DOI: 10.5007/2175-8026.2017v70n3p245

LEXICAL ACCESS IN L2 SPEECH PRODUCTION: A CONTROLLED SERIAL SEARCH TASK

Gicele Vergine Vieira*
Instituto Federal Catarinense
Blumenau, BR

Abstract

When it comes to lexical access in L2 speech production, working memory (WM) seems to play a central role, as less automatized procedures require more WM capacity to be executed (Prebianca, 2007). With that in mind, this article aims at claiming that bilingual lexical access qualifies as a controlled serial strategic search task susceptible to individual differences in WM capacity. Evidence in support of such claim is provided by the results of Prebianca (2010) study conducted so as to investigate the relationship between L2 lexical access, WMC and L2 proficiency. Prebianca (2010) findings indicate that bilingual lexical access entails underlying processes such as cue generation, set delimitation, serial search and monitoring, which to be carried out require the allocation of attention. Attention is limited and, as a result, only higher spans were able to perform these underlying processes automatically.

Key-words: L2 speech production; Lexical access; Controlled serial search; Working memory capacity.

1 Introduction

The production of intentional and fluent speech[1] has been said to require the orchestration of a number of mental operations involving conceptual and linguistic processes (Levelt, 1989). From an information-processing perspective (Shiffrin; Schneider, 1977), some of these processes are performed automatically and some are performed under attentional control. In his L1 adult speech production model, Levelt (1989) acknowledges that conceptualizing a message to be verbalized in one's language and monitoring the output of such verbalization are processes that require attention to be executed, since they are controlled by the speaker himself. On the other hand, linguistic processes such as selecting and retrieving words to express the conceptual message, giving the message sound and articulating it, are highly automatic processes that do not depend on attention to be performed.

As proposed by Levelt, the core process of speaking is word selection upon which all other linguistic processes operate. Word selection or lexical access,[2] as is usually referred to in the literature on speech production, is said to occur under competition. That is, when a concept specified in the conceptual message, activates a word in the mental lexical, this activation spreads along the lexico-semantic network, and several related words (words that share meaning or any other related characteristic) also become activated, competing for selection. The extent to which such competition interferes with the selection of the appropriate word is said to be related to how strong the connections between words are (de Groot, 1992). How this competition is solved by the lexical retrieval system is still a matter of

* Doutora em Letras Inglês pela Universidade Federal de Santa Catarina, com ênfase em Linguística Aplicada. Coordena o grupo de pesquisa Linguagem, Cognição e Tecnologia. Investiga a aquisição de língua estrangeira (L2) mediada pelo uso da tecnologia, bem como os aspectos cognitivos de aquisição e uso da L2. Atua como professora no ensino médio integrado, nos cursos de graduação e no Programa de Pós-Graduação em Educação do Instituto Federal Catarinense. Seu e-mail é gicelevpreb@gmail.com.

contention but, more importantly, because accessing words in L1 is so automatized, few selection errors are made and speech production generally proceeds smoothly to articulation.

The panorama seems to be a very different one when speech is produced in L2.[3] It is now widely accepted that L2 speakers hold a great amount of explicit and underdeveloped knowledge of the second language, thus resorting to more controlled processing, especially in initial learning[4] phases (Kormos, 2006). Because the L2 lacks automatization, speech production in the second language runs serially, thus causing L2 speech to be more hesitant, disfluent and open to L1 influence (Poulisse, 1997; Fortkamp, 2000; Kormos, 2006). Word retrieval, in this scenario, besides suffering from lack of automaticity, is also affected by lexical representations that lack strong connections with the L2 conceptual system, forming a less integrated lexicon in relation to L1(Kormos, 2006; de Groot, 1995), and by competition from other L2 and L1 related items. Serial processing of explicitly stored retrieval procedures, weaker lexical representations, and lexical competition render L2 lexical access an attention-demanding task.

Although L1 speech production models such as Levelt (1989) and Levelt et al. (1999) have provided a comprehensive account of the mental/cognitive processes involved in monolingual lexical access, several issues remain unresolved when turning to bilingual speech processing. Adding an L2 component to L1 models is far from being an easy and straightforward solution, for it raises a bunch of questions particularly in relation to message conceptualization, mental lexicon organization and lexical retrieval, such as language-specificity, the storage of L1 and L2 lexical items in the same network, and thus, lexical competition.

In this sense, it is important to remember that the retrieval of lexical items is directly affected by the way speech is conceptualized and the way the mental lexicon is conceived. This is particularly true because it is the concepts specified in the preverbal message that will determine which lexical items will be selected to verbalize the message one wants to express. These items will then be retrieved on the basis of their primitives or another mental representation and may activate related items in the intended and non-intended languages. The question of whether lexical items of both languages compete for selection has received great attention in bilingual research. There is nowadays plenty of evidence in favor of a language non-specific approach to lexical selection. On this view, items of both languages that share the same characteristics with the preverbal message are activated, competing for selection at the semantic and phonological level. Cascading models of L1 speech production have shown that not only the selected items send activation to their corresponding phonemes, but also non-selected ones have their phonological information activated (Morsella and Miozzo, 2002).

As an attempt to address the issues involved in L2 speech production, Kormos (2006) proposed an L2 speaking model based on Levelt et al.'s (1999) revised blueprint for the speaker. In this recent model, Kormos makes important assumptions regarding knowledge automatization in L2 and the way it affects speech production processes. According to her, because several lexical encoding procedures are not fully automatized in L2, bilingual speakers must have access to an additional knowledge store – a declarative store for syntactic L2 rules. With increasing proficiency the declarative knowledge of L2 rules may become automatized and then lexical processing may develop on a continuum, from serial to parallel processing, allowing for a more native-like speech production. As long as speakers depend on the use of declarative knowledge, lexical encoding can only be serially carried out, requiring more attentional control to be executed.

Research on bilingual lexical access and working memory capacity have shown that accessing and retrieving words in an L2 under competition of related lexical representations in the language in use (usually the speaker's L1) is an attention-demanding cognitive task subject to individual differences in goal maintenance and inhibition of distracting information (Kroll, Michael, Tokowicz and Dufour, 2002; Tokowicz, Michael and Kroll, 2004; Christoffels, de Groot and Kroll, 2006). Most of these studies, however, have not examined the extent to which WMC affects L2 lexical retrieval under within-language competition of semantically-related words at different levels of proficiency. In other words, the empirical evidence gathered so far does not tell us which processes are common to both bilingual lexical access (when retrieval entails L2 response competition) and WMC that cause them to be related or whether those processes change with increased proficiency in L2.

Given that retrieving words in L2 is not only about blocking L1 activation but also about fighting off L2 lexical competitors by adequately delimiting the search set and monitoring for adequate retrieval within a language system under development, in which lexical connections lack rich conceptual specifications and lexical retrieval procedures are not yet fully automatized relative to the ones in L1 (Kormos, 2006), there seems to be room for an investigation concerning the relationship between individual differences in WMC, bilingual lexical access and proficiency level in an L2 picture naming task conducted under the picture-word interference paradigm.

Having said that, the next section presents the controlled-attention view of WMC and its relevance for explaining memory retrieval and, consequently, bilingual lexical access.

2. Memory retrieval and the controlled-attention view of WMC

Since the earliest accounts of information processing (Atkinson and Shiffrin, 1968), there seems to be a consensus among cognitive psychologists that processing and maintenance of information in service of higher order cognition entails the use of limited capacity resources. These computation and storage processes are assumed by most cognitive researchers to be the basic executive functions of Working Memory (Baddeley and Hitch, 1974; Daneman and Carpenter, 1980; Turner and Engle, 1989; among others). When procedures for executing a certain task are not fully automatized, working memory resources are needed to selectively direct attention to those aspects of the task that need controlled processing to be executed. On this view, working memory limitations refer to limitations in the ability to control attention in order to focus on information which is relevant to the execution of the task by ignoring irrelevant stimuli.

The controlled attention view of WMC is adopted as a the major theoretical framework in the present discussion because it is closely related to the cognitive task being addressed, namely lexical access. For the sake of this discussion, lexical access involves naming a picture in the face of interference. To be able to execute this task efficiently, L2 speakers need to block interference in order to keep the main objective of the task active in WM and, thus, retrieve the lexical items from long-term memory quickly and accurately. Being able to suppress interfering stimuli is essential to perform the picture-naming task, and it is also one of the processes performed by WM. In addition, because the L2 is usually the less dominant and less practiced language, retrieving words in this language involves dealing with weaker connections between words and concepts, semantic competition due to overlap in meaning, as well as procedures which operate on explicit L2 knowledge, especially in initial L2 learning stages (Kormos, 2006). Therefore, it is reasonable to propose that bilingual lexical access constitutes a controlled serial strategic search susceptible to individual differences in working memory capacity. The studies reviewed in what follows contribute to strengthen this claim.

Conway and Engle (1994) set out to investigate the role of WM in retrieval by hypothesizing that individual differences in WMC might affect the retrieval of information from what they called primary and secondary memory in different ways. Primary memory (PM) was assumed to be the storage of information in an active state, that is, working memory. Secondary memory (SM) was taken as the repository of information stored for a longer period of time, that is, long-term memory.

A series of experiments using speeded search and verification tasks was carried out with 20 high-span subjects and 20 low-span subjects as determined by their scores on the OSpan. Subjects were submitted to a learning phase and a verification phase. In the learning phase, subjects were required to memorize 4 or 6 sets containing from 2 to 12 letters or words. During verification, two procedures were adopted: either the letter/word was preceded by the presentation of a number indicating the set in which it appeared during the learning phase, or it was displayed together with the set number. Whereas the first procedure was meant to measure primary memory search, the second ensured that information was inactive in second memory until both set number and probe were presented.

Experiments 1 and 2, designed to allow for interference effects due to the overlap in set membership, showed that high-span subjects differed from low-span subjects in retrieval from primary memory but not from secondary memory, as measured by RT scores. Experiments 3 and 4 were designed so as to avoid interference effects. That is, a letter/word could be the target in only one specific set. Results of experiment 3 replicated previous results showing that high- and low-span subjects did not differ in retrieval from secondary memory. However, the same experiment revealed that in the absence of interference, high- and low-spans performed similarly. Experiment 4, which aimed at replicating these findings with word rather than letter retrieval, also showed that high- and low-spans' RTs were not statistically different, suggesting that their performance was similar when retrieving items from primary memory without having to deal with response competition.

Taken together, the experiments conducted by Conway and Engle support the idea that when retrieval from primary active memory involves handling response competition, individual differences in the ability to suppress misleading information will account for better task performance. In other words, subjects with greater WMC are better able than those with less capacity to execute a set search that requires attentional and inhibitory resources. Set searching in secondary memory, on the other hand, was taken as an automatic process, since the time taken to bring the relevant information to an active state in primary memory did not vary as a function of individual differences in the ability to inhibit the activation of the wrong set. Clearly, the role of WM in retrieval proposed by the researchers is only prominent when competition and conflict need to be resolved.

Rosen and Engle (1997) further addressed the role of individual differences in WMC in retrieval. The basic assumption underlying their study follows Moscovitch's idea (1995 apud Rosen and Engle, 1997) that retrieval can occur either through associative or strategic processes. In associative retrieval, the presentation of a cue automatically leads to the retrieval of the target information. In strategic retrieval, on the other hand, the cue functions as a clue to where controlled search should start from. In other words, strategic retrieval implies that attention is necessary to delimit the search set appropriately. Consequently, WMC, which is supposedly unimportant to automatic retrieval, seems to play a salient role in strategic retrieval.

Most important to the present argumentation is that picture-naming in L2 seems to entail strategic procedures. When a picture is displayed, activation spreads along the lexical-semantic network and subjects associate the picture with its concept, either through L1 mediation or directly through L2 conceptual links. Once the concept is activated, a search for the correct word is initiated. This search is potentialized when the semantically-related word distractor is presented. Because the L2 lexical network is likely to be less intricate in relation to the L1 network, any item that shares common characteristics with the target will probably facilitate retrieval.

Rosen and Engle's main objective was to examine the importance of WMC to strategic retrieval. The set of experiments was designed basically to test whether high- and low-spans differed in the number of category exemplars they were able to retrieve while avoiding repetitions in load and no-load conditions. In a no-load condition, higher-span subjects generated more category exemplars than lower-spans. In contrast, under cognitive load, only higher-spans reduced the number of names retrieved. Lower-spans were unaffected by the concurrent digit-tracking task. The researchers suggested that because lower-spans did not have sufficient attentional resources to avoid repetitions, to generate cues to retrieving new names and to track digits simultaneously, they were unable to inhibit previous responses thus retrieving items more automatically than higher-spans. Higher-spans, on the other hand, experienced a reduction in the number of exemplars retrieved due to their greater ability to monitor for repetitions and search for new names at the same time, leading them to retrieve items in a more controlled fashion.

The explanation provided by Rosen and Engle for their findings seems to imply that the ability to suppress proactive interference is not the only one necessary to guarantee efficient retrieval from secondary memory. It seems that generating cues to delimit the response set and guide the search is also an important controlled attention task to be performed if retrieval is to be accomplished successfully. This claim has been further supported by Rosen and Engle (1998). Through a series of paired-associate tasks, the researchers demonstrated that lower WMC subjects had problems generating internal cues to guide the search for the correct item in secondary memory in relation to higher WMC subjects. Lower spans were both slower and less accurate during recall of items that were previously learned with a different pair-associate because they could not block intrusions from previous items.

In a more recent study, Unsworth and Engle (2007) demonstrated that retrieval of information from secondary memory, that is, from information outside the focus of attention (WM), stored in long-term memory, is governed by a discrimination process that involves the use of adequate contextual cues and controlled attention. Those contextual cues can be set by the task context and/or internally generated by the speaker and determine what information is relevant for the retrieval process and what must be displaced. Success in retrieval, then, as proposed by Unsworth and Engle (2007), depends on individuals' ability to use contextual cues effectively to delimit the search set. That is to say, the greater the number of items activated by the contextual cues and consequently included in the search set, the lower the probability that retrieval will occur fast and accurately.

Extending Unsworth and Engle's ideas to a bilingual context, retrieving L2 words from secondary memory is likely to function in basically the same way as in L1. However, an observation must be made. Because lexical retrieval procedures are not fully automatized in L2, any semantically-related cue presented close to the retrieval period is likely to help bilinguals to execute the serial search for the appropriate word, facilitating performance. In other words, semantically-related lexical items tend to belong to the same lexical semantic field and thus may serve as cues to delimit the search set adequately. Once the search set is efficiently delimited, sampling and retrieval become easier. Without such cues, more non-target items are possibly included in the search set and a more extensive search is needed. The extent of such a controlled search is likely to be related to the quantity and quality of L2 knowledge one possesses. Less proficient bilinguals, on the one hand, may have to perform the search more extensively, looking for items either in their less complete L2 mental lexicon or possibly in their L1 lexicon, which would probably be more time-consuming, increasing their reaction times and chances for error. The opposite is likely to be the case for more advanced bilinguals.

In Unsworth and Engle's proposal for retrieval from secondary memory, after the search set is delimited, a serial sample for the correct item is initiated. Once an item is sampled, a decision/monitoring process is responsible for checking whether the item is the target one to be retrieved. Again, in L2 lexical retrieval, a search for the correct item is even more likely to occur serially, since retrieval procedures operate under controlled processing. Another assumption regarding bilingual retrieval is that decision/monitoring processes may be a function of L2 proficiency. That is, because less proficient bilinguals tend to have a smaller repertoire of L2 lexical items stored in secondary memory and most of them might lack strong conceptual representations, it seems plausible to suggest that less proficient bilinguals will face greater difficulty to decide whether the selected item is indeed the more adequate one to be retrieved. More proficient bilinguals, on the other hand, know more L2 words for which conceptual connections are well established and therefore will probably monitor for mismatches more easily.

The studies reviewed in this section indicate a strong link between WM and retrieval. What seems to sustain this link are the processes assumed to be involved in determining individual differences in

WMC – active maintenance of relevant information and controlled serial search (and decision/monitoring process, in the case of L2 lexical retrieval), both made possible through the allocation of attention.

3. L2 lexical access: a controlled serial search task

The strategic controlled serial search view of L2 lexical access is supported by the results of Prebianca (2010). The piece of investigation reported in this section, conducted by Prebianca, was part of a bigger study carried out as a requirement for the attainment of a PhD degree. The main objective of the study was to investigate whether bilingual lexical access is predicted by working memory capacity (WMC) and proficiency level in L2. Deriving from that and based on the assumption that lexical access qualifies as a controlled serial search task (Engle and Oransky, 1999) subject to (i) working memory limitations, and (ii) the amount of automatized L2 knowledge one possesses (Kormos, 2006), three main hypotheses were addressed. The first one posed that WMC and L2 proficiency would both predict bilingual lexical access. The second hypothesis supposed that higher spans would retrieve lexical items faster than lower spans. The third one predicted that more proficient bilinguals would retrieve lexical items faster than less proficient bilinguals.

One hundred learners of English as a foreign language (L2) were submitted to three data collection sessions which comprised three tests to measure WMC (the L1 and L2 Speaking span test and the Operation span test), two tests to measure L2 proficiency (the Toefl iBT Speaking Test and a semantic categorization task) and one test to assess bilingual lexical access (an L2 picture-naming task).[5]

3.1 The WMC tests

The Brazilian Portuguese version of the Speaking Span Test (SST) administered in Prebianca's (2010) study was designed by Fortkamp (1999), based on Daneman's (1991) test, and was partially adapted by the present researcher so as to be more similar to the L2 version of the test with 3 test blocks rather than 2 as in the original. It consisted of 60 unrelated words presented in sets of 2, 3, 4, 5 and 6 words each. The words were 7 letters long and were displayed in the center of a computer screen for 1 second. After 10 milliseconds the next word of the set would appear. After all words of a specific set had been displayed, question marks on a black screen followed by a beep would signal it was the time for participants to start formulating the oral sentences for each word they had seen in that set. The number of question marks always referred to the number of words participants should recall and make a sentence with. Though there was no restriction in terms of complexity and length for the oral sentences, participants were informed that only semantically and syntactically accurate sentences in Brazilian Portuguese, produced for words in their exact form and order of presentation, would be accepted. Scoring procedures followed Daneman and Green (1986) and Daneman (1991), in which 1 point was awarded to every syntactically and semantically accurate L1 sentence generated for the words in the exact form and order of presentation thus yielding a strict score. The total SST score for each participant was calculated by summing up all points credited to the sentences correctly formulated. The measure yielded by this scoring procedure was named **SSTL1**.

The SST used to measure participants' working memory capacity in L2 in Prebianca's (2010) investigation was an L2 version of Daneman's (1991) original test and was designed by Weissheimer (2007). Like the L1 version, in the L2 SST, participants were required to memorize words in English for later recall and use them in the production of L2 (English) semantically and syntactically accurate oral sentences. There were 60 unrelated words displayed in sets of 2, 3, 4, 5, and 6 words each. Detailed instructions were given in the participants' native language together with three blocks of practice. Participants' individual span scores corresponded to the total number of words for which they were able to produce a grammatical and meaningful sentence in English, mirroring the criteria used to calculate the scores of the L1 SST. The measure of WMC resulting from this analysis was called **SSTL2**. This variable was then transformed into standardized

(z) scores to be inserted into the model for the multiple regressions run to investigate hypothesis 1 and was named **zSSTL2**.

The version of the Operation Span Test (OSpan) applied in the study differed from the original task designed by Turner and Engle (1989) in that the words to be recalled were in Portuguese instead of English. The L1 words used in the test were all dissyllabic words unlikely to be unknown by native speakers. The OSpan consisted of three test blocks of four sets each. Within each set, trials could vary from 2 to 5 in a pre-established order. For instance, block 1 was composed of 4 sets of 3, 5, 3 and 2 trials, respectively. In total, there were 42 trials – 19 displayed a mathematical operation string with a correct response and 23 displayed operations with an incorrect one.

In each trial, participants saw an operation-word string with a possible outcome followed by a word. The stimulus was displayed in the center of a computer screen. The math operations were a composite of multiplication or division problems followed by the subtraction or addition of an integer, for example, (9/1) -5, and were the same used in the original OSpan test (Turner and Engle, 1989). The outcome of the math operation should be verified YES or NO, depending on whether it was the correct result for the problem. For the above example, participants would see (9/1) -5 = 4 ?, and were expected to say whether 4 was the correct outcome of this operation. After reading and solving the operation, participants were required to read aloud the subsequent word for later recall, in this case, **balde**. Thus, the whole trial consisted in reading aloud the math operation, solving it as fast and as accurately as possible without pausing, and then immediately after verification, reading the word out aloud. As soon as participants completed each trial, the researcher pressed the space bar on the keyboard so that the next operation-word string appeared in the center of the screen. At the end of all trials of a set, question marks cued participants to recall the words they had read in that particular set in the exact order and form of presentation. The number of question marks corresponded to the number of the words they were supposed to recall.

According to Prebianca (2010) , the scoring procedure used to calculate individual OSpan scores was meant to be consistent with the procedure adopted for the speaking span tests, in which both processing and storage demands on WMC were given equal importance. Thus, following this criterion, 1 point was credited to each word recalled in the exact form and order of presentation, originating an **OSPan** index of WMC. This index, then, generated, together with SSTL1, a new variable calculated by converting OSPan and SSTL1 into standardized scores and averaging them. The resulting index was called **WMCL1z** and was inserted in the model for the multiple regressions run so as to investigate hypothesis 1.

3.2 The L2 proficiency tests

In order to assess learners' proficiency level in L2 speaking, a sample task of the TOEFL iBT Speaking Test was used. The task selected for eliciting participants' speech production was an independent task in which learners were asked to talk about a familiar topic – giving opinion about the best way to get the news. Following the original rubrics for the TOEFL iBT Speaking Test, three external raters rated participants' speech samples from 0 (no attempt from participants to respond to the question) to 4 (maximum score), taking into consideration speech delivery, language use and topic development. The resulting scores were labeled as **PROFToe**.

The Semantic Categorization task implemented in Prebianca's (2010) investigation was devised based on Dufour and Kroll (1995). In this task, participants were presented with names of L2 superordinate categories followed by L2 subordinate target nouns. Their task was to decide whether the subordinate nouns belonged to the superordinate categories. The stimuli for the task consisted of 50 English concrete nouns divided into 10 categories: clothing, color, occupation, fruit, transportation, drink, body part, vegetable, school object and animal. There were 6 target categories – animal, body part, fruit, transportation, vegetable and school object – and 4 filler categories – clothing, color, drink and occupation. All categories were randomly chosen and defined as targets or fillers.

Every single trial within the 5 experimental blocks was run as follows. First, a fixation point represented by the symbol (+) in Arial font, 30 point, appeared on the computer screen for 300 milliseconds (ms). Then, the superordinate category name replaced the fixation point for 400 ms. 450 ms after category onset, a subordinate noun appeared in the center of the screen for 450 ms. At this moment, participants should press 1 on the keyboard if the subordinate noun was a member of the category preousvily displayed and 2 if it was not. After an intertrial interval of 1.5 seconds, the next superordinate category name appeared on the screen automatically. At the end of each experimental block, participants were required to press the space bar to proceed to the next block. At the end of all blocks, a message indicated the experimental session was over. Individual scoring consisted of the total number of subordinate nouns correctly categorized for target superordinate categories only (**TOTCateg**). So as to allow for the investigation of differences between proficiency levels in L2 picture naming task, this variable was converted into standardized scores yielding a new index of L2 proficiency named **zTOTCateg**. Finally, a third proficiency index was obtained by calculating the average of z scores for zTOTCateg and zPROFToe (standardized scores for PROFToe), yielding a new variable referred to as **Meanz**.

3.3 The L2 lexical access test

In Prebianca's (2010) study, the picture-naming task was designed to assess participant's lexical access in L2 in terms of retrieval speed following most studies conducted under the picture-word interference paradigm (Roelofs, 1993; Damian and Martin, 1998; Costa et al., 1999). In this task, participants were required to name pictures in the presence of word distractors (the experimental condition). Pictures portrayed concrete objects visually displayed as black line drawings on a white computer screen and were to be named as fast and accurately as possible. Word distractors also referred to concrete objects and were of two different types: (1) semantically related and (2) phonologically related to the name of the picture. For instance, the picture of a DOG appeared with the word distractors cat and fog, respectively. Distractors were presented at three different points in time, following the Stimulus Onset Asynchrony Paradigm (SOA). Word distractors appeared together with the picture (SOA=0), 100 ms before picture onset (SOA=-100) and 100 ms after picture onset (SOA=+100).

All word distractors and names of the pictures were monosyllabic words in order to avoid an effect of articulation time on RT's. Semantically related distractors were words from the same category such as dog and cat. For phonologically related distractors, words that shared the greatest number of phonological segments with the name of the target picture were selected. Word distractors were presented in capital letters, Arial font, bold, 25 point. To avoid the matching of pictures and letters, all word distractors were displayed in blue font.

The task was divided into two different testing sessions – a control and an experimental session. A split half design was applied in this study regarding the picture-naming task in order to avoid practicing effect. That is, 50 participants were run in the control session first, followed by the experimental session, and 50 were run in the inverse order. Pictures were divided into 3 sets – a set of 25 target pictures displayed in the experimental and control conditions, a set of 30 filler pictures to complete the experimental condition, and a set of 20 training pictures to be presented in the training session. Fillers and training pictures were paired with unrelated word distractors presented at picture onset. The pictures and word distractors of the training session were not used in the main experiment. In total, participants saw 75 different pictures and produced 267 vocal responses.

The experimental session consisted of 6 blocks of 40 trials which, in turn, consisted of 25 target pictures plus 15 filler pictures, summing up a total of 240 responses per participant. Overall, the 25 target pictures produced 150 different combinations since each one was paired with two different types of distractors – semantically and phonologically related – and was presented in three different time conditions – -100 ms, 0ms and +100ms.

Every experimental trial had the following structure. First, a fixation point appeared in the center of the computer screen for 700 ms followed by a blank interval of 500 ms. Then, the picture was presented in the center of the screen. The picture and the word distractor remained on the screen until the participant produced a vocal response or a maximum of 1500 ms and then disappeared (Damian and Martin, 1998; Roelofs, 1993). There was an intertrial interval of 1.5 s. Then, the next picture appeared on the screen automatically (Costa et al., 1999). At the end of each block of the experimental condition, that is, after each 40 trials, participants were instructed in their L1 to press the space bar to proceed to the next block.

Although the picture-naming experiment was designed so as to provide data concerning the effects of different kinds of word distractors (namely, phonologically- and semantically-related) displayed at different moments in relation to picture onset (100 ms before, together with and 100 ms after picture onset), only the data regarding semantically-related word distractors presented 100 ms before picture onset were taken into consideration for data analysis.

Four lexical access measures were obtained through the picture-naming task: reaction time scores for the control and experimental conditions, and, naming accuracy scores for control and experimental conditions. Whereas in SOA= -100, reaction time measures in the experimental condition reflected the time participants took to name the picture from the offset of picture presentation on, in SOA=+100 they corresponded to the time participants took to name the picture from the onset of distractor presentation on. In addition, pictures were also named in a control condition, that is, without any word distractor presentation. This was done in order to generate a baseline measure to be compared with reaction time and naming accuracy measures produced by participants when naming pictures in the face of interfering stimuli. Naming accuracy (NA) was operationalized as the number of pictures participants were able to name correctly, regardless of how long they took to name them. This measure was useful in selecting which responses would be taken into consideration to calculate the mean reaction times for the control and experimental conditions. That is, only the RT's for pictures correctly named were included in the calculations for the mean; thus pictures named inaccurately or not named at all were excluded from the calculations – a procedure frequently adopted in lexical access studies. The mean RT's for the control and experimental conditions analyzed in Prebianca's (2010) study were labeled **RTctr** and **RTexp**, In order to conduct the statistical tests for analyzing the extent to which WMC and L2 proficiency predict bilingual lexical access in terms of retrieval speed, RTexp was converted into standardized scores so as to be in line with the standardized score for L2 proficiency - Meanz. The new index of RTexp was named **zRTexp**.

4. Results and discussion of Prebianca's (2010) study

This subsection attempts to report and explain the main results of the statistical analysis conducted in Prebianca's (2010) study regarding the relationship between WMC, bilingual lexical access and L2 proficiency. It addresses, more specifically, the aforementioned hypotheses.

Hypothesis 1 predicted that WMC and L2 proficiency would predict bilingual lexical access. Results of multiple linear regressions showed that WMC and L2 proficiency significantly predict bilingual lexical access. The main effects for WMC in L1 and L2 proficiency were significant, that is, as individual predictors, WMCL1z and Meanz both proved to contribute uniquely to zRTexp: t = -3,85, p = .000 and t = -5.01, p = .000, respectively, as depicted in Figure 1.

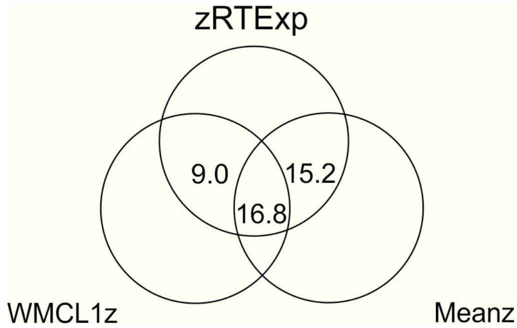

Figure 1. Percentage of unique and shared contributions of WMC in L1 and L2 proficiency to bilingual lexical access

The main effect for WMC in L2, on the other hand, was only significant in the absence of L2 proficiency. When WMC in L2 was inserted in the regression model as a separate set of predictors (apart from proficiency), its main effect proved to be statistically significant: $t = -5.02$, $p = .000$. Overall, these results support hypothesis 1.

In order to explain why WMC successfully predicts variation in the performance of a higher-order cognitive task such as bilingual lexical access, two important issues need to be taken into account: (1) the nature of WM processes, and (2) the nature of retrieval processes.

The basic view of WM taken in AUTHOR's (2010) study is that WM refers to a set of memory traces activated above threshold and temporarily maintained in a short-term buffer for further processing (Kane, Conway, Hambrick and Engle, 2007). Activation and maintenance of information are considered attention demanding tasks, especially when distraction drives attention away from the information being currently maintained. In this sense, attention is also needed to prevent irrelevant representations from entering the WM focus. Under this view, WMC reflects one's ability to (i) retrieve task relevant information from long-term memory when it has been already displaced or could not be kept in the attentional focus; (ii) keep it active and readily accessible, and (iii) inhibit distraction (Kane, Bleckley, Conway and Engle, 2001). Research has consistently shown that high and low span individuals are equally able to retrieve information from long-term memory in terms of speed and accuracy in the absence of interference (Conway and Engle, 1994). When interference is at play, however, only high span individuals can effectively block irrelevant stimuli. According to Kane et al. (2007), "the extent to which executive attention is engaged by a task, for maintenance, retrieval, or for blocking, is critically determined by the degree of interference or conflict presented by the context."(p. 22-23).

In the context of the complex span tasks used to measure WMC in Prebianca 's (2010) study, it is likely that the interference was caused by the intrinsic characteristics of the WMC tests. Remind that both the SST and the OSpan required participants to recall sets of an increasing number of unrelated words in serial order while shifting attention to process intermittently the L2 sentences or mathematical operations, leading to the building up of interference across test blocks and trials.

In other words, as the number of to-be-remembered items increased from block to block and began to accumulate across trials, access to relevant information became more difficult. Proactive interference resulted, in this case, from the competition between the number of words presented in previous blocks and the words that should be recalled in that particular block. Access to relevant information (to the to-be-remembered items, in the case of the span tests) is disrupted, as explained by Kane et al. (2007), because the processing task sentence formulation or solving the math operations in the case of the SST and OSpan, respectively – prevents the rehearsal of the to-be-remembered items, thus increasing the chances for proactive interference to grow. Controlled attention is then necessary to recover or keep access to the target items under proactive interference.

In the picture-naming task, different from what happened in the WMC tasks, interference does not seem to have originated from test stimuli specifically, but to have been caused by the association between the name of the target picture and other semantically-related items in the mental lexicon. For instance, when the picture of a dog is presented activating the word DOG in long-term memory, all other lexical items that share semantic constituents with the word DOG also become active thus interfering with retrieval[6] and possibly leading to cue overloading (Watkins, 1979 apud Unsworth and Engle, 2007a).

Cue overloading, according to Unsworth and Engle, occurs when several memory representations are subsumed to the same cue. If the cue to retrieval is overloaded, more items are selected from memory to be part of the search set and, as a result, retrieval will take longer and be more susceptible to errors. Coming back to the previous example, the word DOG and its semantically-related competitors would be subsumed to the same retrieval cue[7] – animals. Because there are several lexical candidates within the category animals, controlled attention is needed to execute a serial search and sample the most adequate one. Based on what has been said so far, it seems that one reason why measures of WMC and bilingual lexical access covary is due, at least in part, to the need to allocate controlled attention to block interference (proactive or retroactive), by keeping access to target items and retrieving task relevant information in the presence of activated competitors (Unsworth and Engle, 2007a).

Another possible explanation for why WMC significantly predicted variation in bilingual lexical access is that WM resources are usually required to impede automatic behavior when the context calls for a new response (Unsworth and Engle, 2007b), similar to what occurs in Stroop tasks. In the picture naming task conducted in Prebianca's (2010), for instance, attention may have been used to override automatic responses such as reading the word distractors presented prior to picture onset, instead of focusing on retrieving the name of the picture. Take the example of the picture of a dog. One hundred milliseconds before visualizing this picture, participants would see the word distractor **cat**. If one does not make use of his/her attentional resources to overcome this intrusion, it is likely that one will automatically read the word distractor instead of naming the picture of the dog displayed right after it. Moreover, the fact that task instructions emphasized the need to ignore the word distractors and name the pictures might also have contributed to the use of controlled processing to solve response competition. It is also important to highlight that the word distractors were semantically related to the names of the pictures and thus may have been even harder to suppress than unrelated word distractors which are not likely to belong to the same semantic field. Therefore, Prebianca suggested that, in order to perform the L2 picture-naming task properly, the bilingual speakers of the study needed to use controlled attention to maintain the task goal active in memory thus impeding irrelevant information to enter the focus and disrupt performance.

Still regarding the relationship between WMC and bilingual lexical access, another possible explanation for the pattern of results born out in AUTHOR's (2010) study might be related to the nature of the cognitive processes involved in determining individual differences in WMC and retrieval. In a recent model of WMC and retrieval, Unsworth and Engle (2007a, b) view WMC as the ability to maintain relevant information active in primary memory plus the ability to reactivate relevant information from secondary memory in situations where there is internal or external competition. According to the authors, the key to recovering relevant information from secondary memory lies on one's efficiency at delimiting the search set appropriately. In order to do that, one needs to attend to cues provided by the task and use them to restrict the number of possible target representations to search among. Once the search set is delimited adequately, a sampling process starts. In this process, controlled attention is needed to execute the serial search for the target representation. As soon as it is selected, monitoring is initiated. This process is then responsible for checking whether the selected representation is the correct one and can proceed to retrieval.

With that in mind, Prebianca (2010) suggested that what causes WMC and bilingual lexical access to be related is the fact that they share common processes such as cue-generation, set delimitation, sampling and monitoring, all of them being subserved by the allocation of attention. Thus, it is feasible to argue that bilingual lexical access qualifies as a controlled serial search that taps the same processes executed by WM in situations in which representations need to be kept active in the presence of interference.

The relationship between WMC and bilingual lexical access also seems to be mediated by L2 proficiency. As revealed by the results of the Prebianca 's (2010), WMC in L2 predicted bilingual lexical access significantly only when proficiency was excluded from the multiple regression model. However, L2 proficiency significantly predicted bilingual lexical access above and beyond WMC in L2 (about 18%): $t = -4.51$, $p = .000$.

This finding can be explained by the fact that, in order to perform the picture naming task, participants inevitably searched for L2 words they knew and were stored in their mental lexicon. In other words, they needed to know L2 names to be able to perform such a task. In sum, the relationship between L2 proficiency and bilingual lexical access lies on the L2 word knowledge one possesses.

This idea is supported by the Revised Hierarchical Model (RHM) of bilingual lexical representation proposed by Kroll and Stewart (1994), which assumes that semantic and conceptual mental representations evolve in a bilingual mind as a function of proficiency. That is, the model presumes that at the beginning of the learning process, when the L2 knowledge base is still incomplete and underdeveloped in relation to the L1, L2 words are learned and used through associative links with their L1 counterparts. This is so because L1 words are more strongly connected to their meanings, in the conceptual store, than L2 words. Connections between concepts and words in L2 are considered not to be fully established and, as a result, may lack some conceptual specifications (Poulisse, 1993). L1 connections, on the other hand, are stronger, well practiced and fully established in the lexicon. If this reasoning is correct, it seems safe to conclude that retrieval of L2 words for learners with less L2 practice takes place by means of an L1 conceptually-driven matching process. Put differently, because there is no lexical retrieval without conceptual activation, in order to select L2 words, one has first to access the L1 word and its concept. The L1 concept will then activate L1 and L2 words that match its semantic characteristics. Because the goal is to retrieve the word in L2, the lexical item that shares the L2 language cue and the greatest number of conceptual features with the L1 concept will then receive a boost of activation and be selected.

According to the RHM, with increased proficiency the initially weaker connections between L2 words and their concepts become stronger, allowing for a direct access to meaning. That is to say that L1 meanings are not needed for L2 retrieval anymore, which probably saves cognitive effort and processing time. Further attention will be given to this point later on in this discussion section.

With regard to individual differences in WMC and bilingual lexical access, it was hypothesized that higher spans would retrieve lexical items faster than lower spans (Hypothesis 2). Previous research relating WMC and L2 speech production has systematically shown that individuals with larger WMC produce speech which is more fluent, grammatically accurate and complex and less lexically dense (Fortkamp, 2000; Finardi and Prebianca, 2006; Weissheimer, 2007) than individuals with shorter WMC. The statistical analyses carried out in the study conducted by AUTHOR (2010) showed that, in fact, higher spans retrieved words in L2 faster than lower spans irrespective of proficiency level, as observed by their shorter naming response times, as depicted in Figures 2 and 3, thus supporting hypothesis 2.

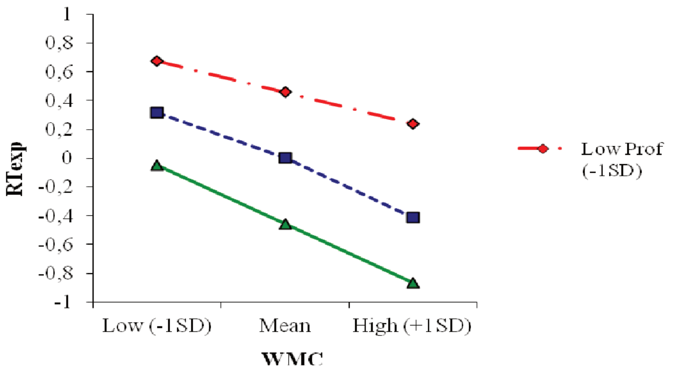

Figure 2. Less and more proficient bilinguals' and higher and lower spans' (WMC in L1) behavior in bilingual lexical access (mean standardized scores for RTexp)

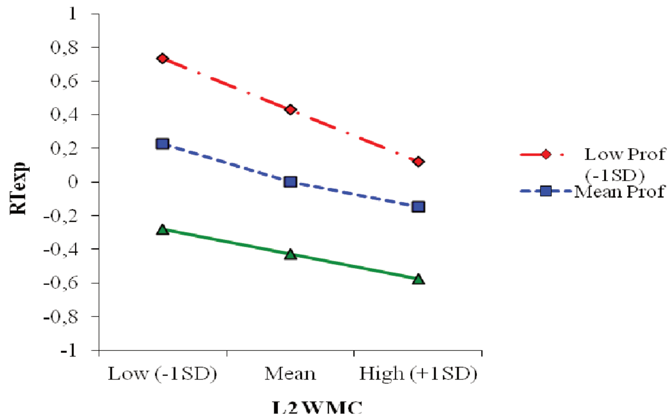

Figure 3. Less and more proficient bilinguals' and higher and lower spans' (WMC in L2) behavior in bilingual lexical access (mean standardized scores for RTexp)

The fact that higher spans outperformed lower spans in a bilingual lexical access task involving retrieval of L2 words can be explained by the interplay of processes which underlie both WMC and retrieval from secondary memory. As Engle (2001) noted, WMC is not simply about storage and processing, but rather about one´s ability to maintain pieces of information in an active memory in the presence of distraction. On this view, differences in WMC mean that misleading information is more efficiently kept outside the focus of attention and that controlled processes are more effectively used to resolve response competition.

The bilingual lexical access task applied in Prebianca's (2010) study was a task in which bilinguals were supposed to retrieve L2 names under the interference of semantically-related information. In order to accomplish this task fast and accurately, participants needed to (i) notice and/or generate cues to help delimiting the search set adequately; (ii) perform a strategic search in order to retrieve the most appropriate item to match the to-be-verbalized concept, and (iii) engage into a decision/monitoring process so as to ensure the selected L2 name was the one that should be retrieved. Because all these underlying retrieval processes were carried out in the face of competition, accessing words in L2 in the case of Prebianca 's (2010) study required a great amount of attention to be executed and thus, only higher spans, who are presumably more efficient at allocating attention to these sub-processes, were better able to perform the picture naming task. Put differently, it might be that lower spans did not have enough attentional capacity to devote to all sub-processes involved in retrieval, which may have hindered their performance. It is important to remember that the build-up of interference in the picture -naming task was not only triggered by the distractors themselves, but also by the activation of items in memory which were semantically related to the target one. This internal competition was probably better handled by higher spans due to their ability to maintain task relevant information active and block irrelevant competitors.

In fact, previous empirical research has demonstrated that high and low spans differ significantly in their ability to block proactive interference. In Rosen and Engle's (1997) study, for example, higher span subjects were better able than lower spans to exclude the names retrieved in a non-load condition from the set of names retrieved in a load condition. Lowers spans, on the other hand, included more repetitions when retrieving category exemplars in a load condition. These findings suggest that only higher spans had enough controlled attention to inhibit interference, monitor for repetitions and generate cues to retrieve new exemplars simultaneously.

Likewise, Rosen and Engle (1998) also showed that lower spans were unable to block intrusions from words they had previously associated with a particular item, when these words appeared again with a new item. Again, higher spans were faster and more accurate to recall words that were learned with a different pair-associate relative to lower spans.

The retrieval deficits demonstrated by lower spans in Prebianca 's study may be more specifically accounted for by their inability to generate cues that could lead to a search set composed of target items only. The bilingual lexical access task in this study required participants to name pictures in L2 by ignoring semantically-related items displayed 100 ms before picture onset (the experimental condition). Under these conditions, there were two possible ways to efficiently use cues to delimit the search set: (i) through a conceptually driven process based on the picture itself or (ii) by using the semantic-related item as a cue itself. Both alternatives would lead to the activation of a set of lexical items that are likely to belong to the same semantic field or, at least, to share some constituent parts. Either way, in order to selectively attend to this specific set of words, controlled attention was needed to inhibit active, but irrelevant representations (Unsworth and Engle, 2007a). If this was the case, the reason why higher spans searched from a more specific search set than lower spans, besides using controlled attention more efficiently, was because they were either better able to notice the cues given by the task context or to internally generate them based on task stimuli and requirements. As noted by Unsworth and Engle (2007b), lower spans usually make use of noisier cues to guide their search, thus including more representations in the search set. The greater the number of representations in the search set, the lower the probability that retrieval of relevant information will be successful.

The poorer performance of lower spans in L2 naming in relation to higher spans may also be related to differences in their ability to recover lexical items from memory and to monitor for errors. Contrary to Unsworth and Engle's assumption that "individuals differ only in the ability to use cues to delimit the search set and not in either the recovery process or the decision/monitoring process (p. 109), it might be that individual differences are likely to influence the recovery and monitoring processes in L2 due to lack of automatization of L2 retrieval procedures and weak connections between words and their conceptual representations.

Concerning hypothesis 3, Prebianca (2010) predicted that more proficient bilinguals would retrieve lexical items faster than less proficient bilinguals irrespective of performing the control or the experimental condition first. This prediction was confirmed by an analysis of variance which showed that the mean retrieval speed of more proficient bilinguals was statistically different from the mean retrieval speed of less proficient speakers for all three measures of proficiency investigated (namely, PROFToe, TotCateg and Meanz), irrespective of task order (performing the experimental or the control condition first). In fact, more proficient bilinguals were always faster than less proficient bilinguals, regardless of WMC, as can be seen in Figures 4, 5 and 6.

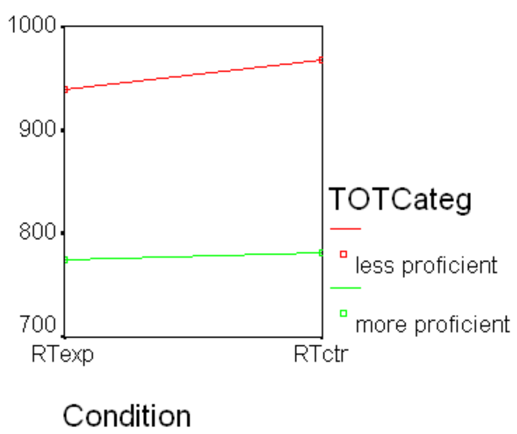

Figure 4. Less and more proficient bilinguals' behavior in RT-ctr and RTexp based on TOTCateg standardized scores

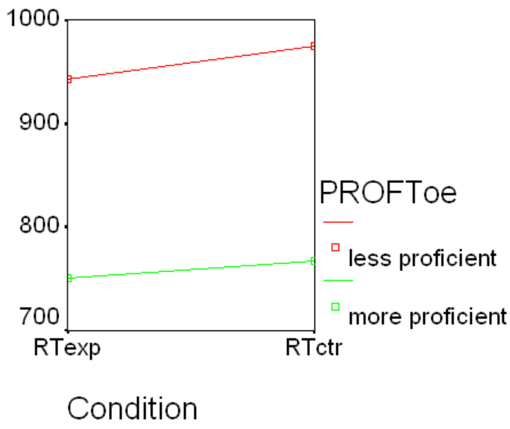

Figure 5. Less and more proficient bilinguals' behavior in RT-ctr and RTexp based on PROFToe standardized scores

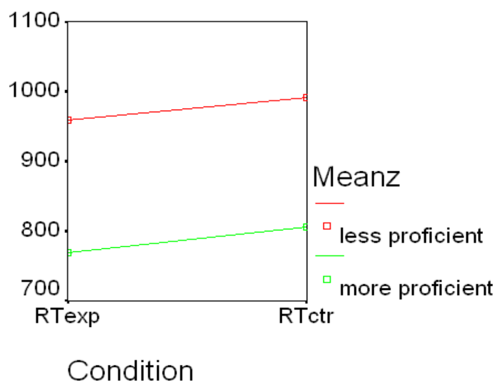

Figure 6. Less and more proficient bilinguals' behavior in RTctr and RTexp based on the mean standardized scores for TOTCateg and PROFToe

The fact that more proficient bilinguals were faster to retrieve L2 names relative to less proficient bilinguals may be a consequence, as already suggested by Prebianca (2007), of their more automatized L2 retrieval procedures. According to Kormos (2006), in the beginning of the learning process, several L2 lexical encoding procedures are not fully automatized yet and tend to be represented in a declarative, explicit fashion. This reasoning is in line with Kormos' proposal of a declarative store for L2 rules in her bilingual speech production model. If that is true, it is feasible to suggest that less proficient bilinguals were slower to retrieve L2 names because their procedures to do so were underdeveloped and/or incomplete. In other words, retrieval for them was based on knowledge of the L2 which was explicitly stored and processed serially instead of in parallel. As a consequence, learners took longer to execute the retrieval procedures involving delimiting the search set, sampling the right lexical item and checking for adequate selection. It is also worth mentioning that these lexical retrieval procedures in L1 are supposed to be part of the encoding system and stored implicitly.

Therefore, accessing words in L1 is a highly automatic process which runs in parallel to other sub-processes involved in the production of speech. As explained by Levelt (1989), the great speed with which speaking is produced in L1 can only be accounted for by what he calls incremental processing. That is, the components responsible for processing speech are made up of sub-components able to work in parallel at different stages from message generation to articulation as long as each component is instantiated with a fragment of its characteristic input.

In L2 speech production, because of the status of the L2 (being the less practiced language), processing is likely to be less incremental, which implies that the processing in one component may only start when the complete output of the previous component is delivered. Once again, the reason for that is the less automatized nature of L2 speech procedures - including lexical access ones -, and underdeveloped L2 knowledge, rendering the speech process a slow, attentional consuming serial task.

Another argument appealing to the idea that level of language automatization may affect lexical retrieval mechanisms is the one proposed by Roelofs (1998). According to him, the retrieval of L2 words is made possible through production rules that specify the concept to be verbalized and the language in which it is to be produced. An example of such rules would be: IF the concept is DOG and the language is Spanish, THEN select "perro". Productions of these kind are similar to the productions proposed by Anderson (1983) in his ACT* model of skill acquisition. Anderson advocates that a skill only becomes automatic when the procedures for its execution are created and retrieved from memory as a whole, without recourse to declarative knowledge. What determines the probability with which these productions will be retrieved over and over again so as to become automatic is a processing Anderson called strengthening, that is, practice. The more practiced a production rule is, the greater the likelihood it will be used again when the context calls for it.

On this view, the production rules proposed by Roelofs for L2 lexical retrieval can evolve from an explicit, declarative stage of representation to a stage where they are performed effortlessly, leading then to quantitative and qualitative changes in performance. In other words, as knowledge of the L2 develops, it is possible that not only the speed of processing is altered during lexical retrieval but also the way in which the underlying processes are organized and/or carried out by the bilingual speaker. As noted by Segalowitz and Hulstijn (2005), automatic processing should not reflect faster cognitive functioning only, but rather may encompass a set of modifications that can occur beneath the cognitive process surface. It might be, then, that the more proficient bilinguals of Prebianca 's (2010) study retrieved L2 words faster than the less proficient ones because they have been restructuring, reorganizing and re-elaborating the underlying processes involved in retrieval during their longer run in learning the L2 (see Cheng, 1985 for a similar view).

According to Kormos (2006), bilingual lexical access can be considered automatic when the to-be-verbalized concepts strongly activate their corresponding words. Under this reasoning, successful lexical access seems to depend only on the development of well established connections between the conceptual and the lexical store, as suggested by the RHM mentioned previously in this discussion section. However, Prebianca 's (2010) findings showing that bilingual lexical access is significantly predicted by WMC indicates that the strengthening of connections cannot be the only reason why the SST and the OSpan contribute to score variations in L2 picture-naming. As previously discussed in this paper, retrieval involves underlying processes that require controlled attention to be executed. Clearly, a well established network of concepts and lexical items seems to be of great help when one needs to select a word to match the conceptual specifications of the pre-verbal message (this point will be discussed in details later). However, there seems to be more than meets the eyes. If we consider that bilingual lexical access entails generating relevant cues for delimiting the search set appropriately, serial search and monitoring, it appears safe to suggest that each one of these underlying processes may be automatized to a different extent depending on the quantity and quality of L2 knowledge one possesses. Contrary to Kormos' claims, research on memory retrieval has indicated that retrieval from long-term memory is not likely to be carried out as a "direct, one-step" (p. 47) process.

Therefore, as suggest by Prebianca (2010), less proficient bilinguals may face greater difficulties in noticing the contextual cues needed to select the set of words from the most adequate category (semantic field) in the lexical network to start searching from, for instance, due to their poorer experience in using the L2. Besides, they may also execute the binding by checking process (as Levet et al. (1999) call the process in charge of checking for the match between concept and the lexical item selected for verbalization), which basically searches for the correct lexical item, in a more serial fashion, by looking for each primitive that makes up any possible lexical candidate and the overlap of these primitives with the primitives of the intended concept.

Monitoring for mismatches, in the case of less proficient bilinguals, also tends to be defective since their reduced L2 knowledge makes it more problematic for them to decide whether the selected item is the correct one. Support in favor of this argument comes from the view that L2 word knowledge evolves as a function of L2 proficiency, as discussed previously (Kroll and Stewart, 1994; de Groot and Hoeks, 1995). Thus, Prebianca (2010) reiterated that, in her study, more proficient learners accessed L2 lexical items faster than less proficient ones because they were able to perform the underlying processes involved in retrieval more efficiently, that is, more automatically.

Another interesting way to exemplify the changes that may occur in underlying retrieval processes as L2 proficiency increases is to look at how the connections between words and their meaning representations develop in a bilingual memory. The revised hierarchical model proposed by Kroll and Stewart (1994) postulates that the connections between L2 words and their meanings are established through associative links to L1 words. Access to meaning in initial L2 learning phases is then accomplished only by accessing L1 meaning first. In the same vein, de Groot and Hoeks (1995) claim that different lexical representations co-exist in a bilingual memory: word-association and concept-mediation representations – and develop as proficiency in L2 increases. That is to say, the lexical connections in the bilingual mental lexicon develop in a somewhat continuous fashion, from weak and indirect links to strong and direct links between words and their conceptual representations (meanings). A less proficient bilingual memory, in this case, would consist of two word stores (L1 and L2 lexicons) and a single common conceptual store with access to meaning occurring via L2/L1 associative links (the word-association hypothesis). Because the conceptual store would be shared between the two languages and because the L1 lexicon is likely to contain stronger, direct and automatic links with the conceptual store (Heredia, 1996), it is likely that in order to understand and produce L2 words, a less proficient bilingual needs to access L1 meaning first.

On the other hand, in a highly proficient bilingual memory, although the L1 and L2 mental lexicons also share the same conceptual store, access to meaning is not mediated by L1 lexical representations anymore. Instead, conceptual meaning is accessed via strong and direct connections between words and the conceptual store in each of the languages (the concept-mediation hypothesis). That is to say, comprehending and speaking in L2 for high proficiency bilinguals is likely to occur in a similar fashion as comprehending and producing speech in L1. According to de Groot (1995) and de Groot and Hoeks (1995), bilingual speakers would start accessing L2 meaning via L1 representations at the word level, but with practice they would develop stronger and direct connections between the L2 lexical and conceptual stores.

Because more proficient bilinguals have a greater amount of L2 practice than less proficient ones, it is likely that they have more L2 words represented, and consequently, processed in a concept-mediation mode. For less proficient bilinguals, who presumably have practiced the L2 for a lesser extent, L2 representation and processing probably relies more on a word-association mode, since the connections between words and their conceptual representations will still be weaker in relation to the same connections in a more proficient L2 memory. Consequently, it seems reasonable to argue that for word processing (accessing and retrieving) through a word-association mode, more controlled attention is necessary. Because the connections are weaker, a more serial search is required, and consequently biding by checking and monitoring processes will take longer to be performed. Together, these factors might have contributed to the inferior performance of less proficient bilinguals in L2 picture naming in Prebianca ‘s (2010) study.

4. Final remarks

The empirical work conducted by Prebianca (2010) and briefly reviewed in this article started from research on models of L2 word representation and processing, on working memory capacity and retrieval, and on L2 speech production studies to suggest that bilingual lexical access qualifies as a controlled serial search task which entails cue generation, set delimitation, sampling and monitoring, all being sub-served by controlled attention mechanisms. The findings of such investigation have shown that because attentional resources are limited and usually required to perform non-automatic retrieval, only higher spans and more proficient L2 learners were able to perform this task efficiently.

As for the pedagogical implications for the relationship among WMC, L2 lexical access and proficiency, it is important to mention that the most important conclusion one can draw from the present research is that learning a second language after some critical period (especially in adulthood) is quite a challenge. The literature on SLA and applied linguistics has consistently raised the point that L2 knowledge is usually less automatized than L1 knowledge with procedures operating under attentional control (McLaughlin, 1987; Poulisse, 1997; Fortkamp, 2000; Kormos, 2006). Likewise, L2 lexical items are in a smaller number and weakly established in the mental lexicon as compared to their L1 counterparts (Kroll and Stewart, 1994; Poulisse, 1997). Following from that is the fact that the development of automatization as well as strong lexical representations are intimately related to practice.

With regards to L2 practice, it is worth remembering that the role of oral production as a tool to promote L2 learning and language automatization has already been emphasized in the SLA literature by Swain (1985, 1995) and Skehan (1998). In her Output Hypothesis, Swain has proposed that language production affords opportunities for learners to engage into L2 syntactic processing especially when they are pushed to convey meaningful messages which are both grammatically and socially appropriate (Ellis, 2003). Skehan, building on Swain’s ideas, has suggested that production, besides fostering syntactic processing, allows learners to engage into hypothesis-testing of particular language forms as well as aids the automatization of L2 knowledge already learned. With that in mind, it seems that if the ultimate L2 teaching goal is to foster the automatization of L2 skills, teachers should concentrate on incorporating to the curriculum classroom activities with the potential to promote language production, knowledge restructuring and repetition.

Communication strategies (CS) can also be fruitful tools to be explored in order to teach learners how to overcome their lexical retrieval problems during communicative interactions. As proposed by Prebianca (2004, p. 108), CS instruction provides L2 learners with chances for

> (i) overcoming their possible communicative problems; (ii) optimizing communication; (iii) bridge the gap between what they know and what they want to say; (iv) developing their metalinguistic awareness, so that they can be able to decide on the best way to reach their communicative goals; (v) playing a more active role in their learning process , thereby expanding their communicative resources through hypothesis-formation processes; (vi) automatizing certain functions of the language such as expressing uncertainty, paraphrasing, asking for help, and using formulaic language.

On top of that, it is believed that certain CS can serve as tools to strengthen the connections among lexical items in the lexico-semantic network. That is, they have the potential to make learners aware of other linguistic forms to convey the same concepts, thereby enlarging the set of conceptual links of these concepts to other concepts in the network as well as to the lexical representations connected to them. Because the connections between concepts and lexical items are weak in the L2 mental lexicon and because some items are still underdeveloped in terms of meaning specifications, it is also important to add more lexico-semantic activities to the daily classroom practice such as the study of antonyms, synonyms, cognate words, homophones and hyperonyms in order to help learners to enlarge their lexical networks.

To conclude, the main objective of this article was to discuss the nature of bilingual lexical access mechanisms and how those mechanisms would operate and develop in the course of L2 learning with limited working memory capacity resources. In this sense, it is believed the article has contributed to refine our knowledge of the processes involved in L2 speech production and their specificities, particularly the features that render bilingual lexical access a controlled serial search task.

Notes

1. Following Schmidt (1992) and Skehan (1996), fluency is taken here as speakers' ability to mobilize their linguistic resources in order to produce speech in real time.
2. For the purposes of the present article, lexical access is the act of "retrieving a word […] from the mental lexicon, given a lexical concept to be expressed" (Levelt et al, 1999, p. 4). Throughout this article, the term lexical access will be used interchangeably with the terms *word selection*, *lexical retrieval*, and *lexical selection*.
3. Hereafter, the terms foreign and second language will be used interchangeably and will be referred to as L2. The term L2, in turn, is understood as a language one speaks other than his mother tongue (L1).
4. Throughout this article, the terms acquisition and learning will be taken as synonyms.
5. Due to space limitations, some details concerning the instruments of data collection needed to be omitted. However, comprehensive information about each test as well as on data collection, scoring procedures and statistical tests can be found in Prebianca (2010).
6. This kind of interference is known as retroactive interference. According to Searleman and Herrmann (1994), it "…occurs when newer information acts backward in time to inhibit recall of older information." (p.108). Proactive interference, on the contrary, "…occurs when previously learned information acts forward in time to inhibit recall of more recently learned material." (Searleman and Hermann, p. 108).
7. As will be further discussed in this subsection, cue generation is a crucial sub-process underlying retrieval from secondary memory. Unsworth and Engle (2006; 2007) have demonstrated that retrieval of information from secondary memory, that is, from information outside the focus of attention (WM) – stored in long-term memory – is governed by a discrimination process that involves the use of adequate contextual cues and controlled attention. Those contextual cues are set by the task context and determine what information is relevant for the retrieval process and what must be displaced. Success in retrieval, as proposed by Unsworth and Engle (2007), depends on individuals' ability to use contextual cues effectively to delimit the search set.

References

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

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory* (Vol 2), 89-195. New York: Academic 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. 47-89). New York: Academic Press.

Bialystok, E. (2005). Consequences of bilingualism for cognitive development. In: J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of Bilingualism: psycholinguistic approaches*. (pp. 417-432). New York: Oxford University Press.

Christoffels, I. K., de Groot, A. M. B., & Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: the role of expertise and language proficiency. *Journal of Memory and Language, 54(3)*, 324-345.

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123*, 354-373.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in: do words in bilingual's two lexicons compete for selection? *Journal of Memory and Language, 41*, 365-397.

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

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

Damian, M. F., & Martin, R. C. (1998). Is visual lexical access based on phonological codes? Evidence from a picture-word interference task. *Psychonomic Bulletin and Review, 5(1)*, 91-96.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Leaning and Verbal Behaviour, 19*, 450-466.

De Groot, A. M. B. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition 18*, 1001–18.

_______. (1995). Determinants of bilingual lexicosemantic organization. *Computer Assisted Language Learning, 8(2-3)*, 151-180.

De Groot, A. M. B., & Hoeks, J. C. J. (1995). The development of bilingual memory: evidence for word translation by trilinguals. *Language Learning, 45(4)*, 683-724.

Dufour, R., & Kroll, J. F. (1995). Matching words to concepts in two languages: A test of the concept mediation model of bilingual representation. *Memory & Cognition, 23*, 166-180.

Ellis, R. (2003). *Task-based language teaching and learning.* Oxford: Oxford University Press.

Engle, R. W. (2001). What is working-memory capacity?. In H. L. Roediger III & J. S. Nairne (Eds.), *The Nature of Remembering: Essays in Honor of Robert G. Crowder* (pp. 297-314). Washington, DC: American Psychological Association.

Engle, R. W., & Oransky, N. (1999). The evolution from short-term to working memory: Multi-store to dynamic models of temporary storage. In R. Sternberg (Ed.), *The Nature of Cognition* (pp. 514-555). Cambridge, MA: MIT Press.

Finardi, K.; & Prebianca, G. V. V. (2006). Working memory capacity and speech production in L2: evidence from a picture description task. *Revista de Estudos da Linguagem, 14 (1),* 231- 260.

Fortkamp, M. B. M. (2000). *Working Memory Capacity and L2 speech production: an exploratory study.* Unpublished doctoral dissertation. Florianópolis: Pós-Graduação em Inglês e Literatura Correspondente, UFSC.

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

Heredia, R. R. (1996). Bilingual Memory: a re-revised version of the hierarchical model of bilingual memory. *The Newsletter of the Center for Research in Language, 10*, 3-6.

Kane, M. J., Bleckley, K. M., 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., Conway, A. R. A., Hambrick, D. Z., Engle, R. W. (2007). Variation in Working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M .J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in Working Memory* (pp. 21 - 48). New York: Oxford University Press.

Kormos, J. (2006). *Speech production and second language acquisition*. Hillsdale, N.J.: Erlbaum.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149-174.

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research, 18*, 137-171.

Levelt, W. J. M. (1989). *Speaking: from intention to articulation*. *The Speaker as Information Processor*. Cambridge, MA: MIT Press.

Levelt, W. J. M., Roelofs, A., & Meyer, A. (1999) A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.

McLaughlin, B. (1987). *Theories of second-language learning*. London, England: Edward Arnold Publishers Ltd.

Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: learning, memory and cognition, 28 (3)*, 555-563.

Poulisse, N. (1993) A Theoretical Account of Lexical Communication Strategies. In R. Schreuder & B. Weltens (Eds.), *The Bilingual lexicon* (pp. 157-190). Amsterdam: John Benjamins.

Poulisse, N. (1997). Language Production in Bilinguals. In de Groot, A. M. B. & Kroll, J. F. (Eds.), *Tutorials in bilingualism: psycholinguistic perspectives* (pp.201-224). Mahwah, NJ: Lawrence Erlbaum.

Prebianca, G. V. V. (2004). *Communication Strategies and L2 Speech Production*. Unpublished Master's Thesis. Florianópolis: UFSC.

Prebianca, G. V. V. (2007). Working memory capacity, lexical access and proficiency level in L2 speech production: an exploratory study. Unpublished research paper. Universidade Federal de Santa Catarina, Florianópolis.

Prebianca, G. V. V. (2010). *Working Memory Capacity and Foreign Language Speech Production:* a look at lexical access processes and level of proficiency. 1. ed. Saarbrücken: LAP Lambert Academic Publishing AG & Co. KG, v. 1.

Roelofs, A. (1998). Lemma selection without inhibition of languages in bilingual speakers. *Bilingualism: Language and Cognition, 1*, 94-95.

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

______. (1998). Working memory capacity and suppression. *Journal of Memory and Language, 39*, 418-436.

Searleman, A., & Hermann, D. (1994). *Memory from a broader perspective*. Singapore: McGraw-Hill, Inc.

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In: J. F. Kroll & A.M.B. de Groot (Eds.), *Handbook of Bilingualism: psycholinguistic approaches*. New York: Oxford University Press, 371-388.

Schmidt, R. (1992). *Psychological mechanisms underlying second language fluency.* Cambridge University Press, 357-385.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing. II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127-190.

Skehan, P. (1996). A Framework of the implementation of Task-based instruction. *Applied Linguistics, 17 (1)*, 38-62.

______. (1998). *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in Second Language Acquisition* (pp. 235-256). New York: Newbury House.

______. (1995). Three functions of output in second language learning. In G. C. B. Seidlhofer (Ed.), *Principle and Practice in Applied Linguistics: Studies in honor of H.G. Widdowson* (pp. 125-144). Oxford: Oxford University Press.

Tokowicz, N., Michael, E. B., & Kroll, J. F. (2004). The roles of study-abroad experience and working-memory capacity in the types of errors made during translation. *Bilingualism: Language and Cognition, 7 (3)*, 255–272.

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

Unsworth, N., & Engle, R. W. (2007a). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*, 104 - 132.

______. (2007b). Individual differences in working memory capacity and retrieval: A Cue-dependent search approach. In J. S. Nairne (Ed.), *The Foundations of Remembering: Essays in Honor of Henry L. Roediger, III* (pp. 241 -258). NY: Psychology Press.

______. (2006). A temporal-contextual retrieval account of complex span: An analysis of errors. *Journal of Memory and Language, 54,* 346 – 362.

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

Weissheimer, J. (2007). *Working memory capacity and the development of L2 speech production: an exploratory study.* Tese de doutorado. Florianópolis: Pós-Graduação em Inglês e Literatura Correspondente, UFSC.

Recebido em: 16/02/2017
Aceito em: 24/04/2017