WORKING MEMORY CAPACITY AND FLUENCY, ACCURACY, COMPLEXITY, AND LEXICAL DENSITY IN L2 SPEECH PRODUCTION

MAILCE BORGES MOTA

Universidade Federal de Santa Catarina
mailce@cce.ufsc.br

Abstract
This study investigated whether there was a relationship between working memory capacity and L2 speech production. 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) 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 fluency, accuracy, complexity, and weighted lexical density. To explain the relationship between working memory capacity 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.
Keywords: individual differences, working memory capacity, L2 speech production

Resumo
Este estudo investigou se há relação entre a capacidade de memória de trabalho e a produção oral em L2. Os participantes do estudo foram 13 alunos de inglês como segunda língua, de nível avançado, da Universidade de Minnesota. A capacidade de memória de trabalho dos participantes foi medida através do speaking span test (Daneman, 1991), aplicado na L2. A produção oral em L2 foi eliciada através de uma tarefa de descrição e uma de narração. Quatro aspectos da produção oral foram avaliados: fluência, precisão gramatical, complexidade gramatical e densidade lexical. A análise estatística demonstrou que, em ambas as tarefas de produção oral, a capacidade de memória de trabalho, medida pelo speaking span test, se correlaciona positivamente com fluência, precisão gramatical e complexidade gramatical, como previsto, e negativamente com densidade lexical, contrariamente à previsão. A análise também demonstrou que o speaking span test prevê de forma significativa o desempenho oral dos participantes no que diz respeito à fluência, precisão gramatical e complexidade gramatical, explicando parcialmente a variação individual de desempenho. Por fim, a análise indica uma tendência de interação entre fluência, precisão gramatical, complexidade gramatical e densidade lexical. Para explicar a relação entre capacidade de memória de trabalho e desempenho oral em L2, propõe-se a codificação gramatical como uma subtarefa do processo de produção oral cuja natureza cognitivamente complexa exige o controle e regulação de recursos atencionais.
Palavras-chave: diferenças individuais, capacidade da memória de trabalho, produção oral em L2.

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; Wiese, 1984). Speaking is the 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 (Bock, 1996; Ratner & Menn, 2000).

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, which 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 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 (e.g., 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).

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.

Researchers have consistently shown that the limited capacity of working memory differs among individuals (cf. Barrett, Tugade, & Engle, 2004). Thus, evidence accumulates demonstrating that individual differences in working memory capacity are related to several aspects of L1 reading comprehension (e.g., Daneman & Carpenter, 1980, 1983; Daneman & Green, 1986; Masson & Miller, 1983; Miyake, Just, & Carpenter, 1994; Tomitch, 1995), 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 (Barrett, Tugade, & Engle, 2004; Cantor & Engle, 1993; Conway, Kane, & Engle, 2003; 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 (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, 2000). 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 study 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 objective of the present study is to investigate whether there is a relationship between working memory capacity and L2 speech production. Following mainstream research, working 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.

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.

The Study

Research question and hypotheses

The present study pursued the following research question: is there a statistically significant relationship between working memory capacity, as measured by the speaking span test, and measures of fluency, accuracy, complexity, and weighted lexical density in L2 speech production? From this research question, 4 hypotheses follow:

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.

Method

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.

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 Turkish. 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.

Instruments

Materials and equipment

The experiment consisted of three tasks: one task aimed at measuring working memory capacity and two tasks aimed at eliciting speech production in the L2. The working memory task was conducted using an IBM ThinkPad Laptop Computer with a VGA monitor. Participants' responses in all three 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).

Assessment of working memory capacity

Participants' working memory capacity was assessed by means of the speaking span test (Daneman, 1991). 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. 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.

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.

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.

Picture description task

Participants were presented with a colorful picture taken from a popular magazine. 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. 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. The task would begin when participants signaled they were ready.

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. Participants were also instructed to talk as much as possible.

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. All participants' speech samples were tape-recorded and later transcribed[1].

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. 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.

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:

(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; 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.

Accuracy

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.

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.

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.

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.

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 was 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]

Reliability of span tests

Reliability estimates were computed for the speaking span test using Cronbach's alpha formula of internal consistency, also used by Turner and Engle (1989) and Engle et al. (1992). Following Turner and Engle (1989), the reliability estimates were based on three composite scores computed for every subject in the span test. As already explained, the span test was organized in three trials each of two, three, four, five, and six sequences of to-be remembered items. 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 reliability estimate of the speaking span was .88. 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.

Data analysis

The goal of this research was to examine the relationship between working memory capacity, as measured by the speaking span task, and L2 speech production, as measured by several variables covering fluency, accuracy, complexity, and lexical density of speech. 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.

Results

Tables 1 and 2 present the descriptive statistics of the speaking span test and the four measures of L2 speech production, respectively:

Table 1

Descriptive Statistics for the Speaking Span Test (SST)

	SST
Mean	24.5
SD	10.82
Minimum	11
Maximum	47

N=13

		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

For the span measure, the highest possible score was 60. As can be seen from Table 1, the speaking span test scores varied over a 36-point range with a large standard deviation.

For the L2 speech production measures, the results from the descriptive statistical analyses 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 counterbalanced 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.

The speaking span test and fluency, accuracy, complexity, and lexical density

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 tables. Table 3 presents the results from the Pearson Product-Moment Correlation performed for the relationship between the speaking span test and speech rate unpruned, speech rate pruned, and mean length of run. Table 4 presents the results for the relationship between the speaking span test and number of silent pauses and hesitations per minute.

Table 3

Pearson Product Moment Correlations 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	.73**	.72**	.70**	.69**	.68**	.62*

Pearson Product Moment Correlation (*r*)
N=13
* $p < 0.05$
** $p < 0.01$

As can be seen in Table 3, results from the Pearson Product Moment Coefficient of Correlation 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.

Table 4

Pearson Product Moment Correlations 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	-.43	.20	-.22	.42

Pearson Product Moment Correlation (*r*), N=13
* $p < 0.05$ (one-tailed)
** $p < 0.01$ (one-tailed)

As can be seen in Table 4, 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.

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 presents the results from The Pearson Product Moment Coefficient of Correlation:

Table 5

Pearson Product Moment Correlations 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

[a] Pearson Product Moment Correlation (r), N=13
* $p < 0.05$
** $p < 0.01$
Note. SST: speaking span test

As can be seen in Table 5, Hypothesis 2 seems to be supported. Results from The Pearson Product Moment Coefficient of Correlation 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, $p < 0.05$. 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.

The correlation coefficients reported in Table 5 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 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. 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.

Finally, for Hypothesis 4, the results reported in Table 5 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 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.

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.

Simple linear regressions

Table 6 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 6

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, $p < 0.05$

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. 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 with mean performance on these aspects increasing as working memory capacity increases.

As Table 6 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 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 6 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. Together, these results indicate that despite the fact that there seems to exist a linear relationship between working memory capacity, 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, a great proportion of the variation in these dimensions is not explained by working memory capacity alone.

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 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 (Levelt, 1989): more specifically the processes involved in the grammatical encoding of the message, the construction of a syntactic structure for the message. 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* (Engle & Oransky, 1999), 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 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 results from a trade-off between fluency, accuracy, complexity, and weighted lexical density.

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.

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 sub-task 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). Thus, during L2 speech production both L1 and L2 lemmas are activated simultaneously. Activation spreads to the corresponding L1 and L2 word forms. 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; 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.[2]

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, 2001; Kane & Engle, 2001; Rosen & Engle, 1997 and 1998). These activities, according to Engle and colleagues include situations (a) when it is necessary to apply *activa-*

tion 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 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).

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.

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.

The statistical results obtained in the current study 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 discussion 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.

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 and this will be discussed in terms of a trade-off between fluency, accuracy, complexity, 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.

In the present study, grammatical encoding has been suggested to be a sub-task 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 (Griffin and Bock, 1998). Being in a high state of activation, these words 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).

Final remarks

The objective of the present study was to investigate the relationship between working memory capacity and L2 speech production. Being exploratory in nature, the study has a number of drawbacks, two of which being its sample size and the statistical techniques used to analyze the data. The results presented, thus, can be taken 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, Jaattuso, & 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 study constitutes a step towards understanding L2 speech production.

Notes

1 The transcription system used was adapted from Jefferson (1979), which has been used in studies on L2 speech production (Ejzenberg, 1992; Riggenbach, 1989).

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

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

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.

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.

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

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.

Barrett, L.F., Tugade, M.M., & Engle, R. (2004). Individual differences in working memory capacity and dual-processing theories of mind. *Psychological Bulletin,* 130, 4, 553-573.

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.

Butler-Wall, B.A. (1986). T*he frequency and function of disfluencies in native and non-native conversational discourse*. Unpublished doctoral dissertation, University of California at Los Angeles.

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 thought. 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.

Conway, A.R.A., Kane, M.J., & Engle, R.W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences,* 7, 547-552.

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.

Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research*, *2*0, 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.

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. (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., Möhle, D., & Raupach, M. (Eds.). (1984). *Second language productions*. Tübingen: Gunter Narr Verlag.

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.

Ejzenberg, R. (1992). *Understanding nonnative oral fluency: The role of task structure and discourse variability*. Unpublished doctoral dissertation, State University of New York.

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.

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

Engle, R.W., Kane, M.J., & 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., & 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.

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.

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.

Fortkamp, M.B.M. (2000). *Working memory capacity and L2 speech production: An exploratory study*. Unpublished doctoral dissertation, Universidade Federal de Santa Catarina.

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.

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

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.

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

Howell, W.C., & Cook, 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.

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. (2001). A controlled-attention view of working memory capacity. *Journal of Experimental Psychology: General*, 130, 169-183.

Kane, M.J., & Engle, R.W. (2001). Working memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 336-358.

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., 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.

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. (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.

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.

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.

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

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.

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.

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.

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.

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.

Perlow, R., Jattuso, De Wayne Moore, D. (1997). Role of verbal working memory in complex skill acquisition. *Human Performance*, *10*, 283-302.

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.

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.

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.

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.

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

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.

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

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

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, Florianópolis.

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.

Torres, A.C. G. (2003). Working memory capacity and readers' performance on main idea construction in L1 and L2. Unpublished doctoral dissertation, 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.

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

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.

Woltz, D.J. (1988). An investigation of the role of working memory in procedural skill acquisition. *Journal of Experimental Psychology: General*, *117*, 319-331.