

# Weighing up the Effects of Working Memory and Cognitive Abilities in CLIL Learning

Laura Sánchez<sup>1</sup>

<sup>1</sup> Stockholm University, Sweden

Correspondence: Laura Sánchez, Stockholm University, Sweden.

Received: October 1, 2018 Accepted: January 7, 2019 Online Published: January 10, 2019

doi: 10.5539/elt.v12n2p113 URL: <https://doi.org/10.5539/elt.v12n2p113>

## Abstract

In an era of globalization, multilingualism is no longer the exception but the rule. To fully promote the communicative and functional development of multilingual learners, different educational programs at school have been implemented in the past few years. The study presented here is a contribution in this direction, by investigating the linguistic development in a second foreign language (L3 English) of students who were exposed to asymmetric input exposure in this and another non-native language. This other language was the L2 (German), which the learners ( $n= 50$ , 11-13 years old) were studying in a CLIL context together with the native languages of the participants (Spanish and Catalan). Within such a context, the purpose of the study is two-fold. On the one hand, it attempts to investigate the linguistic attainment in the L3, measured in terms of various dimensions of writing performance in a composition task, namely, fluency, accuracy, structural complexity and lexical complexity. On the other hand, the study aimed at exploring the contribution of various components of working memory (short-term memory and attention switching) to performance in these linguistic dimensions. The results suggest that individual differences in storage capacity and in capability for attention and task switching had an effect on the accuracy of the learners' performance, and to a lesser extent, also lexical complexity. Moreover, it was found that in general, attentional abilities had a greater impact than memory. These results are discussed in light of previous findings, while proposing directions for further investigation in the field.

**Keywords:** CLIL, working memory, attention, cognitive abilities, L2 writing, L3 acquisition, third language learning, English as a third language, CAF, multilingualism in school contexts

## 1. Introduction

Research on the role of individual differences (IDs) in second language learning and performance has increased in recent years thanks to a renewed interest in inter-individual variation caused by differences in cognitive abilities such as working memory, and also to the re-examination of working memory as a component of language aptitude explaining success in language learning (Doughty, 2018; Singleton, 2017). The present study is a contribution in this line of investigation, as it addresses the potential effect that different components of working memory may have on the accuracy, fluency and complexity of L2 written production.

Most research on the role of working memory in second language learning is based on Baddeley's model (2003), in which working memory is seen as "the temporary storage and manipulation of information that is assumed to be necessary for a wide range of cognitive abilities" (Baddeley, 2003, p. 189). In this model (see Table 1 adapted from Gilabert & Muñoz, 2010), working memory essentially consists of an attentional control system (i.e., the central executive), and its two slave systems, namely the phonological loop and the visuo-spatial sketchpad, which are responsible for the storage of verbal and visual information, respectively. In Baddeley's definition, storage capacity relates to the short-term memory component of working memory and refers to memory ability, that is, "memory maintenance" (Kane & Engle, 2003). This view is consistent with other definitions that highlight the retention element of working memory, such as Paradis (2009, p. 49), who defines it as the "retention of items held for immediate use" (see also Young & Pigott, 1999; Stewart, 2002; but see arguments for the separability of short-term memory from working memory in Conway, 2008, Engle, 2002; Miyake, Friedman, Emerson, Witi, & Howerter, 2000). In turn, the manipulation part of Baddeley's definition relates to the processing component of working memory, which involves various executive functions, including updating of information, inhibition of irrelevant information, and attention allocation or task switching (also referred to as

'shifting' by other authors, e.g. Miyake et al., 2000).

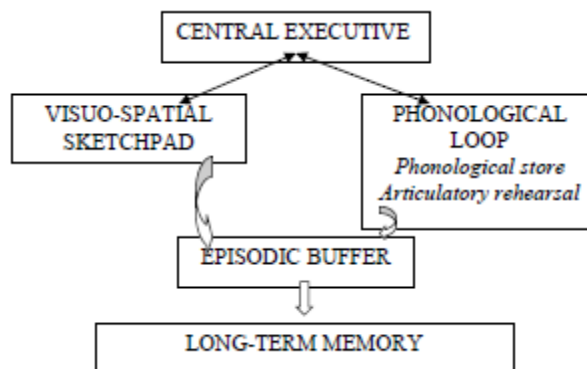


Figure 1. Baddeley's model (2003) working memory

For the overall aims of this contribution it is necessary to highlight the multicomponential nature of working memory, or more precisely, the fractionary view of it in that different executive functions may play different roles (e.g. Kane & Engel, 2003; Miyake et al., 2000) and are associated with different L2 cognitive processes (Bialystok, 2010; Doughty, 2018; Skehan, 2016; Wen, Biedrón & Skehan, 2017). According to these studies, working memory would be engaged in cognitive processes, such as input processing, noticing, mapping, pattern recognition, complexification, error avoidance or even handling feedback. Likewise, different components and executive functions, though related to each other, seem to make differential contributions to performance on complex executive tasks (e.g. Baddeley, 1996; Friedman & Miyake, 2004; Miyake et al., 2000). Miyake et al. (2002, p. 52) claim that there are “dissociations in performance” related to different executive functions, which constitutes a compelling argument to fraction the central executive to obtain a clearer representation of the specific contribution of each executive function to different aspects and dimensions of linguistic performance in oral and written production. This is precisely the primary aim of the study presented later, which explores the contributions of storage capacity and attention switching in various dimensions of written performance.

## 2. Background of the Study

The cognitive factor investigated in the present study, working memory, has been found to be a robust predictor of second language learning (for recent reviews, see Wen, 2016; Wen et al., 2017; Singleton, 2017), above all in studies on vocabulary and grammar learning (Sanz, Lin, Lado, Stafford & Bowden, 2014). However, few studies (Gilbert & Muñoz, 2010; Guara-Tavarés, 2013; Kormos & Safar, 2008; Kormos, 2012b; Kormos & Trebits, 2011; Mota, 2003; Tagarelli, Ruiz, Moreno & Rebuschat, 2016; Weissheimer & Mota, 2011; Zalbidea, 2017) have examined the relationship between working memory and different dimensions of language performance, namely, complexity (both structural and lexical), accuracy and fluency (henceforth, CAF). Furthermore, these studies have come to mixed results, as will be discussed in more detail below, and they often address only one or two dimensions, with only some studies including the four dimensions of linguistic performance. Besides, most studies on the relationship between working memory and CAF performance have tapped into the updating executive function, using tasks that measure both the processing and storage components of working memory by means, for instance, of reading or operational span tasks, or by backward digit or letter span tasks (which are different from tasks that measure only the storage component of working memory, as for example forward digit and letter span tasks). In contrast, very little attention has been paid to other executive functions such as attention switching (with remarkable exceptions such as Recio, 2011 and Pownall, 2016), or inhibition. Among the most common tests used to measure attention switching, the Trail Making Test (TMT), originally developed by Reitan (1958) is a good example, and considered a reliable assessment of executive control (Arbuthnott & Frank, 2002; Bialystok, 2010) in its pen-and-paper version.

The TMT is a neuropsychological test assumed to measure cognitive domains such as processing speed, sequencing, and mental flexibility (Bowie & Harvey, 2006), and consists of two parts, Trails A and Trails B. Both parts consist of 25 dots (that is, consecutive targets) on a sheet of paper. In part A all targets are numbers 1 to 25, and participants are asked to connect all of them in sequential order as quickly as possible. Hence, this part of the test involves visual scanning and spatial skills (Gaudino, Geisler & Squires, 1995) and helps to determine

a person's ability to control attention, that is, to direct and maintain attention and thought to a selected goal (Arbuthott & Frank, 2000, 518). In part B, where the targets alternate numbers and letters, learners have to connect them alternatively (1A, 2B, 3C...), which involves alternating attention between two simultaneous goals. Hence, part B is a measure of task alternation ability (Arbuthott & Frank, 2000), that is, "the ability to flexibly shift the course of an ongoing activity" (Kortte, Horner & Windham, 2002, p. 107), because it offers information about the shift in processing efficiency when attention is directed to more than one stimulus.

Different types of scores are commonly reported when calculating the scores of this test. For some authors "[t]he primary variables of interest are the total time to completion for parts A and B" (Bowie & Harvey, 2006: 2277; Bialystok, 2010; Kortte et al., 2002). This procedure is supported by Fellows, Dahmen, Cook & Schmitter-Edgecombe's consideration (2017, p. 155) that "both conditions of the test have been used under the assumption that they are measuring overlapping but also different cognitive processes". Hence, using the scores of the primary variables can be a way to preserve the uniqueness of the contribution of the abilities measured by each part. Other authors calculate derived scores based on the B/A ratio and the B-A difference, arguing that the switch cost between the primary scores reflect "the additional cognitive control required to switch between sequential numbers and letters" (Arbuthott & Frank, 2000, p. 519). However, this type of score is "best conceptualized as reflecting a combination of several cognitive functions" (Kortte et al., 2002), which complicates the identification and isolation of specific executive functions (see also Fellows et al., 2017). Naturally, both the primary and the derived scores are valid measurements in the TMT, and the choice of one or another may be driven by the particular needs of each study.

Back to the relationship between working memory and CAF performance mentioned at the beginning of this section, it is necessary to briefly make a case for further research in this line of investigation, especially as regards its theoretical relevance and motivation. For a start, the rationale for a (causal) relationship between working memory and language production stems from the fact that, apparently, learners with a higher working memory might be more able to engage in self-monitoring. This, presumably, would help to explain why learners with higher working memory are sometimes found to produce fewer errors and more error-free clauses (e.g. Bergsleithner, 2010; Gilabert & Muñoz, 2010; Mota, 2003; Rosen & Engel, 1997). Another argument in favour of this relationship would have to do with retrieval, focusing on the role played by attention during the formulation stage of speech production. More specifically, there is some evidence to suggest that during language production, the ability to control attention in the formulation stage is more important than in the conceptualization phase, at least as far as the accuracy of the language produced is concerned (e.g. Guara-Tavarés, 2013).

In addition, two more arguments that have been cited as possible explanations for a relationship between working memory and CAF in L2 production are the learners' ability to notice regularities in the input (Kormos, 2012a, 2012b), and the speed of their lexical searches. As regards noticing (Truscott, 1998) and the attention component of working memory, Sawyer and Ranta (2001) claim that "[a]ssuming that noticing is crucial to learning, and attention is required for noticing, and attention at any moment is limited by working memory, then there must logically be a close relationship between amount of learning and size of working memory". With attention being fundamental for noticing, some studies have found that learners with higher working memory might be more capable of noticing formal aspects of language (Schmidt, 1990, 2010), which would explain why they exhibit higher levels of accuracy in their performance (e.g. Bergsleithner, 2007; Mackey, Adams, Stafford & Winke, 2010). However, the results here are not consistent, as other studies have not found any relationship between noticing and performance (Mota, 2007).

Likewise, it has been argued that learners with lower working memory might also be slower in their lexical searches. This would explain why their production seems to be worse in terms of fluency and lexical density (e.g. Weissheimer & Mota, 2011), even though clearly much more research is needed in the field to validate these findings and to have a better understanding of these cognitive effects. Conversely, learners with higher working memory would have a better coordination of processing and storage, resulting in a higher lexical variety in their production (e.g. Gilabert & Muñoz, 2010; Weissheimer & Mota, 2011). Because language production is a complex process that involves a series of simultaneous operations that require conscious attention, working memory may play a role in explaining different levels of success in performing highly demanding cognitive tasks (as would be the case in writing, i.e. the modality of language production investigated in the present study). Gilabert and Muñoz (2010, p. 24), for example, suggest that "[d]uring L2 production conceptual chunks of the conceptualizer need to be held in memory while a set of other cognitive operations take place (e.g. accessing words and organizing them in a sequence, providing their morphological and phonological form, while monitoring the whole process and maintaining fluency).

On the other hand, attentional resources are limited (among others: Anderson, 1983, Baddeley, 2003; Kane & Engle, 2003; Lepsien, Thornton & Nobrea, 2011; Magimairaj & Montgomery, 2013; Miyake et al., 2000; Paradis, 1993), and different aspects of performance compete for attentional resources (e.g. Guara-Tavarés, 2013; Forthkamp, 1999; Forthkamp & Bergsleithner, 2010; Robinson, 2001; Skehan, 1998). This, for instance, would be the case of the CAF dimensions of linguistic performance investigated in the empirical part of this paper. Precisely because of this competition and because attentional resources are limited, it is reasonable to argue that an individual cannot pay attention to several things at the same time with the same level of efficiency and efficacy. As this regards, Skehan (2009, p. 511, emphasis added) proposes that “if performance in each of these areas, complexity, accuracy, and fluency (...) require attention and working memory involvement, then *committing attentional resources to one may have a negative impact on others*”. The author further argues that there is “a tension between form (complexity and accuracy), on the one hand, and fluency, on the other” (Note 1).

Some of the studies mentioned in the preceding paragraphs offer evidence that learners with higher working memory may exhibit a more accurate, fluent, grammatically complex and lexically rich production. However, at this point results are still inconclusive, because other studies have not found any relationship between working memory and these dimensions of linguistic performance, or have found advantages in some of them but not in others (e.g. Bergsleithner, 2010; Guara-Tavarés, 2013; Forthkamp, 1999; Mizera, 2006; Mota, 2003; Prebianca & D’Ely, 2008; Rezaei & Okhovat, 2016; Weissheimer & Mota, 2009, 2011). In addition to these mixed results, some of the findings in these studies could be explained by inter-learner variability in proficiency level, which would blur the identification of cognitive effects on performance (Gilabert & Muñoz, 2010; Kormos & Sáfár, 2008). In this sense, it is worthy of note that in many studies on the effect of working memory on L2 learning the proficiency factor is not taken into account. This is methodologically relevant in that the role played by memory and attention components of working memory might be different at different stages of development in the target language (see review in Serafini & Sanz, 2015). On the other hand, it is possible that at least in adults, cognitive factors correlate with proficiency (e.g. Gilabert & Muñoz, 2010; Mizera, 2006). If so, in order not to confound them, it is necessary to use a research design and methodology that make it possible to tease apart the effects of these factors. Because this has hardly been done, the findings on cognitive effects or lack thereof reported in these studies could instead be the result of proficiency rather than reflect cognitive shortcomings on performance. Likewise, it could also be plausible that cognitive effects are mediated by proficiency.

In addition to these methodological considerations, it must also be noted that many theoretical claims in the literature on cognitive effects in CAF performance have been proposed within the Task-Based Language Teaching framework. Studies within this framework are conducted with an eye to providing detailed explanations of how, if at all, cognitive overload may constrain the development of CAF in L2 production. However, less of an effort is made to overtly examining whether IDs in working memory have a bearing on CAF performance. As a consequence, cognitive effects are sometimes assumed rather than demonstrated, as the explanations offered in some of these studies are sometimes theoretical and need to be empirically substantiated using actual data. As a corollary, most studies cited here rely on correlational research designs. As Guara-Tavarés (2013, p. 26) warns us, “these correlations do not reveal whether the differences between the performance of lower and higher spans are significant”, but simply that they are somehow related. It goes without saying that the verification of a relationship is an important step, yet insufficient for comparison and discrimination of performance between learners with different cognitive abilities.

Finally, the results in previous studies need to be expanded in terms of the type and amount of dependent variables that are used in the assessment of linguistic performance. Probably the most urgent need is the area of fluency, given that the great majority of studies have addressed performance in oral production. In the case of written production, these variables are not an option because fluent texts cannot be determined on the grounds of syllable per minute or pauses. Bearing in mind the methodological issues raised here, the study presented in the next section tries to shed some light on the role played by IDs in working memory on CAF written performance in English as a foreign language. This is done by using a larger number of analytical measures both of cognitive abilities and of linguistic performance, and also by employing more robust tests in the statistical treatment of the data. In addition to this, a novelty of the study is that it is conducted within a CLIL (Content and Language Integrated Learning) context.

### 3. The Present Study

The study presented here is exploratory, and it sets out to investigate the relationship between various components of working memory and different linguistic dimensions of CAF performance. More precisely, it tries to find out whether IDs in storage capacity (phonological loop) and switching ability (central executive) have

any impact on L3 written production when it comes to the fluency, accuracy and complexity of the texts produced by school-age learners in a CLIL context. Specifically, the research question guiding the study is formulated as follows:

RQ: Do IDs in storage capacity and switching ability have an effect on the fluency, accuracy, structural complexity and lexical complexity of the L3 English written production of CLIL learners?

### 3.1 Participants

In order to answer this research question, data were gathered from learners of English as a foreign language ( $n=50$ ). They were instructed learners of English in Catalonia, and fell within the age range 11-13. The native language of the participants was Spanish, but they were also functional users of Catalan. The participants were selected from a larger pool of data (Sánchez, 2015b) on the basis of their responses in a language background questionnaire they were administered before the actual data collection started. Only learners were selected who were born and raised in Spanish-speaking homes and whose both parents were Spanish. As far as the target language English is concerned, none of the participants had extracurricular exposure outside school in terms of formal instruction or naturalistic exposure in stays abroad.

More relevant for the spirit of this volume is the fact that the participants had as well knowledge of German as an L2, a language they were learning in a CLIL programme that combined language learning and contents (Sánchez, 2014; 2015a; 2015b). For the sake of clarity, some terminological clarification is needed on the status of each language in the linguistic repertoire of the learners. Even though from a chronological point of view German would be the L3 and English the L4, the status of German in this study was that of an L2, in the understanding that a true L2 corresponds to “the first encounter with a non-native language” (Falk & Bardel, 2010, p. 61). Hence, English was a second non-native language, and therefore an L3. Following Cenoz and Valencia (1994) the relationship between Spanish, Catalan and German would be one of additive trilingualism, whereas English was learnt as a foreign language in a setting where it was not present in the wider environment and in a regular language teaching model within a communication-oriented program.

In this CLIL programme, around half the school subjects were taught in the L2 German, as in history, geography or arts (for more details see Sánchez, 2015b), whereas in the other half the subjects the language of instruction was either Spanish or Catalan. The learners’ exposure to German was essentially restricted to the classroom, with the exception of occasional extracurricular activities organized by the school. Besides, as indicated above, the parents were not native speakers of German and they had little or no knowledge of German. A final remark on the learning of the two non-native languages investigated in the present study is the concurrent learning (Todeva & Cenoz, 2009) of the L2 German and the L3 English in spite of the earlier age of onset of the former. The learners had started learning the L2 German in Grade 1 around the age of 6 (some of them a bit earlier, when they were 5). In turn, they had started learning English in Grade 4, when they were 9 years old and with the L2 German still under development (Sánchez, under review).

### 3.2 Instruments

The battery of tests employed in the data collection included the following instruments: a language background questionnaire, an EFL proficiency test, two cognitive tests, and a composition task meant to elicit written data in the target language (L3 English). As mentioned above, the participants were selected upon completion of a language background questionnaire. This questionnaire was adapted from an already existing one (Sánchez, 2014), which had already been piloted and used with learners of equivalent socioeconomic status and linguistic profile (Sánchez, 2012). Furthermore, the proficiency level in English of the participants was measured by means of the Oxford Placement Test (OPT), which gives an indication of overall proficiency, assessed by questions addressing vocabulary and grammar points as well as listening comprehension. Even though proficiency was not a factor under investigation in the present study, it was necessary to control for its possibly interfering effects and prevent them from biasing the results. This decision was made in light of the theoretical and methodological issues raised in the literature review. In congruence with these, proficiency in the L3 English was measured and later taken into account when conducting the statistical analysis of the data.

Two cognitive tests were used in order to measure different components of the participants’ working memory. The learners’ storage capacity, that is, the short-term memory component of working memory, was measured by means of a Forward Letter Span Task (FLST) designed and piloted within the GRAL research group. The FLST is a computerized task in which learners are presented with blocks of 9 letters that appear consecutively after each other on the screen. After each block learners are confronted with 12 letters, and they have to tick the letters they have just seen, and in the same order in which they saw them. The learners’ capacity for memory maintenance corresponds to the longest string of letters they can recall and repeat correctly in the same order in

which they were presented (Miller, 1956). A score is automatically calculated in the output of the test based on a partial credit unit scoring procedure. The Trail Making Test (TMT) in its pen-and-paper version was used to tap into the central executive, by measuring the learners' ability to control attention (Trails A) and their task alternation ability (Trails B), respectively. Finally, the language sample elicited for the analysis of fluency, accuracy, structural complexity and lexical complexity of the participants' written production in English was based on a composition entitled 'My life: past, present, and future expectations'. In this task, the participants were asked to write about their childhood, their life at present, and the expectations they had for their future life.

### 3.3 Procedure

With the purpose of guaranteeing the most adequate conditions for the data collection, all the tasks except for the cognitive tests were carried out in class time and in presence of both the teacher and the researcher. The instructions for the questionnaire were given in Spanish, while those for the composition task and OPT were given in English. Irrespective of the language of the instruction for each task, requests for clarification in Spanish were addressed in this language.

As stated in the description of the participants, the age range of the participants investigated was 11-13. The data collection involved three classes of EFL learners (Grades 5, 6 and 7), and took place in three sessions. In the first session each of the three classes filled out the questionnaire and wrote the composition task in the classroom. This last task was time-controlled, and participants were given 20 minutes to do so without using dictionaries, grammar books, ipads or any other reference tool. Before starting the task, they were informed that the results would not count as a class score. In the second session they carried out the OPT, for which they had as much time as they needed (with a maximum of 1 hour, which was the duration of the class). The cognitive tests were administered on the third session, and only to the 50 participants who met the criteria for inclusion in the sample analysed in this study. For the completion of these tests, each participant was taken individually to a small computer room in the school, so that they could carry out the cognitive tasks in as silent and quiet way as possible.

### 3.4 Data Analysis

The research design of the study reported here relied on a series of Ancovas (Analysis of Co-variance) using the scores obtained in the cognitive tests measuring the participants' short-term memory and attention switching as fixed factors. In order to avoid undesired proficiency effects biasing the results, it was necessary to control for the effects of this factor. With this purpose, the OPT tests were proofread by the researcher, and the raw scores registered in a continuous variable in SPSS (Statistical Package for the Social Sciences, version 23, the research software chosen for the statistical treatment of the data). The raw scores were used as covariates. In addition to this, and despite the fact that the ages of the participants were very similar, it was felt that pubertal changes associated to these ages might eventually distort the picture. To avoid this risk, the learners' biological ages at time-of-testing were also included in the Ancova(s) as a covariate. The dependent variables in the Ancovas were the different CAF measures that were used in the assessment of the participants' linguistic production.

In order to use the cognitive measures as fixed factors, it was necessary to transform the original continuous variables into dichotomous variables that would make it feasible to break down the sample into two groups according to the learners' cognitive abilities (see Table 1). To this aim, the scores obtained in the measure of short-term memory (resulting from the FLST) were used in order to divide participants into either low- or high-memory spans. In a similar manner, the scores for the measures of Trails A and Trails B (corresponding to parts A and B of the TMT) served to classify participants into high and low attention spans. Table 1 shows the classification of participants according to high and low cognitive abilities in each of the three measures (i.e., storage capacity, ability to control attention and task alternation ability).

Table 1. Participants' classification according to cognitive abilities

	<b>Storage Capacity (STM)</b>	<b>Ability to Control attention (TRAILS A)</b>	<b>Task Alternation Ability (TRAILS B)</b>
<b>LOW (n)</b>	20	26	22
<b>HIGH (n)</b>	30	24	27

The L3 language sample elicited (through the composition task) was used for the calculation of different analytical measures that assessed the participants' linguistic production in terms of fluency, accuracy, structural complexity and lexical complexity. These measures, used later as dependent variables in the Ancovas, are summarized in Table 2 below. The measurement of complexity was two-fold. On the one hand, structural complexity was defined here as size, elaborateness, richness and diversity (Housen & Kuiken, 2009) at the sentential level (Bulté & Housen, 2012). The measures employed were the number of coordinate and subordinate clauses, the ratio of clauses per sentence, and the ratio of coordinate and subordinate clauses per sentence. On the other hand, lexical complexity was understood as lexical richness, that is, as "the range of vocabulary and avoidance of repetition" (Malvern, Richards, Chipere, & Durán, 2004, p. 3). To this aim, the indexes of lexical diversity (Guiraud's Index, Uber Index) and lexical density (ratio of lexical words to function words) were calculated, alongside the plain number of types and lexical tokens. The fluency dimension was defined as the ability to process the L2 with ease and rapid speed (Housen & Kuiken, 2009, p. 462; Lennon, 1990, p. 390). In consonance with previous research (Wolfe-Quintero, Inagaki & Kim, 1998), it was measured using a frequency of the total number of words, sentences, and clauses, along with a ratio of words per sentence and words per clause. The last dimension of linguistic production examined was accuracy, which related to the ability to produce target-like use (Ellis, 2003, p. 340; Pallotti, 2009, pp. 591-592). It was measured using the frequency of error-free clauses and error-free sentences (counts), and the percentage of error-free clauses over total number of clauses, and of error-free sentences over total number of sentences.

Table 2. Analytical measures used in the analysis of L3 English written production

<b>STRUCT. COMP.</b>	<b>LEXICAL COMP.</b>	<b>FLUENCY</b>	<b>ACCURACY</b>
<b>CoorC</b> (Nr. of coordinate cl.)	<b>LV</b> (Lexical Variation)	<b>W</b> (Nr. of words)	<b>EFC</b> (Nr. of error-free clauses)
<b>SubC</b> (Nr. of subordinate cl.)	Guiraud's Index Uber's Index	<b>S</b> (Nr. of sentences)	<b>EFS</b> (Nr. of error-free sentences)
<b>C/S</b> (Cl. × Sentence)	<b>LD</b> (Lexical Density)	<b>C</b> (Nr. of clauses)	<b>%EFC</b> (Percentage of error-free clauses)
<b>CoorS</b> (Ratio Coord. Cl × Sentence)	Lexical words/ Function words	<b>W/C</b> (Words × clause)	<b>%EFS</b> (Percentage of error-free sentences)
<b>SubS</b> (Ratio Subord. Cl × Sentence)	<b>Ty</b> (Nr. of types)	<b>W/S</b> (Words × sent.)	
	<b>To</b> (Nr. of tokens)		

#### 4. Results

The study just presented aimed at ascertaining whether IDs in storage capacity and switching ability had any impact on L3 written production when it comes to the fluency, accuracy, structural complexity and lexical complexity of the texts produced by school-age learners in a CLIL context. The report of the results starts with a brief overview of the descriptive statistics of the participants' performance in the cognitive tests according to ability group (i.e. high vs. low) in each measure (Table 3).

Table 3. Descriptive statistics of participants' performance in the cognitive tests

		Mean	Std. Deviation	Minim.	Maxim.	Median	Mode
<b>Storage Capacity</b>	HIGH	.630	.07652	.54	.77	.610	.54 <sup>a</sup>
	LOW	.403	.10658	.07	.53	.410	.34 <sup>a</sup>
<b>Ability to control attention</b>	HIGH	24.39	3.365	16.78	27.78	24.500	16.78 <sub>a</sub>
	LOW	42.64	11.855	28.91	83.00	39.735	38.78
<b>Task alternation ability</b>	HIGH	31.48	6.511	22.33	46.78	31.160	22.33 <sup>a</sup>
	LOW	42.18	13.535	16.78	83.00	41.620	38.78

<sup>a</sup> Multiple modes exist. The smallest value is shown.

The remainder of the section reports the results of the series of Ancovas that were run on the data in order to find out whether the components of working memory investigated here had an effect on CAF performance. For the sake of clarity, the results of the three measures are considered together, and presented for each dimension of language performance separately. They are summarized in Table 4, which roughly indicates the dimensions where significant differences were obtained between high and low cognitive abilities.

Table 4. Summary of cognitive effects on structural complexity, lexical complexity, fluency and accuracy

	Structural Complexity	Lexical Complexity	Fluency	Accuracy
<b>Storage Capacity</b>	-	✓	-	✓
<b>Ability to Control Attention</b>	✓	-	(✓)	(✓)
<b>Task Alternation Ability</b>	✓	✓	✓	✓

In the structural complexity dimension, a clear distinction could be made between the memory and attentional components. After controlling for proficiency and age, short-term memory did not have any significant effect in any of the analytical measures that tapped into structural complexity. Conversely, the attentional components did have an effect on structural complexity, with each component affecting a different aspect of sentential complexity. Whereas ability to control attention turned out to have a significant effect on the measures of subordination (SubC,  $p = .032$ ; SubS,  $p = .024$ ), task alternation ability was significant in measures of coordination ( $p = .000$  for both CoorC and CoorS). Moreover, both attentional components had a significant effect on the ratio of clauses per sentence ( $p = .025$  and  $p = .014$  for ability to control attention and task alternation, respectively). The results were somewhat different when looking at lexical complexity, with short-term memory and task alternation ability (but not ability to control attention) yielding significant effects. Significant differences between low and high memory spans were found in Guiraud's index ( $p = .011$ ), Uber ( $p = .024$ ) and plain number of types ( $p = .034$ ). The same pattern was found when comparing low and high attention switching spans (Guiraud:  $p = .002$ ; Uber:  $p = .014$ ; Types:  $p = .006$ ). Here also the difference in plain number of tokens was significant ( $p = .048$ ).

The pattern found for fluency was similar to that observed for structural complexity, in that attention but not memory turned out to be statistically significant. To be more precise, while no effect was found for short-term memory, both ability to control attention and task alternation ability seemed to have a significant effect on some of the analytical measures. Attention control had an effect on S produced ( $p = .010$ ), whereas attention switching affected the W ( $p = .039$ ) and the ratio of W/C ( $p = .037$ ) and W/S ( $p = .015$ ). The results were more uniform when it came to accuracy, in the sense that the three measures of working memory investigated had significant effects on one or more of the measures employed in the analysis of accuracy. First, the performance of higher memory spans was significantly better in all the accuracy measures (EFC,  $p = .014$ ; EFS,  $p = .035$ ; %EFC,  $p = .003$ ; %EFS,  $p = .012$ ). Second, ability to control attention had a significant effect on EFS ( $p = .014$ ) only, and task alternation ability discriminated between participants with higher and lower attentional abilities in measures at the clause level (EFC,  $p = .008$ ; %EFC,  $p = .007$ ).



## 5. Discussion and Concluding Remarks

The study presented here inquired into the role of working memory in a CLIL learning context. In particular, the study was guided by the question whether IDs in short-term memory and switching ability had any effect on the fluency, accuracy, structural complexity or lexical complexity of the participants' L3 written production. To this aim, various kinds of data were gathered, including an L3 language sample and three measures tapping into different components of working memory, storage capacity, ability to control attention, and task alternation ability. At the same time, the administration of a standardized L3 proficiency test guaranteed that eventual variation in levels of linguistic competence would not distort the picture. The data were analyzed with the purpose of identifying potential variability in the performance of learners with high and low cognitive abilities. In what follows the main findings are discussed, while explanations are sought for the generally poorer linguistic performance of learners with lower working memory. Likewise, based on the data assembled in this study, the discussion here tries to tease apart the effects related to shortcomings in memory storage from those stemming from either insufficient control of attention allocation or of inefficient distributed attention.

All in all, considering the evidence obtained in the various analyses of the data, it seems that the answer to the research question would be confirmatory, in that the different components of working memory analyzed did have an effect on the L3 written production of the CLIL learners investigated. This was true even if these effects were disparate in the different linguistic dimensions investigated, and also as regards the markedly differential effects for memory and attentional components, respectively. Generally speaking, the attentional components of working memory seemed to play a greater role than the memory component.

A clear pattern emerged from the analysis of linguistic performance in the dimension of complexity. The results of the Ancova yielded a significant difference between low and high memory spans in various measures of lexical diversity, and the same was true in relation to task alternation ability. On the contrary, the two attentional but not the memory components were engaged in structural complexity. Specifically, learners with higher ability to control attention and higher task alternation ability made a statistically higher use of subordination and coordination, respectively. In addition to this, both attentional components discriminated between learners with higher and lower abilities in their ratio of clauses per sentence, with the higher ability learners exhibiting a higher ratio. A similar result was obtained in the fluency dimension. Thus, the memory component did not seem to affect the fluency of the participants' production, while the two attentional components did. To be more precise, the texts were longer in learners with higher attentional abilities, as reflected by the significant differences obtained for ability to control attention in the number of sentences, and for task alternation ability in the ratio of words per clause and per sentence. Furthermore, the dimension that appeared to be the most affected by the components of working memory investigated here was accuracy. In this sense, it is relevant to highlight that all the three measures had a significant effect on the accuracy of the texts produced. Thus, learners with higher memory maintenance and attention switching seemed to write more accurate texts than their lower-ability peers. This was also true, though to a lesser extent, for learners with a higher ability to control attention.

At a more abstract level, the discussion of these results calls for a thorough search of explanations that account for the findings in relation to the benefits and drawbacks associated with IDs in working memory. Moreover, these explanations are necessary for a better (and deeper) understanding of the complex interaction between cognitive processes in the mind of the foreign language learner. To begin with, consistent with the assertion that the linguistic dimension found to be most affected by IDs in the working memory components explored here was accuracy, a few preliminary remarks can be made. This impact of working memory on accuracy lends confirmation to findings reported in previous studies discussed in the opening of this paper (Bergsleithner, 2010; Gilabert & Muñoz, 2010; Kormos, 2012a; Mota, 2003; Sawyer & Ranta, 2001; but see Mota, 2007). In explaining this result, some authors have argued that learners with a higher working memory are better ready to notice regularities in the input (Bergsleithner, 2007; Mackey et al., 2010; Sánchez & Bardel, 2016; Bardel & Sánchez, 2017), and they would also be more able to engage in self-monitoring (Bergsleithner, 2010; Gilabert & Muñoz, 2010; Mota, 2003; Rosen & Engel, 1997). In the understanding that noticing draws on linguistic representations that are stored by declarative knowledge (Baddeley, 2003; Ellis, 2001), it might be argued that lower working memory might have obstructed access to these representations (see also Sánchez & Bardel, 2016). To put it another way, a less efficient coordination of the various cognitive processes that occur in working memory would not only have slowed down the speed of mental access, but also impaired the effectiveness of information encoding and retrieval. The data from lower-ability learners in this study seems to suggest that shortcomings related to working memory would lead to a less accurate performance, since these learners produced significantly fewer error-free clauses and sentences.

In addition to this, the efforts made by low-ability learners would have consumed a great deal of cognitive

resources. In the case of high-ability learners, the fluency of their production was not compromised by an extra effort put in the attempt to enhance (or at least keep) the quality of the texts written. Notwithstanding, to have a more comprehensive view of the picture, future studies should examine more carefully the presence or absence of trade-offs between different linguistic dimensions in low-ability learners in a more contextualized framework comparing Skehan's limited attentional capacity and Robinson's cognition hypotheses (see comparison of both in Skehan, 2015). It is also likely that the higher speed of access of high-ability learners might have saved them time and cognitive resources which they could otherwise have then relocated to other aspects of their production. With complexity (both structural and lexical) being a case in point, the results yielded by the Ancova indicated that learners with higher abilities to control attention and alternate between tasks used significantly higher coordination and subordination, respectively. Hence, the data proved these high-ability learners to be able to efficiently pay simultaneous attention to different aspects of their production at a time, or to be more precise, to continually shift attention back and forward between different aspects.

By not exhausting their cognitive resources on a sole linguistic dimension, accuracy in this case, learners with higher switching ability had more chances to concentrate on more sophisticated aspects of their text production. While not compromising accuracy, these learners wrote longer texts which were also fluent, as indicated by their higher rate of words per sentence. This reflects a more thorough, thoughtful and carefully planned writing process, including the conceptualization of the message in the pre-writing stage and the subsequent execution of the written speech plan (Kormos, 2012b). In light of these results, higher-ability learners' more elaborated performance in the fluency dimension is suggestive of a better organization and structuring of their ideas and thoughts, and of how they divided them into syntactic units. Likewise, they seemed to be more successful in establishing the necessary connections between ideas, and in determining and expressing the relationship, hierarchy, or chronological order of the events or situations expressed in these syntactic units. Moreover, a larger memory span contributed to enhancing the lexical complexity of the production of learners with a higher short-term memory, as shown, for example, by their significantly higher Guiraud's and Uber's indices. In light of this result, one can speculate that this superiority was due to the fact that these learners were more capable of retaining the lexical items they had employed in their previous discourse. This would have enabled them to avoid repetition and use a richer and more varied lexical repertoire, also aided by their better switching ability.

Another possible explanation for these results might have to do with how learners with different cognitive abilities coped with and integrated the different types of information they need to bear in mind while doing the task (Kormos & Trebits, 2012). In the present study, it seems that participants with lower memory and attentional abilities were less able to handle the content information and planning inherent to the task and the encoding of grammatical and lexical information. To put it another way, participants with higher abilities would have been better at monitoring their attention resources by giving each piece of information sufficient attention, that is, the attention necessary for the successful accomplishment of the task. Conversely, participants with lower abilities had a harder time addressing conceptual information (including, but not limited to) the perception and categorization of the situations described, the ordering of these situations and the establishment of temporal relations between them, as pointed out above.

As indicated in the opening of this section, the discussion here pursued to identify and isolate differential contributions of each executive function (storage capacity, ability to control attention and task alternation ability) to various dimensions of linguistic performance in written production. The data in this study suggests that attentional components had a greater influence than the memory component. Moreover, when comparing the two measures of attention, the ability to alternate between tasks was clearly much more influential than the ability to solely direct and maintain attention. On the other hand, performance in structural complexity and fluency turned out to be related only to the attention switching executive function, but not to storage capacity. In any case, the study presented here is only exploratory, and further research should replicate this study with a larger sample and different-aged participants. Furthermore, the research design should incorporate derived scores of the trail making scores, as well as measures of inhibition, with the purpose of comparing which executive functions are recruited for which L2 cognitive processes. The data from this study also opens an important question as regards the possibility that the CLIL learners investigated here might have an advantage over language learners in a non-CLIL context, perhaps because the language experience of CLIL learners might help them enhance a varied array of cognitive abilities, including those investigated here. This possibility should be explored in the very near future.

### **Acknowledgements**

The author acknowledges support by SGR-2014SGR1089 (from the Catalan Government) to the GRAL Research Group, with which she collaborates.

## References

- Alan, B. (2003). Working memory and language: An overview. *Journal of communication disorders*, 36, 189-208. [https://doi.org/10.1016/S0021-9924\(03\)00019-4](https://doi.org/10.1016/S0021-9924(03)00019-4)
- Anderson, J. (1983). *The architecture of cognition*. Cambridge: Harvard University Press
- Arbuthnott, K., & Frank, J. (2002). Trail making test, Part B as a measure of executive control: validation using a set-switching paradigm. *Journal of clinical and experimental psychology*, 22, 518-528.
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Neuroscience*, 4, 829-839.
- Baddeley, A. & Hitch, G. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47-89). New York: Academic Press.
- Beaman, P. (2010). Working memory and working attention. What could possibly evolve? *Current Anthropology*, 51, 27-38. <https://doi.org/10.1086/650297>
- Baddeley, A. (1996). Exploring the central executive. *The quarterly journal of experimental psychology*, 59A, 5-28. <https://doi.org/10.1080/027249896392784>
- Bardel, C., & Sánchez, L. (2017). The L2 status factor hypothesis revisited. In T. Angelovska, & A. Hahn (Eds.), *L3 syntactic transfer: models, new developments and implications* (pp. 85-101). Amsterdam: John Benjamins. <https://doi.org/10.1075/bpa.5.05bar>
- Bersleithner, J. (2007). *Working memory capacity, noticing, and L2 speech production*. Unpublished PhD, Universidad Federal de Santa Catarina.
- Bersleithner, J. (2010). Working memory capacity and L2 writing performance. *Ciências & Cognição*, 15, 2-20.
- Bialystok, E. (2010). Global-local and trail-making tasks by monolingual and bilingual children: beyond inhibition. *Developmental psychology*, 46, 93-105. <https://doi.org/10.1037/a0015466>
- Bowie, C., & Harvey, P. (2006). Administration and interpretation of the trail making test. *Nature protocols*, 1, 2277-2281. <https://doi.org/10.1038/nprot.2006.390>
- Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency. Complexity, accuracy and fluency in SLA* (pp. 21-46). Amsterdam: John Benjamins. <https://doi.org/10.1075/llt.32.02bul>
- Cenoz, J., & Valencia, J. (1994). Additive trilingualism: Evidence from the basque country. *Applied psycholinguistics*, 15, 195-207. <https://doi.org/10.1017/S0142716400005324>
- Conway, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in brain research*, 169, 323-338. [https://doi.org/10.1016/S0079-6123\(07\)00020-9](https://doi.org/10.1016/S0079-6123(07)00020-9)
- Doughty, C. (2018). Cognitive language aptitude. *Language learning*, 68, 1-26. <https://doi.org/10.1111/lang.12322>
- Ellis, N. (2001). Memory for language. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 33-68). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524780.004>
- Engle, R. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, 11, 19-23. <https://doi.org/10.1111/1467-8721.00160>
- Falk, Y., & Bardel, C. (2010). The study of the role of the background languages in third language acquisition: the state of the art. *International review of applied linguistics*, 48, 185-219. <https://doi.org/10.1515/iral.2010.009>
- Fellows, R., Dahmen, J., Cook, D., & Scmitter-Edgecombe, M. (2017). Multicomponential analysis of a digital trail making test. *Clinical neuropsychology*, 31, 154-167. <https://doi.org/10.1080/13854046.2016.1238510>
- Forthkamp, M. (1999). Working memory capacity and aspects of L2 production. *Communication and Cognition*, 32, 259-296.
- Forthkamp, M., & Bergsleithner (2010). Relationship among individual differences in working memory capacity, noticing, and L2 speech production. *Signo*, 32, 40-53.
- Friedman, N., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of experimental psychology: general*, 133, 101-135. <https://doi.org/10.1037/0096-3445.133.1.101>
- Gaudino, E., Geisler, M., & Squires, N. (1995). Construct validity in the trail making test: what makes part B

- harder? *Journal of clinical and experimental neuropsychology*, 17, 529-535. <https://doi.org/10.1080/01688639508405143>
- Gilabert, R., & Muñoz, C. (2010). Differences in attainment and performance in a foreign language: the role of working memory capacity. *International Journal of English Studies*, 10, 19-42. <https://doi.org/10.6018/ije/s/2010/1/113961>
- Guara-Tavarés, M., (2013). Working memory capacity and L2 speech performance in planned and spontaneous conditions: a correlational analysis. *Trabajos en Lingüística Aplicada*, 52, 9-29. <https://doi.org/10.1590/S0103-18132013000100002>
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy and fluency in second language acquisition. *Applied linguistics*, 30, 461-473. <https://doi.org/10.1093/applin/amp048>
- Kane, M., & Engle, R. (2003). Working memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132, 47-70. <https://doi.org/10.1037/0096-3445.132.1.47>
- Kormos, J. (2012a). New conceptualizations of language aptitude in second language attainment. In G. Granena, & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate L2 attainment* (pp. 131-152). Amsterdam: John Benjamins.
- Kormos, J. (2012b). The role of individual differences in L2 writing. *Journal of second language writing*, 21, 390-403. <https://doi.org/10.1016/j.jslw.2012.09.003>
- Kormos, J., & Safar, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and Cognition*, 11, 261-271. <https://doi.org/10.1017/S1366728908003416>
- Kormos, J., & Trebits, A. (2011). Working memory capacity and narrative performance. In P. Robinson (Ed.), *Second language task complexity. Researching the cognition hypothesis of language learning and performance* (pp. 267-285) Amsterdam: John Benjamins. <https://doi.org/10.1075/tblt.2.17ch10>
- Kortte, K., Horner, M., & Windham, W. (2002). The trail making test, part B: cognitive flexibility or ability to maintain set? *Applied neuropsychology*, 9, 106-109. [https://doi.org/10.1207/S15324826AN0902\\_5](https://doi.org/10.1207/S15324826AN0902_5)
- Lennon, P. (1990). Investigating fluency in EFL. A quantitative approach. *Language learning*, 40, 3387-417. <https://doi.org/10.1111/j.1467-1770.1990.tb00669.x>
- Lepsien, J., Thornton, I., & Nobre (2011). Modulation of working-memory maintenance by directed attention. *Neuropsychologia*, 49, 1569-1577. <https://doi.org/10.1016/j.neuropsychologia.2011.03.011>
- Mackey, A., Rebecca, A., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, 60, 501-533. <https://doi.org/10.1111/j.1467-9922.2010.00565.x>
- Magimairaj, B., & Montgomery, J. (2013). Examining the relative contribution of memory updating, attention focus switching, and sustained attention to children's verbal working memory span. *Child Development Research*, 2013, 1-12. <https://doi.org/10.1155/2013/763808>
- Malvern, D. Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development. Quantification and assessment*. Palgrave: MacMillan. <https://doi.org/10.1057/9780230511804>
- Miller, A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97. <https://doi.org/10.1037/h0043158>
- Miyake, A., Friedman, N., & Emerson, M. (2000). The unity and diversity of functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cognitive Psychology*, 41, 49-100. <https://doi.org/10.1006/cogp.1999.0734>
- Mizera, G. (2006). *Working memory and L2 oral fluency*. Unpublished PhD. University of Pittsburgh, Pennsylvania.
- Mota, M. (2003). Working memory capacity and fluency, accuracy, complexity, and lexical density in L2 speech production. *Fragmentos*, 24, 69-104.
- Pallotti, G. (2009). CAF: Definition, refining and differentiation constructs. *Applied linguistics*, 30, 590-601. <https://doi.org/10.1093/applin/amp045>
- Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of "interference" in bilingual

- speakers: The activation threshold hypothesis. *International Journal of Psycholinguistics*, 9, 133-145.
- Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Amsterdam: John Benjamins.
- Pownall, J. (2016). *Task complexity as mediated by proficiency, working memory, and attention: competition for cognitive resources during L2 oral task performance*. Unpublished PhD dissertation, University of Barcelona.
- Prebianca, G., & D'Ely, R. (2008). EFL speaking and individual differences in working memory capacity: Grammatical complexity and weighted lexical density in the oral production of beginners. *Singótica*, 20, 332-364.
- Recio, M. (2011). *The effects of task complexity on L2 oral production as mediated by differences in working memory capacity*. Unpublished Ms thesis, University of Barcelona.
- Reitan, R. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and motor skills*, 8, 211-216. <https://doi.org/10.2466/pms.1958.8.3.271>
- Rezail, M., & Okhovat, B. (2016). The effect of working memory on EFL learners' oral fluency. *International Journal of English Linguistics*, 6, 74-81. <https://doi.org/10.5539/ijel.v6n5p74>
- Robinson, P. (2001). *Cognition and second language instruction*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524780>
- Rosen, V., & Engle, R. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology* 126, 211-227. <https://doi.org/10.1037/0096-3445.126.3.211>
- Sánchez, L. (2012). 'Luisa and Pedrito's dog will the breakfast eat': interlanguage transfer and the role of the second language factor. In G. De Angelis, & J.M. Dewaele (Eds.), *New trends in crosslinguistic influence and multilingualism research* (pp. 86-104). Clevedon: Multilingual Matters.
- Paradis, M. (2014). The role of age on the development of written competence in L4 English. Evidence from a Spanish/ German CLIL context. In G. De Angelis, & A. Otwinowska (Eds.), *Teaching and learning in multilingual contexts. Sociolinguistic and educational perspectives* (pp. 125-144). Bristol: Multilingual Matters.
- Paradis, M. (2015a). The acquisition of written competence in learners of L3 English in a Spanish/ German content and language integrated learning context. In D. Marsh, M.L. Pérez, & J. Ráez (Eds.), *CLIL in the action: voices from the classroom* (pp. 64-81). Cambridge: Cambridge Publishing Scholars.
- Paradis, M. (2015b). L2 activation and blending in third language acquisition: evidence of crosslinguistic influence from the L2 in a longitudinal study on the acquisition of L3 English. *Bilingualism: Language and Cognition*, 18, 252-269. <https://doi.org/10.1017/S1366728914000091>
- Paradis, M. (under review) From L2 to L3, verbs getting into place: a study on transfer and L2 proficiency. To appear in L. Sánchez, & C. Bardel (Eds.) Eurosla monograph *Multilingualism, age and proficiency*.
- Sánchez, L., & Bardel, C. (2016). Cognitive factors, linguistic perceptions and transfer in third language learning. *Revue française de linguistique appliquée*, 21, 123-138.
- Sanz, C., Lin, H., Lado, B., Stafford, C., & Bowden, H. (2014). One size fits all? Learning conditions and working memory capacity in *ab initio* language development. *Applied Linguistics*, 37, 1-26.
- Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319-353). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524780.013>
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139524780.003>
- Schmidt, R. (2010). Attention, awareness and individual differences in language learning. In W. Chan, S. Chi, K. Cin, J. Itsanto, M. Nagami, J. Sew, T. Suthiwan, & I. Walker (Eds.), *Proceedings of CLAsIC* (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.
- Serafini, E., & Sanz, C. (2015). Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Studies in Second Language Acquisition*, 37, 1-40.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30, 510-532. <https://doi.org/10.1093/applin/amp047>

- Skehan, P. (2015). Limited attention capacity and cognition. Two hypotheses regarding second language performance on tasks. In M. Bygate (Eds.), *Domains and directions in the development of TBLT: A decade of plenaries* (pp. 123-155). Amsterdam: John Benjamins. <https://doi.org/10.1075/tblt.8>
- Skehan, P. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In G. Granena, D. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in L2 processing and acquisition*. Amsterdam: John Benjamins.
- Singleton, D. (2017). Language aptitude: desirable trait or acquirable attribute? *Studies in second language learning and teaching*, 17, 89-103. <https://doi.org/10.14746/ssllt.2017.7.1.5>
- Stewart, C. (2002). Memory. In E. Perry, H. Ashton, & A. Young (Eds.), *Neurochemistry of consciousness* (pp. 65-81). Amsterdam: John Benjamins. <https://doi.org/10.1075/aicr.36.08ste>
- Tagarelli, K., Ruiz, S., Moreno, J.L., & Rebuschat, P. (2016). Variability in second language learning. The roles of individual differences, learning conditions and linguistic complexity. *Studies in second language acquisition*, 38, 293-316. <https://doi.org/10.1017/S0272263116000036>
- Todeva, E., & Cenoz J. (2009). Multilingualism: Emic and etic perspectives. In E. Todeva, & J. Cenoz (Eds.), *The multiple realities of multilingualism* (pp. 1-32). Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110224481.1>
- Truscott, J. (1998). Noticing in second language acquisition: a critical review. *Second Language Research*, 14, 103-135.
- Wen, Z. (2016). *Working memory and second language learning. Towards an integrated approach*. Clevedon: Multilingual Matters.
- Wen, Z., Biedrón, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language teaching*, 50, 1-31. <https://doi.org/10.1017/S0261444816000276>
- Weissheimer, J., & Mota, M., (2011). Working memory capacity and the development of L2 speech production: A study of individual differences. In G. Granena et al. (Eds.) *Selected Proceedings of the 2010 Second Language Research Forum* (pp. 169-181). Somerville, MA: Cascadilla Press.
- Weissheimer, J., & Borges Mota, M., (2009). Working memory capacity and the development of L2 speech production: a study of individual differences. In G. Granena, Koeth, J., Lee-Ellis, S., Lukyanchenko, A., Prieto, G., & Rhoades, E. (Eds.), *Selected Proceedings of the 2010 Second Language Research Forum* (pp. 169-181). Somerville: Cascadilla Proceedings Project.
- Weissheimer, J., & Borges Mota, M. (2011). Working memory capacity and the development of L2 speech production: A study of individual differences. In G. Granena et al. (Eds.), *Selected Proceedings of the 2010 Second Language Research Forum* (pp. 169-181). Somerville, MA: Cascadilla Press.
- Wolfe-Quintero, K., Inagaki, S. & Kim, H.Y. (1998). *Second language development in writing. Measures of fluency, accuracy, and complexity*. Honolulu: University of Hawaii, Second language teaching & curriculum center.
- Young, B., & Pigott, S. (1999). Neurobiological basis of consciousness. *Archives of Neurology*, 56, 153-157.
- Zalbidea, J. (2017). 'One task fits all'? The role of task complexity, modality, and working memory capacity in L2 performance. *The modern language journal*, 101, 335-352. <https://doi.org/10.1111/modl.12389>

Note 1. Within the 'Task-Based Language Learning' framework, this theoretical stance is known as the trade-off hypothesis (or more recently, the 'limited attentional capacity' hypothesis). This stance is in contrast with Robinson's cognition hypothesis, according to which the learner's performance is pushed in the three dimensions (especially accuracy and complexity) if the task is cognitively demanding. This is mentioned here merely for the sake of preciseness and clarity, but testing these two hypotheses is out of the scope of the present paper.

## Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).