store the subproducts of these comprehension processes and later integrate them to form a representation of the text. In this view, working memory capacity is translated as processing efficiency which is represented as span size.

To further verify the relationship between working memory capacity and language, Daneman (1991) designed the speaking span test. This test taxes working memory capacity under language production processes and was used to predict first language oral fluency. In her 1991 study, Daneman defines working memory capacity as "the ability to coordinate the processing and temporary storage functions of working memory" (p. 457) and the speaking span is proposed as a measure of such an ability. In her results, Daneman obtained significant correlations between subjects' speaking span and oral fluency and argued that individuals had a larger working memory capacity as measured by the speaking span test because of their efficient speech production processes: since they applied efficient speaking processes, they had a greater amount of capacity available for storing the to-be-remembered words, thus their larger span. These efficient speech production processes are noticed when individuals engage in an act of speaking: they are more fluent. Nevertheless, there are several drawbacks in Daneman's study, which are of crucial importance to explain why the present study failed to find significant correlations between working memory capacity as measured by the speaking span test and L2 oral fluency.

First, let's scrutinize what happens when subjects are required to perform the speaking span test as designed by Daneman. Subjects are presented with increasingly longer sets of two to six words. After a set is presented, subjects are required to generate aloud a sentence for each word in the set. The test is assumed to require the use of both functions of working memory capacity: storing the words presented and processing a sentence for each one. Such storage and processing is also assumed to take place when we speak. In addition, it is how much the subject is able to store that will define his/her working memory capacity and thus his/her fluency when speaking. However, what is it exactly that we store when we speak, be it in an L1 or L2?

Daneman does give us a clue: "speakers must *plan* what to say and temporarily *store* the plans until ready to execute them as words, phrases, and sentences" (p. 446). Moreover, she adds that execution might start without the speaker having finished planning. Hesitations and pauses are evidence, in Daneman's view, that failures occurred in either the planning or execution phases. This general conceptualization of speaking leaves a problem with the notion of "plan". "Plan" can be understood either as declarative knowledge - speakers must think of the **content** of what they are going to say as well as how they are going to say it, and this definitely requires some room in our working memory if it is to be maintained - or procedural knowledge - the cognitive plans or processes that are actually carried out when we think of the content and of how we are going to say it (for instance, access to and retrieval of concepts, lexical items, grammatical structures and phonological information). Whatever it is that working memory stores when we speak, it is left unclear whether the speaking span test is reflecting such storage. It seems that the speaking span test is measuring how many items one can hold, i.e., the quantity of information, whereas when we speak the question seems to be related to "how much" we can say, i.e., quality of information, which is represented in content.

Working memory capacity, in this view, is taken as "size" for holding a given number of items and such size is a result of processing efficiency, the second drawback in Daneman's (1991) study. It is not clear what is meant by the notion of "processing efficiency", and what is generally implied is that it is equated to speed of processing (Richardson, 1996). In any event, such processing efficiency is not assessed in the speaking span test, and one has to assume that subjects with a larger working memory span applied "efficient" processes, without any clues as to the nature of this efficiency. The sentences produced in the background task of the speaking span test are not judged as to content, adequacy or pragmatic value, for instance. Analyzing the sentences produced by the subjects in the present study shows how difficult it is to talk about processing efficiency in the performance of the background task. To mention but one example, these were the sentences produced by subject 10 for the last set of words:

> The desk is full of paper.
> The road is closed.
> The glass is empty.
> My brain is working.

Subject 8, on the other hand, chose to produce the following sentences:

> On my desk there's a lot of papers.
> There are serious problems with Brazilian roads.
> Nowadays the medical sciences can answer a lot of questions related to the human brain.

That is, while subject 10 was able to recall a greater number of to-be-remembered words, the sentences produced were relatively simpler in content compared to those produced by subject 8, who missed one word. At any rate, my point here is that these qualitative differences are not accounted for in the speaking span test, which takes into consideration only the number of items recalled. What should we do, for instance, when the subject recalls the word "wave" but generates a sentence such as "the wave is in the sea"?

The third drawback in Daneman (1991) has to do with her concept and assessment of fluency. Clearly, fluent speech is taken as speech that is fast and accurate in articulation. The variable used in a picture description task aimed at eliciting a monologue was number of words produced in the allotted time (one minute). The other variable was accuracy of articulation in a reading aloud task and in a spoonerism task.

Such concept and measures of oral fluency seem to be inadequate, especially when taken from the perspective of foreign/second languages, applied linguistics, and pedagogy, where the term has been given careful attention and tools have been developed to better understand what it means to be "fluent". As already pointed out, the notion has different values in first and second languages, but even in our first language, speech rate is only one characteristic of fluency, among many others, and number of words is not a reliable method for determining it. Applied linguists have used, instead, the notion of semantic units to determine speech rate.

To have a more accurate assessment of fluency, in the present study more refined variables were used that have been shown to be salient features of speech production. These variables are interdependent and present in both first and second language speech. They are also all temporal variables that, in one way or another, reflect speed of processing. However, the speed of processing present in the speaking span test does not seem to match that of actual speech production, when this processing is analyzed under more sensitive variables.

The operation-word span test, the second measure of working memory capacity, was used in the present study as an alternative to overcome the limitations of the speaking span test. Originally, the operation-word span test was designed to investigate whether working memory capacity was task specific or a general capacity (Turner & Engle, 1989). The assumption underlying the test is thus different from Daneman and colleagues': the background task (solving a math problem) is not related to the cognitive task the subject will be required to perform and, in Turner and Engle's view, this gives a "purer" measure of working memory capacity: it allows us to determine whether it is general processing efficiency or span size that is driving performance on complex cognitive tasks.

As already reported, no correlations were found between the operation-word span test and variables of fluency. Since the test was subject-paced, it seems that subjects in the present study would take the time they needed to solve the math problems and rehearse the to-be-remembered words as many times as they wanted before pressing the enter key for the next operation-word pair in the set, thus minimizing any pressures in their working memory capacity. The effect of this test was just the opposite of that of the speaking span test: while in the latter there is the pressure of time to produce the sentences in order not to forget the words in the set (thus provoking speed of processing that is unlikely to occur in normal speech production), the former does not impose any time pressure, which is also unlikely to occur in normal speech production. Another factor that indicates that the operation-word span test did not impose much pressure on subjects' working memory capacity is their mean performance on the task: 56.8, SD (2.4). Thus, it seems that, because there was no time pressure, working memory capacity demands were minimized and subjects would allocate their cognitive resources first to solve the operation string, then to memorize the words in the set. As already reported in this study, subjects' performance on the tests was tape-recorded and it is evident from their protocols that the resolution of the operation string was carried in Portuguese. Such evidence might suggest that, for these subjects, the operation-word span test was performed, in fact, as two separate tasks: one focusing more heavily on processing (the math problem) and the other being simply a word span task.

## Summary and Conclusions

The objective of the present study was to investigate whether measures of working memory capacity currently used in cognitive psychology correlate with temporal variables of L2 oral fluency. Eleven English-as-a-foreign-language undergraduate students participated in this study—6 at the intermediate level and 5 at the advanced level (fourth and seventh semesters of the Letras course at the Federal University of Santa Catarina, respectively). Subjects' working memory capacity was assessed by means of the operation-word span test (Turner and Engle, 1989) and the speaking span test (Daneman & Green, 1986; Daneman, 1991). Subjects' L2 oral fluency was assessed by means of a picture description task and a narrative task. There were no significant correlations between measures of working memory capacity and variables of L2 oral fluency. There was no statistically significant difference among subjects' oral fluency, even when they are compared as intermediate and advanced groups. The results obtained in the present study were explained by arguing that the working memory capacity measures used in the test are not good predictors of L2 fluency since they are not sensitive to the cognitive processes involved in L2 speaking. As regards L2 oral fluency, the findings of the present study tend to corroborate results of previous results in that there seem to be few statistically significant differences between more and less fluent speakers, and in that those speakers with a faster speech rate also tend to present a greater amount of speech, with fewer unfilled pauses and a greater number of micropauses. Moreover, there seems to be no statistically significant task effect on the display of fluency of the subjects who participated in this study.

Although the findings of the present study contradict an impressive body of research investigating the role of working memory capacity in the performance of complex cognitive tasks, it is still worth insisting on the relationship between this construct and L2 oral fluency. It is now well accepted in various areas of research dealing with learning that the human beings are limited-capacity information processors and that this limited capacity constrains our performance on a wide range of tasks in various ways. This limited-capacity system been conceptualized in different forms, but a widely accepted one is that it is an active portion of our cognition in which information from long-term memory is processed and held temporarily for the performance of a task. Because it is an active portion of our cognitive architecture, where information is handled on line and processes are managed so that they can be applied, this system has been called a working memory. A number of research paradigms have been used to better explain the structure and function of working memory, but attention has been given with much more frequency to the storage capacity component and function of the system. The active coordination of the cognitive processes we need to accomplish a task is poorly understood in current cognitive psychology and a lot of what we know about it is inferred from the studies focusing on the memory aspect - the span size - of the system. It seems that so far we already know that this span size is limited and that individuals vary in this size. However, how important it is to accomplish a task satisfactorily seems to be irrelevant when we take into consideration evidence such as that obtained in the present study. A much more interesting question seems to be how human beings coordinate mental processes and how this relates to performance; i.e., what it is that efficient human beings do to perform a task well. Equally important is to investigate the role that the status of representations plays on these mental processes. L2 speaking, in this sense, seems a case in point: from the speech samples of the subjects who participated in this study, it is possible to observe that their cognitive processes for speaking are normally executed, but their speech is disrupted by the quality of the knowledge they have of the L2—mostly declarative knowledge about words and grammar. Thus, it might very well be that display of fluency in L2 is related to the nature of knowledge representation rather than cognitive capacity. A longitudinal study focusing on the development of L2 oral fluency and inspecting the strategic behavior of the speakers might help cast some light on these issues.

**Notes**

1 Short-term memory as it has been described in current cognitive psychology is assumed to be a small component of the working memory system, responsible for the storage of information for up to 20 seconds (Ashcraft, 1994).

2 For an extensive review of the literature on individual differences in working memory capacity and reading comprehension, see Tomitch (1995).

3 See Sinsabaugh and Fox (1986) for previous criticism of the Oral Slip Task as designed by Motley, Baars, and Camdem (1983).

4 Extending Lennon's point of view, Ejzenberg (ibid.) states that, in Applied Linguistics, fluency has generally been equated to language proficiency, that is, to the individual's overall ability in the L2.

## References

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

Baddeley, A. D. (1990). *Human memory: Theory and practice*. Hove, UK: Lawrence Erlbaum Associates.

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*, *44A*, 1-31.

Baddeley, A. D. & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation, Vol. 8*. New York: Academic Press.

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.

Carlson, R. A., Sullivan, M., & Schneider, W. (1990). Practice effects and composition: A reply to Anderson. *Journal of Experimental Psychology: Learning, Memory, and Cognition*,_*15*, 531-533.

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.

Crookes, G. (1991). Second language speech production research: A methodologically-oriented review. *Studies in Second Language Acquisition*, *11*, 367-383.

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

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

Dechert, H. W. (1983). How a story is done in a second language. In C. Faerch & G. Kasper (Eds.), *Strategies in interlanguage communication.* London: Longman.

Dechert, H. W. & Raupach, M. (1980).(Eds.). *Temporal variables in speech.* The Hague: Mouton.

Dechert, H. W. & Raupach, M. (1987). *Psycholinguistic models of production.* Norwood, NJ: Ablex.

Deschamp, A. (1980). The syntactical distribution of pauses in English spoken as a second language by French students. in H. Dechert & M. Raupach (Eds.*), Temporal variables in speech: Studies in honor of Frieda Goldman-Eisler.* The Hague: Mouton.

Ejzenberg, R. (1992*). Understanding nonnative oral fluency: the role of task sturcture and discourse separability.* Unpublished doctoral dissertation, State University of New York at Albany.

Ejzenberg, R. (1995). *Discourse markers and context-dependency in oral fluency assessment.* Paper presented at the 1995 American Association of Applied Linguistics Conference, Long Beach, California.

Engle, R., Cantor, J., Carullo, 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., Carullo, J., & Collins, K. (1991). Individual differences in the role of working memory in comprehension and following directions in children. *Journal of Educational Research*, *84*, 253-262.

Fillmore, C. (1979). On fluency. In C. J. Fillmore, D. Kempler, & W. S. Y. Wang (Eds.). *Individual differences in language ability and language behavior*. New York: Academic Press.

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: John Benjamins.

Fromkin, V. (1973). *Speech errors and linguistic evidence*. The Hague: Mouton Publishers.

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

Grosjean, F. (1982). *Life with two languages*. Cambridge, Massachusetts.: Harvard University Press.

Harrington, M., & Sawyer, M. (1993). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*, 25-38.

Hitch, G., & Baddeley, A. . (1976). Verbal reasoning and working memory. *Quarterly Journal of Experimental Psychology*, *28*, 603-21.

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

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

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

Macdonald, M. C., Just, 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 .

Maxwell, J. & Heritage, J. (1984). *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press.

Myiake, 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.

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 production*. Tübingen: Narr.

Mota, M. B. (1995). *Working memory capacity and L2 fluent speech production*. Unpublished MAdissertation. Florianópolis: UFSC.

Olynik, M., D'anglejan, A., & Sankoff, D. (1990). A quantitative and qualitative analysis of speech markers in the native and second language speech of bilinguals. In C.R. Scarcella, E.S.Andersen, & S.D. Krashen. (Eds.), *Developing communicative competence in a second language*. New York, NY: Newbury House.

Olynik, M., Sankoff, D., & D'anglejan, A. (1983). Second language fluency and the subjective evaluation of office cadets in a military college. *Studies in Second Language Acquisition*, *5*, 213-249.

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.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and native fluency. in J.C. Richards & R.W. Schmidt. (Eds.), *Language and communication* (pp.191-226). New York: Longman.

Postma, A., Kolk, H. & Pole, D. J. (1990). On the relation among speech errors, disfluencies, and self-repairs. *Language and Speech*, *33*, 1, 19-29.

Raupach, M. (1980).Temporal variables in first and second language speech production. in H. Dechert & M. Raupach (Eds.), *Temporal variables in speech: Studies in honor of Frieda Goldman-Eisler*. The Hague: Mouton.

Rehbein, J. (1987). On fluency in second language speech. In H.-W. Dechert & M.Raupach (Eds.), *Psycholinguistic models of production*. Tübingen: Narr.

Riggenbach, H. (1991). Toward an understanding of fluency: A microanalysis of nonnative speaker conversations. *Discourse Processes*, *14*, 423-41.

Richardson, J. T. E. (1996). 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). New York: Oxford University Press.

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

Shute, V. J. (1991). Who is likely to acquire programming skills? *Journal of Educational Computing Research, 7*, 1-24.

Temple, L. (1992). Disfluencies in learner speech. *Australian Review of Applied Linguistics*, *15*, 29-44.

Tomitch, L. M. B. (1995). *Reading: Text organization perception and working memory capacity*. Unpublished doctoral thesis. Florianopolis: Universidade Federal de Santa Catarina.

Tomlin, R. S. & Gernsbacher, M. A. (1994). Cognitive foundations of second language acquisition. *Studies in Second Language Acquisition*, *16*, 129-132.

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

## APPENDIX A

Operation-word span test (sample)

| | | |
|---|---|---|
| (6X2)+2=14 | V | HALL |
| (3/1)+1=4 | V | NEAR |
| (7/1)+6=12 | V | SKILL |
| (7/1)+2=7 | V | MEAL |
| (10/1)-1=11 | V | NOSE |
| (6X4)-1=25 | V | BUY |
| (9/1)-8=18 | F | GREEN |
| (5X1)+1=5 | V | DANCE |
| (9X1)-9=1 | F | GUEST |
| (7X7)+1=50 | V | LOCK |

## APPENDIX B

Speaking span test in English (SSTE)

| | | | |
|---|---|---|---|
| (2) | Cake | (5) | bank |
| | hand | | shirt |
| (2) | week | | egg |
| | rain | | date |
| (3) | duck | | hair |
| | pen | (6) | clock |
| | gas | | wave |
| (3) | club | | tool |
| | spring | | coat |
| | knife | | map |
| (4) | arm | | year |
| | sky | (6) | cow |
| | deer | | pair |
| | ball | | drum |
| (4) | desk | | sea |
| | road | | bus |
| | glass | | west |
| | brain | | |
| (5) | sun | | |
| | mouth | | |
| | key | | |
| | bag | | |
| | file | | |