WORKING MEMORY CAPACITY AND THE IMPLEMENTATION OF PLANNED IDEAS INTO L2 SPEECH PERFORMANCE OF TRANSLATION TASKS

Maria da Glória Guará-Tavares¹
Sâmia Alves Carvalho²

ABSTRACT

The present study is part of a larger scale research (Guará-Tavares, 2011, 2013, 2016) that investigates the relationship among working memory capacity, pre-task planning and L2 speech performance. The aim of the study was to analyze whether higher working memory capacity individuals are better able to implement planned information into L2 oral performance of translation tasks. Learners’ planned ideas were accessed by means of think aloud protocols. Working memory capacity was measured by the Speaking Span Test. Results indicate that there are no significant differences between higher and lower spans concerning retrieval of planned lexical items that were implemented into task performance. As for the percentage of clauses per c-unit retrieved, the differences between higher and lower spans only approached significance. Results are discussed in terms of (Engle’s model of working memory (Engle, 1996; Engle & Oransky, 1999; Engle, Kane & Tuholsky, 1999).

Key words: Working memory. Translation tasks. Retrieval. Performance

¹ Professora Asssociada I da Universidade Federal do Ceará, membro permanente da Pós-Graduação em Estudos da Tradução – POET-UFC.
² Professora da Casa de Cultura Britânica da Universidade Federal do Ceará, membro permanente da Pós-Graduação em Estudos da Tradução – POET-UFC.
RESUMO

O presente estudo faz parte de uma pesquisa maior (Guará-Tavares, 2011, 2013, 2016) que investiga a relação entre capacidade de memória de trabalho, planejamento pré-tarefa e desempenho oral em L2. O objetivo do estudo foi analisar se os indivíduos com maior capacidade de memória de trabalho são mais capazes de implementar informações planejadas no desempenho oral de tarefas de tradução. As idéias planejadas dos participantes foram acessadas por meio de protocolos verbais. A capacidade de memória de trabalho foi medida pelo teste Speaking Span. Os resultados indicam que não existem diferenças significativas entre indivíduos com maior e menor capacidade de memória de trabalho no que diz respeito ao resgate de itens lexicais planejados; e no que concerne ao resgate de sentenças por período, as diferenças apenas se aproximaram da significância. Os resultados são discutidos com base no modelo de memória de trabalho de Engle (Engle, 1996; Engle & Oransky, 1999; Engle, Kane & Tuholsky, 1999).


1 INTRODUCTION

Several studies have examined the impact of pre-task planning on L2 performance (Ellis, 1987; Foster & Skehan, 1996; Menhert, 1998; Sangaran, 2005; Ortega, 1999, 2005; Guará-Tavares, 2011, 2013, 2016; Yuan & Ellis, 2003; Abdi Tabari, 2016, 2017). In general, studies have shown a positive impact of planning on L2 performance, with several studies demonstrating that planning leads to gains in fluency3 (Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999). Planning also leads to gains in accuracy, although results have been more mixed and inconsistent in this respect (Ellis, 1987; Wendel, 1997; Mehnert, 1998; Ortega, 1999; Foster & Skehan, 1999). Finally, studies have also shown that planning enhances complexity (Crookes, 1989; Foster & Skehan, 1996; Mehnert, 1998; Ortega, 1999; Yuan & Ellis, 2003).

One interesting finding of the research on the impact of planning on L2 performance is the evidence of attentional trade-off effects among the goals of fluency, accuracy, and complexity. Foster and Skehan (1996), Menhert (1998), as well as Yuan and Ellis (2003) discuss results of their studies in terms of an attentional model of learning and performance. In this sense, these researchers propose that there are trade-off effects among the goals of fluency, accuracy, and complexity in the context of the use of learners’ limited capacity attentional resources. In other words, because attentional resources are limited, planning benefits cannot be achieved to the same extent simultaneously for fluency, accuracy, and complexity of L2 performance. The trend of research results shows that there are gains in fluency and complexity at the expense of gains in accuracy.

We take the perspective that working memory resources are attentional. In the present study, working memory is defined as “a system consisting of those long-term memory traces above a

---

3 According to Skehan (1996, 1998), fluency is related to the temporal aspects of speech production; accuracy is related to grammatical correctness; complexity is related to language elaboration (e.g., subordination).
threshold, the procedures and skills to achieve and maintain that activation, and limited-capacity, controlled attention” (Engle, Kane, & Tuholski, 1999, p. 102). Despite the fact that researchers in task-based planning (e.g. Foster & Skehan, 1996; Menhert, 1998; Yuan, Ellis & 2003) explain results of studies in terms of learners’ limited capacity attentional resources, individual differences in working memory capacity have not been taken into account in any of these studies (e.g. Foster & Skehan, 1996; Menhert, 1998; Yuan & Ellis, 2003) as a feasible variable for affecting learners’ performance under planning conditions.

Guará-Tavares (2011, 2013) provided evidence that higher working memory spans outperform lower working memory spans in L2 oral performance after pre-task planning, and Guará-Tavares (2016) found that higher working memory spans outperform lower working memory spans in the use of metacognitive strategies during pre-task planning. Planning is a problem solving activity (D’Ely, 2006), and it seems to assist performance by triggering a range of strategic, metalinguistic, and metacognitive behaviors (Ortega, 2005). One’s ability to engage in such strategic behaviors may successfully, to some extent, explain benefits achieved from planning. Since individuals with higher capacity tend to be more strategic (McNamara & Scott, 2001), individual differences in working memory capacity seems to reflect differences on how successful one is in the process of planning (Guará-Tavares, 2016).

In addition to that, the benefits of planning on performance may also depend on the ability to actually retrieve what was planned and implement it into online performance (Ortega, 2005). According to Rosen and Engle (1997), working memory plays a crucial role in retrieval, that is to say, individuals with higher capacity tend to retrieve information more effectively during the performance of complex cognitive tasks. The objective of this study is to investigate whether higher and lower working memory spans differ in the amount of planned ideas that are retrieved and implemented into task performance.

2 REVIEW OF THE LITERATURE

2.1 WORKING MEMORY

Engle et al. (1999) view working memory as a cognitive system comprising (1) a store in the form of long-term memory traces active above a threshold; (2) processes for achieving and maintaining this activation and (3) controlled attention. Nevertheless, when they refer to working memory capacity, it is the limited capacity of the element of controlled attention that is being referred to.

More specifically, for Engle and his associates the term working memory capacity refers to “attentional processes that maintain task-relevant information activated in an accessible state, or to retrieve that information under conditions of interference, conflict, and competition” (Kane, Conway, Hambrick & Engle, 2003, p. 23). When referring to working memory capacity, Engle and his associates mean the limited capacity of the mechanism that Baddeley and Hitch (1974), Baddeley and Logie (1999) have called central executive, which is, in turn, similar to the supervisory attentional system proposed by Norman and Shallice (1986, as cited in Baddeley, 1990; Engle et al. 1999).

According to Engle et al. (1999, p. 104), “working memory is not about storage or memory per se, but about the capacity for controlled sustained attention in the face of interference” (em-
phasis in the original). They view the nature of working memory limitations in terms of the capacity for controlled attention, which will emerge in situations that require controlled processing. In a controlled processing activity, it is attention that is controlled and the cognitive mechanisms that encompass a controlled processing activity include activation, suppression, serial search and retrieval, and monitoring (Engle, 1996; Engle & Oransky, 1999; Engle, Kane, & Tuholsky, 1999).

Rosen and Engle (1997) found that working memory plays a role in retrieval. Four experiments explored the role of individual differences in working memory capacity on verbal fluency under various secondary load conditions. High working memory individuals consistently recalled more exemplars. Unsworth, Billar and Spillers (2013) examined the role of working memory in retrieval of information. Participants performed a prolonged category fluency task that required them to retrieve as many animals as possible during 5 min. The results suggested that working memory capacity (WMC, from now on) differences emerged in the numbers of animals retrieved. Moreover, an analysis of differences in retrieval strategies suggested that higher working memory spans were more strategic than lower working memory spans and that these differences in retrieval strategies accounted for the overall differences in the numbers of animals retrieved. Results suggest that low-WMC individuals are less able than high-WMC individuals to select and utilize appropriate retrieval strategies to self-generate cues to access information in long-term memory. These results are consistent with research suggesting that working memory capacity is important for controlled search from long-term memory.

Guará-Tavares (2016) found that learners mainly engage in organization of ideas, rehearsal, lexical searches and monitoring when they plan an oral task. Moreover, higher spans employ significantly more metacognitive strategies during planning when compared to lower spans. Based on the findings of Rosen and Engle (1997), Unsworth, Billar and Spillers (2013) and Guará-Tavares (2016), this study sets out to examine whether higher and lower working memory spans differ in the amount of information they are able to retrieve from pre-task planning and implement into performance of a translation task.

2.2 MULTIMODAL TEXTS AND TRANSLATION TASKS

In our postmodern society where the dizzying advances of technology in all fields have tremendously increased the amount of data and information in our hands, students very often come across texts composed not only by words but also by images, that is, texts produced through different semiotic modes4. Therefore, we have chosen to use a broader notion of text in order to accommodate various forms of expression. Text is defined here as Halliday (1985, p. 52) stated, “in the simplest way…as language that is functional, by functional we simply mean language that is doing some job in some context”. In addition, by language we mean a system of signs in which the verbal sign is only one among many others like images, sounds, gestures etc.

According to Jewitt (2011), communication has always happened through more than one mode of expression. However, this perception of the existence of multimodal communicative compositions has become more intense with the development of visual representation technologies and with the expansion of the various forms that people use in communicating and interacting (synchronously and asynchronously) and thereby opening space for the insertion of other modes, different from writing, and nowadays very recurrent in people’s daily lives.

---

4 “Mode is a socially shaped and culturally given resource for making meaning. Image, writing, layout, music, gesture, speech, moving image, soundtrack are examples of modes (emphasis in the original)” (Kress, 2011, p. 54).
Lemke (1998a, p. 283) states that

[...] today our technologies are moving us from the age of ‘writing’ to an age of ‘multimedia authoring’ in which voice-annotated documents and images, and written text itself, are now merely components of larger meaning-objects.

In order to be coherent with these kinds of text, these larger meaning objects students deal with in their everyday lives; we have decided to use an image-text relation in our experiment. We used a narrative text made of images, which students had to read and translate into an oral story. That is, they had to perform an intersemiotic translation task, grasping meaning from one mode and translating into another.

Even though performing an intersemiotic translation task may seem difficult for some people, it is actually a natural process. The very perception of the world around us can be considered a translation process in the perspective that what is perceived through our senses is translated into signs and these signs structure other essential processes, such as communication and thinking. According to Jakobson (1959), translation may occur in three different levels: 1) Intralingual translation or rewording which is an interpretation of verbal signs by means of other signs of the same language; 2) Interlingual translation or translation proper, which is an interpretation of verbal signs by means of some other language; 3) Intersemiotic translation or transmutation, which is an interpretation of verbal signs by means of signs of nonverbal sign systems. The latter was the one our participants used to perform the task that were assigned.

We see texts made of images every day and we very often translate them into verbal texts. When we see an accident on the streets and we tell friends what happened, for example, we are performing an intersemiotic translations task. We get involved in a meaning-making process and the meanings we make are produced socially in the sense that they are governed by certain rules in force at the time of their production, by a specific context, and they are also influenced by the interests of those who produce them. We read meanings from one semiotic mode (moving images) and translate them into another (an oral text) more often than we actually notice. Reading is considered here as in Walsh’s (2009, p. 2) perspective, it “may involve viewing, listening and responding”. Most of the time the process occurs naturally, even though we are not usually aware of this as an intersemiotic translation task; therefore, we do not consciously plan it.

When this process is conscious, before approaching a text the translator makes a series of decisions taking into account the context of the source text, the context of the target text and the purpose and the cultural aspects of target audience.

Our aim is to provide the participants with the opportunity of manipulating an image-text, similar to the ones they see every day; use their reading images skills and perform one kind of translation that happens naturally in our daily lives. With this research set established, we aim to analyze whether higher-WMC individuals are better able to implement planned information into L2 oral performance of translation tasks.
3 METHOD

3.1 RESEARCH QUESTION

The study was motivated by one research question: Do higher and lower working memory spans significantly differ in the amount of information they are able to retrieve from pre-task planning and implement into performance of a translation task?

3.2 CONTEXT AND PARTICIPANTS

Participants of the present study were twenty-five intermediate learners from the Letras\(^5\) Licenciatura, Letras Secretariado\(^6\), and also from the Extracurricular Language Courses at the Federal University of Santa Catarina (UFSC). Participants of the Extracurricular Language courses were all undergraduate students at the Federal University of Santa Catarina from a variety of backgrounds (Biology, Engineering, Law, and History, among others). 15 were female and 10 were males. Their ages ranged between 18 and 29 years old, thus an adult population.

3.3 DATA COLLECTION PROCEDURES

Data collection of the present study was divided into three phases. The first phase was the selection of participants which aimed at controlling for proficiency level. Participants performed the proficiency trial task at the language laboratory, and all students of the same class did the task together (See Guará-Tavares, 2011, for a detailed description of the selection of participants). The translation task of the selection of participants was carried out under no planning conditions.

The second phase consisted of the Speaking Span Test to measure participants’ working memory capacity. Participants carried out the speaking span test individually with the researcher in a computer lab. A training session on how to take the test took place before test performance itself.

The third phase of data collection consisted of the second translation task. Participants carried out the second task under a planning condition, they had 10 minutes to plan the oral task prior to actual performance. During planning, verbal protocols were carried out. When planning time was over, participants carried out the task, then, a retrospective interview was also conducted.

3.4 THE SPEAKING SPAN TEST

A version of Daneman and Green (1986) Speaking Span Test was used to measure individuals’ working memory capacity. A training phase (20 words) preceded the testing phase (60 words). The test contained 60 unrelated words organized in three sets each of two, three, four, five and six words.

Each word was presented individually, on the middle line of a computer screen for one second. Participants were instructed to read each word aloud. At the end of each set, question marks appeared. These marks signaled the number of words that had to be stored and the number of

\(^5\) Undergraduate Language Teaching program.
\(^6\) Undergraduate Bilingual Secretary program.
sentences to be produced. Participants were instructed to use the words in the exact form and order they appeared to generate syntactically and semantically acceptable sentences, aloud, in English.

There were no restrictions concerning the length or complexity of the sentences produced. For instance, after being presented a set of three words: guy - point - train, a participant produced the following sentences: “I am a guy”, “what’s your point?”; “The train was dirty”. Participants’ speaking span score was defined as the maximum number of words for which they could generate grammatically and semantically acceptable sentences in English.

Following Daneman (1991), in this study, participants’ responses, which were recorded, transcribed and analyzed, generated two different speaking span scores: a speaking span strict score, when all the sentences the subject produced contained the target word in the exact form and order of presentation; and a speaking span lenient score, when credit was given for sentences that contained the target word in a form other than that of presentation (e.g., target word being ‘guy’ and the word in the sentence produced being ‘guys’), and half credit was given to words recalled in a different order. No credit was given to ungrammatical sentences in terms of syntax and semantics.

3.5 THE SPEECH GENERATION TASK

As we discussed previously in this article, we chose an intermodal translation task because the aim is that participants deal with the kind of text they often come across nowadays. It was also intended to provide participants with one kind of task that we naturally do but do not plan, then give them time and opportunity to plan, so the following procedures were set.

Participants were instructed to look at the set of pictures for 50 seconds. Then, the picture was removed from them and they had 10 minutes to plan their oral performance. These procedures followed Mehnert (1998). Planning was not guided.

3.6 THE VERBAL REPORTS: RETROSPECTIVE ON-LINE PROTOCOLS AND RETROSPECTIVE INTERVIEWS

Following retrospective on-line procedures (Leow; Morgan-Short, 2004), participants were given 10 minutes to plan and were required to verbalize what they were planning in breaks of every one minute. After every one minute of planning they were prompted with the question: ‘What were you just thinking about?’ However, in some moments in which participants stopped taking notes and seemed to be thinking hard or when participants erased part of their notes, the present researcher asked different questions: ‘What were you just thinking when you stopped writing?’ or ‘What did you just erase from your notes?’ The ten minute planning time was counted with the aid of a chronometer which was stopped during the verbalization so that participants could have ten minutes of actual planning. Also, the instances of verbalization were made the shortest possible so that they would not take participants away from the planning task itself. Basically, I asked the question and accepted whatever answer they gave me and instructed them to go on planning. In general, each verbalization was no longer than 30 seconds (including my question and the answer).

After the retrospective on-line protocols, participants performed the task. After performance of the task, an interview was also carried out for the purpose of complementing the retrospective on-line protocols. This combination of protocols is suggested by Wigglesworth (2005) and Leow and Morgan-Short (2004).
3.7 DATA ANALYSIS

The analysis of the protocols consisted of three phases. Two interraters helped in the analysis of protocols. Initially, a content analysis of the protocols was carried out individually by the present researcher and the first interrater. This content analysis consisted of going through the protocols and writing down our first general impressions on them.

The second phase consisted of identifying the information present in the protocols that was also present in participants’ performance. This analysis was done one by one; each participant had the planning protocol compared to the performance after planning. Two categories of information were included: a) number of isolated lexical items (content words- nouns, verbs, adjectives and adverbs) per a hundred words and b) percentage of clauses per c-unit. When learners mentioned the same lexical items and/or clauses in both protocol and performance, they were included in the counting. When lexical items and clauses were present in participants’ protocol and performance, these lexical items and clauses were considered as planned ideas/information retrieved from pre-task planning and implemented into task performance.

After counting all the lexical items present in each participant’s protocol and performance (retrieved lexical items), the number of retrieved lexical items per a hundred words was calculated by dividing participants’ total number of retrieved lexical items by the total number of words produced in performance and multiplying the result by 100. After counting all the clauses that were present in each participant’s protocol and performance (retrieved clauses), the percentage of retrieved clauses per c-unit was calculated by dividing the total number of retrieved clauses by the total number of clauses produced in performance, and the resulting figure was multiplied by 100.

After calculating the number of lexical items retrieved per a hundred words and the percentage of clauses retrieved per c-unit, descriptive statistics analysis was carried out. Then, an ANOVA was performed to verify whether there were any statistically significant differences between lower and higher spans concerning the number of lexical items retrieved per a hundred words and the percentage of clauses retrieved per c-unit.

In order to scrutinize the differences between lower and higher spans, we also followed the extreme-group design procedures and excluded intermediate spans from the analysis. For all statistical analyses of the present study, a probability level of \( p < .05 \) was used to determine statistical significance.

4 RESULTS

In this section, the results of the analysis of the protocols are reported. Table 1 displays the results of the ANOVA.

---

7 When analyzing c-units in the present study, I followed Foster et al. (2000) criteria. Utterances that were abandoned were not counted as a unit; phrases or full clauses that were repeated verbatim were counted once, with only one instance being considered as either a c-unit or belonging to a c-unit; verbatim repetition of words including those used for rhetorical purposes were considered as parts of the c-unit they belonged to; and whenever self-corrections took place, only the final version was counted as belonging to the c-unit.
As can be seen in Table 1, there were no significant differences between lower and higher working memory spans concerning the percentage of retrieved clauses that were implemented into task performance ($f = 2.663$, $p = .121$). Moreover, the difference between lower and higher spans concerning the retrieved number of lexical items per a hundred words only approached significance ($f = 3.903$, $p = .064$) but did not reach significance.

### 5 DISCUSSION

The study was guided by one research question: Do higher and lower working memory spans significantly differ in the amount of information they are able to retrieve from pre-task planning and implement into task performance? Retrieval of information from pre-task planning and implemented into task performance was measured by the number of retrieved lexical items per a hundred words and percentage of retrieved clauses per c-unit. There were no significant differences between higher and lower working memory spans in any of the two measures.

A tentative explanation for these unexpected results may be that higher and lower spans not only perform the task by retrieval of planned information but also by creation of new ideas online during performance. Guará-Tavares (2011, 2013) showed that higher working memory spans outperform lower working memory spans in task performance after planning. Guará-Tavares (2016) provided evidence that higher working memory spans use pre-task planning time more effectively than lower working memory spans by using a significantly greater amount of metacognitive strategies during planning. However, participants in Guará-Tavares (2016) and in the current study also reported that they used planned information but also created new ideas online during performance. Ortega (1999, 2005) claims that successful performance after pre-task planning depends on being able to retrieve what was planned into performance and create of new ideas online.

Since the present study showed no significant differences in terms of retrieval, possibly the key difference between higher and lower spans performance after pre-task planning revealed in previous studies (Guará-Tavares, 2011, 2013) may not be the amount of retrieved planned information into performance, but the ability to coordinate the combination of planned and new information during performance. Individual differences in working memory capacity reflect differences in the capacity for sustaining, maintaining and shifting attention among the various aspects of task...
performance (Engle, 1996; Engle & Oransky, 1999; Engle, Kane, & Tuholsky, 1999). Possibly, higher spans were able to retrieve more information but decided not to use all of it because they had more ideas online during performance and aimed at combining planned and new ideas in an effective way.

Another plausible explanation for the lack of significant differences between higher and lower spans in terms of retrieval may be that although verbal protocols were effective to establish the strategies employed during planning (Gará-Tavares, 2016), they may not be a suitable tool to assess the actual language being planned. Providing learners with instructions to write a draft of their stories during planning may be more effective. The language present in the drafts of the stories could be more thoroughly compared to language used in task performance.

The present study has its limitations. The sample size was small and only one test was used to measure working memory capacity. Despite its limitations, the study is relevant since it is a first step towards scrutinizing the role of working memory in retrieval of planned information that is implemented into performance of a translation task.

REFERENCES


