

# **The Influence of Working Memory on L2 Oral Fluency: An Exploratory Study**

## **Abstract**

The goal of this exploratory study is to investigate the effect of working memory capacity on L2 oral fluency in 79 learners of English as a foreign language. Three tasks were used as measures of working memory (the reading span task, letter span task and an attention-switching task). Twelve measures of fluency were used spanning across speed, breakdown and repair fluency. Positive correlations were found with measures of repair fluency, specifically morphosyntactic, difference, and other repairs whereas negative correlations were found for lexical repairs. When participants were divided into groups based on proficiency, potential relationships were found between working memory and speed/breakdown fluency suggesting the possible existence of proficiency thresholds affecting the relationship between working memory and fluency. The results are discussed in light of previous research and De Bot's (1992) model of L2 speech production.

## 1. INTRODUCTION

The cognitive revolution in Psychology has had a huge impact on the field of applied linguistics. The exploration of how individual differences in cognitive abilities can affect second language (L2) attainment and performance is gaining momentum as important predictors of L2 acquisition are being found (Sawyer & Ranta, 2001). The goal of this paper is to investigate the role of working memory (WM) capacity on L2 oral fluency. The introduction outlines the main concepts and existing research into the area. Subsequently, the study itself will be described and the results will be presented and discussed in relation to the studies' research questions, previous studies and future directions in this field.

### *1.1. What is L2 fluency?*

In order to measure something correctly it is necessary to first define what exactly it is that is being measured. The way in which fluency is measured depends on how it is conceptualised and defined. There exists a plethora of definitions and measures of fluency. People may use it to refer to global L2 proficiency, to be able to speak publicly in an L2 or to express thoughts and opinions quickly (Fillmore, 1979). The term is also associated with the metaphorical movement-like aspects of spoken language such as how well language flows or how fluid it is (Segalowitz, 2010). These are considered broad, qualitative definitions of fluency (Kormos, 2006). However, it is the narrow sense of fluency which the present study will focus on.

Narrow definitions regard fluency as just one aspect of performance (Kormos, 2006). Lennon (1990) states that fluent speech is speech in which the listener receives the impression that the underlying psycholinguistic processes responsible are working effortlessly. This definition alludes to fluency as a listener's impression of efficiency in terms of a speaker's cognitive processes. Schmidt (1992) stated that fluency also relates to the cognitive processes themselves and that fluent speech emanates from automaticity and proceduralization. This led to Lennon re-defining fluency as "rapid, smooth, accurate, lucid and efficient translation of thought or communicative intention into language under the temporal constraints of on-line processing" (2000: 26).

Segalowitz (2010) outlined a new framework for the conceptualization of fluency pertaining 3 types of fluency; cognitive, utterance and perceived fluency (see figure 1). This 3-way distinction advises that these 3 types be thought of separately. Previous conceptions of L2 fluency have been imprecise and do not reflect the multi-dimensional nature of fluency as a construct. These three definitions provided by Segalowitz will be

described below.

Cognitive fluency refers to the extent to which the underlying cognitive processes which mediate L2 speech are working efficiently and rapidly. Furthermore, it refers to the co-ordination and integration of the processes by the speaker and the extent to which there exists internal interference or crosstalk between them. It is believed that high levels of efficiency and speed paired with low levels of interference promote greater fluency. These mechanisms responsible for speech are described in section 1.2.

Utterance fluency refers to the measurable features of speech which indicate fluency. Features used to measure utterance fluency include speed, hesitation and repair phenomena. There are many potential features of fluency which can be found in speech and a number of ways to operationalize each one (which will be reviewed in section 1.3. - *Measures of fluency*).

Perceived fluency refers to an impression received by the listener of a speaker's cognitive fluency, a judgement of their cognitive processes based on the utterance fluency of their speech. Of course, as this is an 'on-line' impression the listener may take into account different factors when judging how features of the utterance affect their fluency.

The current study will be focusing on and measuring utterance fluency. With fluency defined, this review turns to how individual differences in oral fluency can be explained using speech production models.

### *1.2. Models of L2 speech and how fluency is manifest*

De Bot's 1992 adaptation of Levelt's 1989 blueprint of first language (L1) speech production is widely cited by L2 researchers as an accurate model of the bilingual speaker (Segalowitz, 2010). The model outlines the act of speaking step by step. De Bot's model (1992) is presented below along with the 7 points at which fluency can be affected according to Segalowitz.

To begin with, the model assumes that there are three separate encoding modules; the conceptualizer, the formulator and the articulator (Kormos, 2006). Furthermore there are three knowledge stores; encyclopaedic knowledge, the mental lexicon and the syllabary (Kormos, 2006). Speech begins with the conceptualizer and with macroplanning, where the speaker plans what is to be said. During macroplanning, the speaker uses information from their encyclopaedic knowledge of the outside world and knowledge of their interlocutor's thoughts and beliefs. Here, speech register and language choice are also selected. The knowledge used for macroplanning is not language-specific and as such does not affect L2 fluency. The next step is

microplanning which is the preparation of speech and the process which leads to a preverbal message. While still being based on concepts, microplanning is much more specific than macroplanning because it only deals with concepts which can be lexicalized. The output of microplanning is the preverbal message. This is a conceptual code which can but has not yet been lexicalized. The preverbal message takes into account the speaker's own perspective of the event. Microplanning is the first potential source of L2 dysfluency because in forming the preverbal message the speaker must only use the concepts which they know they can express in the L2. L2 knowledge limitations will therefore lead to a slowing down of this process and thus, dysfluency – despite the case being that formation of the preverbal message is not language-specific.

Grammatical encoding, or the formulator, is the next step. Here the preverbal message is encoded into appropriate words, it is given linguistic shape. Grammatical encoding entails which words are to be used but also how they relate to each other to express the perspective of the speaker. This process requires the use of information from another knowledge store, the mental lexicon. The mental lexicon contains representations of all the different lemmas known in a language and their meanings. Syntactic information relating to the lemma is also called upon when it is selected and these syntactic characteristics are also stored in the mental lexicon. Here the second source of potential dysfluency is encountered, where the speaker may face difficulties in using or retrieving information from the mental lexicon because the system may be incomplete in the L2.

Grammatical encoding leads to formation of the surface structure, a verbalized plan. This must then be converted into overt or 'real' speech in the articulator. It begins with morpho-phonological encoding which utilises morpho-phonological codes. These are stored together with lemmas in the mental lexicon. Again, difficulties in retrieval of information from the mental lexicon or when this retrieval is not automatized can lead to dysfluency in L2 speech (the fourth such point). The product of morpho-phonological encoding is the phonological score which is used for overt speech but firstly needs to be converted into an articulatory score. The articulatory score gives the articulatory system (speech apparatus) specific instructions in order to produce overt speech. Here, information is taken from the syllabary. The syllabary is a separate knowledge source which contains information on how to produce the specific motor responses which produce certain speech sounds, known as gestural scores. Here lies the fifth area for potential L2 fluency, occurring if the speaker does not select gestural scores automatically. Overt speech is the result of execution of the gestural scores relating to the articulatory score. Here there is also potential for dysfluency if the speaker does not

carry out articulation of gestural scores automatically (the sixth such point).

Monitoring is where the speaker analyses what has been produced in the speech process. Kormos (2006) states that monitoring occurs in L2 speech just as it does in L1 speech. She outlines three monitor loops. Firstly, after the pre-verbal message is produced, the speaker will compare it to the original concept produced during macroplanning. The second loop is the comparison of the phonological score with the speaker's intentions. Finally, overt, parsed speech is monitored against a speaker's original intentions. This monitoring system is the seventh and final point for potential dysfluency. Kormos (2006) proposed the bilingual speech production model which differs in some ways to the Levelt/de Bot model. For example, it is proposed that there is just one large memory store, long-term memory (LTM), which consists of sub-components analogous to the knowledge stores proposed by Levelt/de Bot but which includes an L2-specific store for declarative rules relating to L2 syntax and phonology.

The notion of automaticity in the underlying cognitive processes which carry out speech is a persistent factor mentioned above in explaining L2 oral fluency. According to Segalowitz (2010), automaticity is a difficult notion to define and is often left poorly operationalized but which in general refers to a process which is functioning at a high level of efficiency. Unfortunately, explanations and definitions of the nature of automaticity are sparse and the area is under researched. One theory is that automaticity can be defined as ballistic processing, the idea that once the specific process has begun it cannot be stopped (Favreau & Segalowitz, 1983).

Here it is worth noting that a key goal of this study is to measure participants' utterance fluency through analysing the speech produced. Utterance fluency will relate to the performance and automaticity of the underlying system at the seven potential 'points of fluency vulnerability' and can be considered as being closely linked with cognitive fluency. L2 fluency is probably affected by many variables, which is why the dynamic systems approach described below may be extremely important to understanding L2 fluency.

At the beginning of the text it was suggested that in order to measure something it must firstly be defined. The current study defines fluency in the narrow sense, as the efficiency, speed and degree of automaticity of the underlying cognitive processes responsible for L2 speech found in de Bot's speech production model. It will take Segalowitz's (2010) idea of utterance fluency as a working definition of and way of measuring fluency. Thus, now fluency has been appropriately defined, measurements of it can now be identified.



It has been shown that some common fluency measures can also be indicative of global proficiency as well as oral fluency. Iwashita et al. (2008) compared L2 oral fluency (measured using 6 measures of L2 fluency) with the global proficiency ratings given to participants by raters. Oral productions were elicited using 5 oral production tasks. 6 measures of L2 fluency were used. They found that unfilled pauses, total pause time and speech rate had a significant impact on global proficiency ratings but filled pauses, repair and mean length of run did not.

To summarise the above it can be seen that there are broad and narrow definitions of fluency, explanations of differences in L2 oral fluency based on L2 speech models and many methods of measuring oral fluency. One of the main goals of the current study is to measure the utterance fluency of participants' L2 speech. This will be carried out using a broad range of fluency measures spanning across Skehan's 3 categories; breakdown, repair and speed fluency. With fluency thoroughly reviewed, the second key part of the current study is now to be described, working memory, a potential predictor L2 oral fluency.

#### *1.4. Working memory*

Working memory (WM) is the name given to the system in the brain which carries out the temporary storage of information and the dynamic processing and manipulation of this information. It is essential for the completion of higher-order cognitive tasks such as reasoning, learning and language comprehension (Baddeley, 2000). WM is seen to be at the very centre of cognition and all mental processing (Mota, 2003). The original model of WM proposed by Baddeley and Hitch (1974, cited in Baddeley, 2000) continues to be a widely accepted model of WM that integrates the limited capacity information storage aspect of previous models (e.g. Atkinson & Schiffrin's (1968) model of short-term memory) with dynamic processing (Juffs & Harrington, 2011). Since the first model was proposed in 1974 it has been subject to a number of revisions. Here the model as revised by Baddeley (2000) is outlined.

WM is a multi-component system. The central executive is responsible for attentional control and it coordinates two subsidiary slave systems. The central executive is the most important and central component of WM. Its subsidiary systems are the visuo-spatial sketchpad which holds visual and spatial information and the phonological loop which holds verbal and acoustic information (thus they are modality-specific). The phonological loop holds chunks of information for a few seconds. This information can be refreshed using the articulatory rehearsal process. This is the sub-vocal repetition of information used for example when trying to remember a phone number. This refreshes

information and avoids decay (Kormos & Sáfár, 2008). The original 1974 model of WM consists of these 3-components. Subsequently, the episodic buffer has been integrated into the model (Baddeley, 2000). The buffer is assumed to be another slave component controlled by the central executive but which is not modality-specific. It temporarily stores information from the other subsidiary components and serves as an intermediary between them and long-term memory (LTM). It binds information from WM to episodic LTM. Episodic LTM refers to conscious, declarative memories related to specific events (Baddeley, 2001). The introduction of the episodic buffer helps explain how WM can represent and store modality-specific information and how this information can be bound to LTM, something which was lacking from the initial model (Baddeley, 2000). As mentioned in the opening paragraph of this study, individual differences in cognitive abilities such as WM and PM can affect L2 acquisition and performance.

#### *1.5. Individual differences and measures of phonological short-term memory/working memory and how they relate to language learning*

The phonological loop has perhaps been the most extensively researched component of Baddeley and Hitch's model of WM (Kormos & Sáfár, 2008). As it is a component of WM it is important to this study. Phonological short-term memory (PSTM) capacity has been shown to have a strong relationship with various aspects of language learning for example; L1 lexical acquisition (Gathercole et al., 1999); L2 lexical access and language aptitude (Kormos & Sáfár, 2008); and function word and subordinate clause usage (O'Brien et al., 2006). Simple span tasks such as the letter span task and the word span task are designed to measure the storage capacity of PSTM (Engle et al., 1999).

WM involves not only the storage of information but also the dynamic processing of information central to many cognitive processes (Baddeley, 2000). Complex tasks which tap not only storage but processing as well can be used to measure individual differences in WM. The storage of to be remembered (TBR) items is interspersed with some kind of mental operation. A number of complex span tasks including the reading span task (where participants are given various sentences and instructed to say if the sentence makes sense and then remember the final word) have been shown to be valid and reliable measures of WM along with the counting span, operation span and backward digit span tasks (Conway et al., 2005; Kormos & Trebits, 2008). WM and PSTM have been shown to be distinct constructs which are highly related to each other (Engle et al., 1999). Essential to memory testing is that tests are administered in the participants' L1 to remove L2 proficiency as a confound (Gilbert & Muñoz, 2010).

Individual differences in WM capacity have also been shown to be influential in many aspects of L2 learning. Examples include; overall L2 proficiency and 3 major skills of reading, listening and speaking (Kormos & Sáfár, 2008); acquisition of morpho-syntax (French & O'Brien, 2008); and lexical complexity (Gilabert & Muñoz, 2010).

Mackey et al. (2010) suggest that increased WM capacity affords language learners the sufficient cognitive resources to reflect on their output whilst speaking. They found that increased WM capacity was significantly correlated to modified output in L2 speech i.e. reformulations and restarts. Inevitably WM has been considered to be one of the main predictors of language learning success (Juffs & Harrington, 2011).

Correlations have been found between WM and language learning aptitude (Kormos & Sáfár, 2008; Hummel, 2009) and between WM and L2 proficiency (Hummel, 2009).

In summary, it has been shown that there are many measures of PSTM and WM and that they are related to many aspects of L2 learning (see Juffs & Harrington, 2011 for a complete review). However, it is oral production and specifically L2 fluency that will be the focus of attention in this study.

#### *1.6. L2 fluency and working memory*

A number of studies have investigated a possible link between working memory and L2 oral fluency. Research within the area is scarce and no substantial body of evidence has accumulated for or against this link. Moreover, much of the research is limited by methodological problems.

Gilabert and Muñoz (2010) examined the influence of WM capacity on proficiency and performance in foreign language acquisition. It was found that WM capacity and L2 fluency were significantly, positively correlated, although a fairly modest correlation was found (.231). When the groups were split into high-proficiency and low-proficiency groups this correlation disappeared although this may have been because of decreased group size. WM capacity was measured using a reading span task. L2 fluency was measured by calculating unpruned speech rate (syllables per minute including all the utterances such as repetitions, self-corrections and false starts). Two limitations of this study are that analyses were correlational (thus non-inferential) and fluency measurement was limited to just speed fluency.

Mota (2003) too found some moderate correlations between WM capacity (as measured by the speaking span test) and L2 oral fluency using measures of speed fluency; filled/unfilled pauses, and mean length of run for 2 narratives. Filled and unfilled pauses however were not significantly correlated with WM. Using a simple linear regression analysis, WM capacity was found to account for 53%, 52% and 49%

of speech rate unpruned, speech rate pruned and mean length of run respectively. This suggests that there is a link between WM and fluency. However, this study has severe limitations in that it used a very small sample ( $n=13$ ) and proficiency in English was regarded as being homogenous by the author (advanced) without it being measured and was not considered a potential covariate. Moreover, the use of the speaking span test as a valid measure of WM capacity is questionable as the task requires spoken production of sentences which have to be generated by the participants themselves. This is essential to test performance and therefore reativity is bound to be a confounding variable. Given that the test itself was administered in the L2, English L2 proficiency will also be a confound.

Mizera (2006) found scant evidence for a relationship between WM capacity and L2 oral fluency. WM was measured using three measures; the speaking span test, the math span test and the non-word repetition test. A wide range of fluency measures were carried out measuring speed, breakdown, repetitions and morphosyntactic accuracy. Morphosyntactic accuracy is a measure normally associated with accuracy not fluency. It was selected on the basis that it correlated with holistic judgements of fluency by raters in the first part of the study. Three working memory tests were also carried out. Of the thirty-five correlations only three were significant and the correlations found were very weak. The same pattern was found when only low-proficiency participants were analysed. The study used measures of many aspects of fluency which is rarely carried out and can be considered a huge positive. However, a limitation of the study is that the sample size used was fairly small ( $n=44$ ), the speaking span test was used (see criticisms of Mota, 2003) and only non-inferential correlational analyses were performed.

Weissheimer and Mota (2009) found no overall correlations between working memory and L2 fluency. When the sample was split into a high and low group based on WM score (measured using a speaking span test) a significant correlation was found for the lower-WM group but not for the higher-WM group. This suggests that the effect of WM on fluency may differ depending on capacity of WM itself. However, the higher-span group was very small ( $n=8$ ) and were heterogeneous when compared to lower-WM group which may have resulted in a non-significant correlation. Moreover, the complete sample used was small ( $n=32$ ). Again, the study only operationalized fluency in terms of speed (pruned and unpruned speech rate), used the speaking span test in the participants' L2, and used correlational analysis only.

Kormos and Trebits (2011) found that WM capacity as measured by the backward

digit span task had no effect on oral fluency in two narrative tasks. Using a MANOVA test the study compared the fluency of four groups with differing WM capacities. The study however had many limitations; only one measure of fluency was used (unpruned speech rate), only one measure of WM was used, and the sample was limited in size and homogenous in terms of proficiency. The possibility that WM may affect fluency depending on the proficiency level of the individual is an area the current study aims to explore. A positive aspect of the study is the use of the MANOVA as an inferential statistical tool where many of the above studies rely solely on correlational analysis.

### *1.7. How might working memory affect fluency?*

At this point, it is important to provide an explanation as to how L2 fluency may be influenced by WM. Current explanations are rooted in de Bot's model of the L2 speaker process. There is the potential for greater or lesser fluency at seven points in de Bot's model (see section 1.2) according to Segalowitz (2010). Higher WM capacity could increase the extent to which L2 speech processes are automatized and proceduralized. Higher automaticity and proceduralization at these seven points will mean that the system relies less on WM itself during the formulation of speech (Mota, 2003) which would lead to greater speed and efficiency in the speech process (that which Segalowitz describes as cognitive fluency). This speed and efficiency would be manifest as greater utterance fluency - that which is measured.

A second theory relates to Sawyer and Ranta's (2001) suggestion that higher WM capacity allows more WM attentional resources to be freed up. This could also mean that throughout the learning process more cognitive resources can be dedicated to the learning of speech strategies and the rules related to production (Weissheimer & Mota, 2009). The above ideas are hypotheses and are speculative in nature.

### *1.8. Objectives and research questions*

The current study will look further into this relationship between WM and L2 oral fluency. The majority of existing research has used just one or two measures of fluency (usually speech rate, an all-round measure of fluency) or has used questionable measures of WM capacity. These are methodological issues which the current study aims to overcome. The objectives of this study are; (a) to use a wide range of fluency measures (twelve in total) spanning across the three categories outlined by Skehan (2009); speed, breakdown and repair fluency; (b) to use a wide range of valid measures of WM; the reading span task, the letter span task, and the attention switching task; (c) to use inferential statistical analysis as well correlational and; (d) to sub-divide participants based on L2 proficiency and WM capacity in order to further explore the

complexities of any relationship between L2 fluency and WM capacity. The study is exploratory in nature and will produce a large amount of data which can be used to guide further research into this specific area.

The research questions are the following:-

- 1) Does working memory capacity as measured by the reading span task, the letter span task and the attention-switching task affect L2 oral fluency?
- 2) Does L2 proficiency affect the relationship between working memory and fluency?
- 3) Does the relationship between working memory and fluency differ between measures of breakdown fluency, speed fluency and repair fluency?

## **2. METHOD**

### *2.1. Participants*

79 participants took part in this study. Their mean age was 21.2. They were taken from a pool of participants from the Age, Input and Aptitude in foreign language learners project (GRAL group, University of Barcelona). They were mainly L1 Spanish/Catalan speakers (89% - all other L1s 11%) Their proficiency in English ranged between intermediate and upper-intermediate.

### *2.2. Tests and Tasks*

L2 English proficiency was measured using the Oxford Placement Test (OPT), a standardized general proficiency test. Mean = 44.7 (1dp), Standard deviation = 6.4 (1dp).

The automated reading span task was used as developed by the GRAL group (Spanish or Catalan) to test both reading span (RS) and letter span (LS). The test has been to have high internal consistency (Cronbach's Alpha .755 (RS) and .787 (LS)). The LS task is considered a simple memory task. It is a test of the information storage aspect of working memory without requiring dynamic information processing. The reading span task (RS) is a complex memory task. This type of task taps into both the storage and the dynamic information manipulation aspects of WM. It can be expected that these tests will correlate weakly as they measure some of the same components of WM (storage) but not all (information processing).

In the automated span task participants have to take three practice tests followed by the actual experiment. In the first test participants are asked to memorize a series of letters ranging from 3 to 9. The scores obtained by means of this practice task constitute

the “letter span score”. The second test asks participants to rate sentences as either making sense or not. The test also serves the purpose of automatically measuring the participants’ mean reaction time in order present sentences in accordance with that mean (i.e. participants with longer reactions times during the practice test will observe the sentences on the screen for longer than participants with lower reactions times during the actual experiment). In the third practice test, participants are asked to practice rating sentences and memorizing letters simultaneously. This is followed by the actual experiment. It does this by giving participants to-be-remembered items and concurrently a distractor task requiring information processing (i.e. to decide whether the sentences make sense or not). 81 sentences in Spanish or Catalan were used here which were between 12-15 words long. They were organized into 15 sets ranging from 3 to 9 sentences. Participants were a) given a sentence and asked to say if it made sense or not and b) asked to remember a letter given to them at the end. The final score is automatically calculated by the test and it factors in accuracy, order of the recalled letters and reaction times.

The attention-switching (AS) task (often referred to as the Trail Making Test) is a measure of attentional control and requires use of the executive control component of WM, but not information storage. Therefore it is expected to correlate only weakly with the RS and not at all with the LS. The task used was made up of two parts which were timed. The first part, participants joined up consecutive numbers (from 1–25) which were scrambled across a sheet of paper. The second task involved joining up numbers and letters (from 1-13 and from A-L), alternating between the two i.e. A-1-B-2-C-3 etc. which were again scrambled. Errors were pointed out by the experimenter and the time taken to correct errors was included. The time taken to complete the first part was taken as a baseline. This was subtracted from the time taken to complete the second task. Both tasks require visual search, visual perceptual ability and motor skills but only the second task requires task shifting and working memory. Therefore, that which remains was considered a measure of executive control functioning, the ‘time cost’ of switching between letters and numbers (Arbuthnott and Frank, 2000).

### *2.3. L2 speech elicitation*

The participants were given a film-retelling task in order to elicit speech in the L2, English. Participants were asked to view two short excerpts from the Charlie Chaplin film, ‘Modern Times’ twice. The excerpt lasted 8 minutes 30 seconds and can be found on this link (available July 2012) [http://www.youtube.com/watch?v=5njTU7a3\\_\\_g](http://www.youtube.com/watch?v=5njTU7a3__g) between 0:50-9:20.

After seeing the first half of the excerpt twice participants were asked, “Can you tell me what happened in the first part of the story?” After the second half they were asked, “Can you tell me what happened in the second part of the story?” Responses were all in English except from asides in L1 Spanish/Catalan. They were recorded using a microphone and a sound recorder. The length of the combined narratives varied from 49 – 564 seconds (mean=273.7s, s.d.=106.7s). The narratives were transcribed using computerized language analysis (CLAN) software. Transcriptions followed the conventions and principles of the CHAT transcription system. The transcriptions were then analysed using measures of oral fluency.

#### *2.4. Measures of fluency*

Twelve measures of fluency were used. As all of these measures are sensitive to length of speech sample, they were divided by total speech time (seconds) and multiplied by 60 to give results per minute.

Speed fluency was measured using two speech rates. Unpruned speech rate (total number of syllables/minute) and pruned speech rate (total number of syllables produced excluding repetitions, repairs and restarts/minute) were calculated (taken from Mora and Valls, in press).

Repair fluency was measured using seven measures (taken from Mora and Valls, in press) Dysfluency ratio was used as a global measure of repair fluency (dysfluencies/minute). Dysfluencies are classified as; repetitions (where one word or groups of words appearing next to each other are repeated by the speaker), repairs (where the speaker corrects themselves) or restarts (where the speaker leaves a sentence unfinished and begins a new one). Repetitions/minute, repairs/minute and restarts/minute were also calculated as measures of repair fluency.

Repairs were sub-classified into three categories (as in Gilabert, 2007); lexical repairs (where the repaired word used is judged to be incorrect at the lexical level e.g. her→him), morphosyntactic repairs (where the repaired word used is part of a syntactic structure e.g. works→worked), and other repairs. Other repairs include; appropriacy repairs (where inappropriate or inadequate information is judged to have been supplied e.g. there is a man→a gentleman) and different repairs (where the original speech plan is changed and different information is encoded).

Breakdown fluency was measured using filled pauses/minute (such as “umm” or “erm”). Filled pauses were sub-classified as intra-clausal filled pauses/minute (pauses within clauses) and inter-clausal filled pauses/minute (pauses at the boundary of two different clauses). Clauses were operationally defined as a grammatical unit consisting

of a verb plus any dependent elements linked to it. Thus, a subject, verb, object and adverbial would count together as a single clause (as in Mora and Valls-Ferrer, in press).

### **3. RESULTS**

Descriptive statistics for the group as a whole can be found in Appendix #1.

The three measures of WM were correlated with each other in order to gain insight into the extent to which they are measuring the same ability. The reading span (RS) and letter span (LS) tasks showed a weak, positive correlation (.257,  $p=.020$ ). The RS and attention span (AS), and the LS and AS tasks did not correlate significantly. This suggests that the assumption made that each test was measuring different abilities can be confirmed, with some overlap between the RS and LS tasks (in line with Engle et al., 1999).

The first research question was addressed i.e. does working memory capacity affect L2 oral fluency. The relationship between fluency and WM was explored using partial correlations in which the effect of L2 proficiency was removed as an extraneous variable (OPT scores were controlled for). These correlations can be found in Table 1.

As can be seen, WM capacity as measured by the RS and LS tasks correlated positively and albeit weakly with ‘morphosyntactic repairs’ and the same was found for ‘other repairs’ and performance on the LS task. Correlations between dysfluencies/min and the LS task were nearly significant. These results suggest that those with higher WM-capacity tend to make more morphosyntactic and other types of repair. Overall there are few significant relationships between fluency and the measures used here.

In order to further explore research question 1 the groups for each WM test were subdivided into two groups, one higher-capacity and one lower-capacity to see if closer analysis of the data would produce more findings. Partial correlations, again controlling for OPT scores were run once more for the lower-capacity and higher-capacity reading span groups (see Table 2).

The groups were split using k-means clustering (RS - lower-capacity N=31, higher-capacity N=48; LS – lower-capacity N=60, higher-capacity N=5; AS – lower-capacity N=65, higher-capacity=11). For LS and AS, some of the groups created were too small. Therefore a median split was utilised, where the median value is taken as a cut-off point (LS – lower-capacity=31, higher-capacity N=30, AS lower-capacity N=37, higher-capacity N=37).

Table 1: *Partial correlations (controlling for OPT score) of three WM measures and twelve fluency measures*

	<b>Reading Span Task (N=79)</b>	<b>Letter Span Task (N=65)</b>	<b>Attention Span Task (N=77)</b>
<b>Rate A</b>	.035	.025	-.056
Sig (2-tailed)	.381	.423	.316
<b>Rate B</b>	.030	.008	-.084
Sig (2-tailed)	.399	.476	.238
<b>Fp/min</b>	-.025	-0.11	-.019
Sig (2-tailed)	.415	.466	.437
<b>Inter-clausal fp/min</b>	.037	.036	.111
Sig (2-tailed)	.373	.390	.172
<b>Intra-clausal fp/min</b>	-0.058	-0.38	-0.88
Sig (2-tailed)	.308	.384	.227
<b>Repairs/min</b>	-.030	.146	.036
Sig (2-tailed)	.398	.125	.380
<b>Lexical repairs/min</b>	-.132	-0.76	-.056
Sig (2-tailed)	.124	.275	.316
<b>Morphosyntactic repairs/min</b>	<b>.195*</b>	<b>.273*</b>	.147
Sig (2-tailed)	.044	.014	.104
<b>Other repairs/min</b>	-.010	<b>.213*</b>	-.053
Sig (2-tailed)	.467	.045	.325
<b>Repetitions/min</b>	.097	.016	.087
Sig (2-tailed)	.199	.451	.229
<b>Restarts/min</b>	.022	-.015	.164
Sig (2-tailed)	.423	.453	.080
<b>Dysfluencies/min</b>	.078	<b>.343</b>	.101
Sig (2-tailed)	.248	.051	.194

\*p < 0.05, \*\*p < 0.01, **bold type** = approaching significance, fp=filled pauses

Table 2. *Partial correlations (controlling for OPT score) of fluency measures and WM measures divided into low- and high-proficiency categories*

	<b>Lower- capacity RS group (N=31)</b>	<b>Higher- capacity RS group (N=48)</b>	<b>Lower- capacity LS group (N=31)</b>	<b>Higher- capacity LS group (N=30)</b>	<b>Lower- capacity AS group (N=37)</b>	<b>Higher- capacity AS group (N=37)</b>
<b>Rate A</b>	.073	-.148	-.073	.144	.116	.168
Sig (2-tailed)	.350	.161	.341	.212	.251	.161
<b>Rate B</b>	-0.25	-.109	-.104	.174	.087	.114
Sig (2-tailed)	.447	.234	.279	.167	.307	.252
<b>Fp/min</b>	.178	.197	-.071	-.087	-.128	.041
Sig (2-tailed)	.173	-.127	.463	.315	.229	.404
<b>Inter-clausal fp/min</b>	.236	-.014	-.017	-.095	.041	.095
Sig (2-tailed)	.105	.462	.463	.299	.406	.288
<b>Intra-clausal fp/min</b>	.100	-.155	-.085	-.065	-.190	.003
Sig (2-tailed)	.299	.149	.317	.359	.133	.494
<b>Repairs/min</b>	<b>.310*</b>	-.57	.186	-.278	.198	-.085
Sig (2-tailed)	.048	.353	.146	.059	.123	.308
<b>Lexical repairs/min</b>	.141	-.160	.110	<b>-.385*</b>	.120	-.018
Sig (2-tailed)	.228	.141	.267	.013	.242	.457
<b>Morphosyntactic repairs/min</b>	<b>.280</b>	-.16	<b>.249</b>	-.088	.076	-.052
Sig (2-tailed)	.067	.459	.078	.313	.329	.379
<b>Other repairs/min</b>	<b>.374*</b>	<b>.239</b>	.055	.238	.144	-.077
Sig (2-tailed)	.021	.053	.379	.091	.202	.325
<b>Repetitions/min</b>	<b>.483**</b>	.069	-.029	-.050	.143	.018
Sig (2-tailed)	.003	.323	.435	.392	.203	.459
<b>Restarts/min</b>	<b>.281</b>	<b>-.336*</b>	-.029	.157	.030	.173
Sig (2-tailed)	.066	.011	.435	.191	.432	.153
<b>Dysfluencies/min</b>	<b>.518**</b>	.011	.026	-.083	.184	.016
Sig (2-tailed)	.002	.470	.442	.323	.142	.464

\*p<0.05, \*\*p<0.01, **bold type**=significant/approaching significance, fp=filled pauses

Beginning with RS, the lower-capacity group showed positive significant correlations for repairs, other repairs, repetition and dysfluencies. Morphosyntactic repairs and

restarts showed a string trend. For higher-capacity RS a significant negative correlation was found with restarts and other repairs reached positive, near-significance. Together, it seems to suggest that repair fluency decreases as WM increases with the exception of restarts which showed the opposite effect. No other significant results were found.

For the lower-capacity LS, near significant, positive correlation was found between morphosyntactic repairs and LS score. In the higher-capacity group a negative, medium-strength correlation was found for lexical repairs. This second result seems to contradict the above findings for RS. However, these two tests are measuring different things.

No significant results were found for AS.

Inferential methods of analysis were then used. The higher- and lower- WM capacity groups were compared for each measure of fluency in order to see if having greater WM inferred greater oral fluency (research question 1). This was carried out using the independent samples T-test (for normally distributed fluency measures) and the Mann-Whitney U-test (for non-normally distributed fluency measures). Normality was assessed using the Kolmogorov-Smirnov normality test. The findings are displayed in Tables 3, 4 and 5.

Table 3. *Mann-Whitney U-tests and independent samples T-tests comparing higher- and lower-capacity reading span task groups (2-split) on twelve fluency measures*

	<b>Mann-Whitney U-test</b>	<b>Independent samples T-test</b>
<b>Rate A</b>	.129	X
<b>Rate B</b>	X	.352
<b>Fp/min</b>	X	.687
<b>Inter-clausal fp/min</b>	.514	X
<b>Intra-clausal fp/min</b>	X	.687
<b>Repairs/min</b>	X	.312
<b>Lexical repairs/min</b>	.491	X
<b>Morphosyntactic repairs/min</b>	.115	X
<b>Other repairs/min</b>	.096	X
<b>Repetitions/min</b>	X	.831
<b>Restarts/min</b>	.338	X
<b>Dysfluencies/min</b>	X	.729

fp=filled pauses

Table 3 shows that for RS scores there were no significant difference between the higher- and lower-WM capacity groups in any of the fluency measures.

Table 4. *Mann-Whitney U-tests and independent samples T-tests comparing higher- and lower-capacity listening span task groups on twelve fluency measures*

	Mann-Whitney U-test	Independent samples T-test
Rate A	.784	X
Rate B	X	.759
Fp/min	X	.986
Inter-clausal fp/min	.901	X
Intra-clausal fp/min	X	.901
Repairs/min	X	.261
Lexical repairs/min	.971	X
Morphosyntactic repairs/min	<b>.021*</b>	X
Other repairs/min	<b>.090</b>	X
Repetitions/min	X	.573
Restarts/min	.744	X
Dysfluencies/min	X	.441

\* $p < 0.05$ , **bold type** = approaching significance, fp=filled pauses

Table 4 shows that for LS scores there was a significant difference between the two groups for morphosyntactic repairs (repair fluency). Analysis of the means shows that the higher-capacity group tended to make more morphosyntactic repairs.

Table 5 shows that for AS scores there were no significant difference between the higher- and lower-WM capacity groups in any of the fluency measures.

For the RS group, k-means clustering produced three groups well-balanced in size (low to high – 22, 35, 22). It was thought that comparing the highest and lowest groups would be of interest as there would be a clear distinction between the groups (as they were separated by the middle-capacity group). The same inferential tests were carried out. The results are found in Table 6. Again, there were no significant differences between the high and low-capacity groups.

Table 5. *Mann-Whitney U-tests and independent samples T-tests comparing higher- and lower-capacity attention-switching task groups on twelve fluency measures*

	<b>Mann-Whitney U-test</b>	<b>Independent samples T-test</b>
<b>Rate A</b>	.320	X
<b>Rate B</b>	X	.105
<b>Fp/min</b>	X	.990
<b>Inter-clausal fp/min</b>	.446	X
<b>Intra-clausal fp/min</b>	X	.687
<b>Repairs/min</b>	X	.785
<b>Lexical repairs/min</b>	.851	X
<b>Morphosyntactic repairs/min</b>	.352	X
<b>Other repairs/min</b>	.475	X
<b>Repetitions/min</b>	X	.556
<b>Restarts/min</b>	.389	X
<b>Dysfluencies/min</b>	X	.530

fp=filled pauses

Table 6. *Mann-Whitney U-tests and independent samples T-tests comparing higher- and lower-capacity reading span task groups (3-split) on twelve fluency measures*

	<b>Mann-Whitney U-test</b>	<b>Independent samples T-test</b>
<b>Rate A</b>	.453	X
<b>Rate B</b>	X	.854
<b>Fp/min</b>	X	.501
<b>Inter-clausal fp/min</b>	.372	X
<b>Intra-clausal fp/min</b>	X	.391
<b>Repairs/min</b>	X	.854
<b>Lexical repairs/min</b>	.851	X
<b>Morphosyntactic repairs/min</b>	.213	X
<b>Other repairs/min</b>	.847	X
<b>Repetitions/min</b>	X	.292
<b>Restarts/min</b>	.709	X

---

 fp=filled pause

In order to answer research question 3, whether proficiency and affects the relationship between fluency and WM the group was split by proficiency (measured by OPT score). K-means clustering was used to create two groups; lower-proficiency (N=53) and higher-proficiency (N=26). Partial correlations were again carried out for all three WM tests (controlling for OPT score). These can be found in tables 7, 8 and 9.

Looking at table 7 The higher-proficiency RS group have significant negative correlations for RS with filled pauses, intra-clausal filled pauses and near-significant negative correlations with all repairs and lexical repairs. The lower-proficiency group has significant, positive correlations for RS with rate A, repairs and morphosyntactic repairs. This suggests that higher-proficiency speakers use fewer filled pauses and less repairs than lower-proficiency speakers. However, lower-proficiency speakers might have greater speed fluency.

Table 8 shows that the higher-proficiency group have significant positive correlations for LS with repairs and morphosyntactic repairs, the opposite to the RS results. This suggests that as proficiency increases the amount of morphosyntactic repairs increases. The lower-proficiency group showed no significant correlations between LS and any of the fluency measures.

Table 9 shows that the higher-proficiency group had no significant correlations between AS and the fluency measures. For the lower-proficiency measures there was a strong, positive correlation between AS and morphosyntactic repairs and near-significance with restarts.

Table 7. *Correlations for higher- and lower-proficiency groups – RS and twelve fluency measures*

	<b>RS higher-proficiency group (N=53)</b>	<b>RS lower-proficiency group (N=26)</b>
<b>Rate A</b>	-0.73	<b>.310</b>
Sig (2-tailed)	.303	.066
<b>Rate B</b>	-0.58	.225
Sig (2-tailed)	.340	.139
<b>Fp/min</b>	<b>-.243*</b>	.276
Sig (2-tailed)	.041	.091
<b>Inter-clausal fp/min</b>	-.125	.262
Sig (2-tailed)	.188	.103
<b>Intra-clausal fp/min</b>	<b>-.245*</b>	.240
Sig (2-tailed)	.040	.124
<b>Repairs/min</b>	<b>-.215</b>	<b>.341*</b>
Sig (2-tailed)	.063	.048
<b>Lexical repairs/min</b>	<b>-.215</b>	.042
Sig (2-tailed)	.063	.422
<b>Morphosyntactic repairs/min</b>	.064	<b>.438*</b>
Sig (2-tailed)	.326	.014
<b>Other repairs/min</b>	-.056	.053
Sig (2-tailed)	.345	.400
<b>Repetitions/min</b>	.086	.125
Sig (2-tailed)	.273	.276
<b>Restarts/min</b>	-.048	.157
Sig (2-tailed)	.369	.227
<b>Dysfluencies/min</b>	.199	.215
Sig (2-tailed)	.078	.151

\*p&lt;0.05, fp=filled pauses

Table 8. *Correlations for higher- and lower-proficiency groups – LS and twelve fluency measures*

	<b>LS task higher-proficiency group (N=43)</b>	<b>LS task lower-proficiency group (N=22)</b>
<b>Rate A</b>	.023	-.147
Sig (2-tailed)	.443	.263
<b>Rate B</b>	.010	.107
Sig (2-tailed)	.474	.323
<b>Fp/min</b>	-.136	.109
Sig (2-tailed)	.195	.319
<b>Inter-clausal fp/min</b>	-.060	.086
Sig (2-tailed)	.354	.355
<b>Intra-clausal fp/min</b>	-.142	.109
Sig (2-tailed)	.185	.320
<b>Repairs/min</b>	.139	.135
Sig (2-tailed)	.190	.279
<b>Lexical repairs/min</b>	-.076	-.118
Sig (2-tailed)	.316	.305
<b>Morphosyntactic repairs/min</b>	<b>.307*</b>	.217
Sig (2-tailed)	.024	.172
<b>Other repairs/min</b>	<b>.265*</b>	.189
Sig (2-tailed)	.045	.207
<b>Repetitions/min</b>	.065	-.070
Sig (2-tailed)	.341	.381
<b>Restarts/min</b>	-.105	-.150
Sig (2-tailed)	.253	.259
<b>Dysfluencies/min</b>	.079	-.016
Sig (2-tailed)	.310	.472

\* $p < 0.05$ , fp=filled pauses

Table 9. *Correlations for higher- and lower-proficiency groups – AS and twelve fluency measures*

	AS task higher-proficiency group (N=51)	AS task lower-proficiency group (N=25)
<b>Rate A</b>	-.053	-.136
Sig (2-tailed)	.357	.262
<b>Rate B</b>	-.069	-.175
Sig (2-tailed)	.317	.206
<b>Fp/min</b>	-.036	-.095
Sig (2-tailed)	.401	.329
<b>Inter-clausal fp/min</b>	.010	.189
Sig (2-tailed)	.472	.188
<b>Intra-clausal fp/min</b>	-.053	.224
Sig (2-tailed)	.357	.136
<b>Repairs/min</b>	-.083	.191
Sig (2-tailed)	.284	.185
<b>Lexical repairs/min</b>	-.084	-.048
Sig (2-tailed)	.281	.411
<b>Morphosyntactic repairs/min</b>	-.103	<b>.432*</b>
Sig (2-tailed)	.237	.017
<b>Other repairs/min</b>	.019	-.209
Sig (2-tailed)	.448	.163
<b>Repetitions/min</b>	.089	.103
Sig (2-tailed)	.269	.316
<b>Restarts/min</b>	.075	<b>.333</b>
Sig (2-tailed)	.302	.056
<b>Dysfluencies/min</b>	.061	.170
Sig (2-tailed)	.337	.214

\*p&lt;0.05, fp=filled pauses

## 4. DISCUSSION

### 4.1. *Research questions 3 and 1*

Going against the status quo for discussing results, this section will begin by answering the third research question posed by the study, ‘Does the relationship between WM and fluency differ between measures of speed, breakdown and repair fluency?’ with a view to answering research question one simultaneously which was, ‘Does WM capacity affect L2 oral fluency?’.

The study strongly suggests that WM capacity influences the 3 types of fluency (as outlined by Skehan, 2009) to different extents. To begin with, it seems that there is probably no relationship between WM capacity and speed fluency (with one potential exception elaborated on in section 4.2., as it appears to be proficiency dependent). A more concrete yet modest relationship was found between WM capacity and breakdown fluency - relating to filled pauses. This may also be linked to proficiency and is to be discussed in section 4.2. By far the strongest link between WM capacity and fluency was found for repair fluency. Results suggest that learners with higher WM-capacity tend to use more morphosyntactic, appropriacy repairs, different repairs and restarts. This relationship was a weak one found in some of the correlational analyses and comparisons of the means.

How could this be explained? A speaker’s output is constantly being monitored and compared to what the speaker knows (the interlanguage system) using a feedback loop (Kormos, 2006). When there is incongruence between speech and the interlanguage system of the speaker, i.e. a mistake has been recognised or something has not been expressed as desired, the speaker will normally stop to either reformulate (in the case of morphosyntactic repairs) or reconceptualise (in the case of appropriacy repairs, different repairs and restarts). All these processes require conscious effort. For example, a morphosyntactic repair will usually require the adaptation of a word e.g. work→works, finded→found. This requires dynamic processing of information stored in WM, attentional resources (in order to notice the incongruence), and memory resources (to hold all the elements of the conversation whilst information manipulation is being carried out). This is a complex process, attentional and information processing resources (in other words, WM resources) are being stretched. Therefore, those with higher WM-capacity are more likely to have the available resources ‘left over’ (since speaking in a L2 is a demanding task in itself) to make these conscious, effortful repairs. This is in line with the theory posited by Sawyer and Ranta (2001) who suggested that higher

WM capacity allowed for more attentional resources to be freed up for L2 speech processing.

Bearing the above in mind, it is appropriate now to introduce lexical repairs, for which the opposite effect was found - higher-capacity WM participants made less lexical repairs. It is suggested that lexical repairs may be qualitatively different from the other types of repairs. Indeed, lexical repairs also require WM resources in terms of storage and attention to feedback. However, it is suggested here that dynamic processing is not part of the process, the speaker must instead return to the mental lexicon in order to search for a new lemma which better represents the preverbal message. It is hypothesised that this requires less WM resources and therefore it is not only higher WM capacity learners who can carry it out. But how can the fact that fewer lexical repairs were carried out by higher-WM learners be explained? Indeed, the process of lexical access may be more automatized in the higher capacity group which is in line with Segalowitz's (2010) idea that lexical access is one point of 'dysfluency vulnerability'. Thus, lexical access is more efficient in higher WM capacity learners and lemmas noticed as being incongruent are less likely to be selected in the first place. But if a higher WM capacity speaker has an advantage for lexical access why then do they not have a greater advantage for forming correct morphosyntactic structures? Their interlanguage system may still be incomplete or underspecified, higher WM does not guarantee high proficiency, they simply have more resources available to deal with morphosyntactic errors when they arise. Of course, an unrepaired utterance does not imply that it is 'correct', the error may not have been noticed. These are speculative hypotheses and certainly warrant further research.

These findings open up some interesting avenues. Schmidt's noticing hypothesis (1993) posits that the features of a language will not be learned unless they have been noticed. Thus repairs, being an overt indication of noticing one's own mistakes, are essential for the restructuring of an interlanguage and thus, learning. This is especially true for a learner using explicit, conscious learning mechanisms. Therefore, the understanding of the different processes involved in repair behaviours, how they are mediated by individual differences and how it relates to learning is an exciting area for future research as they so important for language learning.

Whilst restarts were included with morphosyntactic and 'other' repairs, their relationship with WM was found to be modest and much more complex. In the higher capacity WM group there was a trend for a positive correlation and in the lower capacity WM group there was a significant negative correlation (higher- and lower-

capacity RS groups, respectively). This suggests the existence of a non-linear relationship where some kind of WM threshold indicates a change in the effect of WM on restarting behaviours. Inappropriate statistical methods were used here to investigate non-linear relationships, but this finding suggests that the relationship may be more complex than was imagined.

To conclude this section, a return to research question one. It seems that WM capacity does affect L2 fluency but only some aspects of it.

#### *4.2. Research question 2*

Research question two asked whether proficiency had an effect on the relationship between WM capacity and fluency.

For speed fluency, there was a trend for a positive correlation between WM and speech rate (speed of speech) for lower-proficiency learners. Although only a trend was found here, it may be that WM makes a difference only at a lower-level of proficiency. Past a certain proficiency threshold, WM may no longer discriminate and all speakers will use comparable speeds.

For breakdown fluency it was found that for higher-proficiency learners there was a weak, negative correlation between filled pauses and WM suggesting that WM increases the speed of lexical access. These higher-WM/proficiency learners were not forced to pause to search for a word. This argument is supported by the fact that this pattern was found for intra-clausal filled pauses and not inter-clausal pauses. Inter-clausal pauses are frequent amongst native-speakers as they have a non-linguistic purpose, they allow time for conceptualisation (Tavakoli, 2011). Intra-clausal pauses relate to lexical access however and are much more frequent in non-native speakers. There is a delay in retrieval of the appropriate lemma. This seems to decrease as WM capacity increases. Why might this relationship exist only in higher-proficiency learners? WM capacity may only discriminate once a certain proficiency level is reached. As vocabulary size increases, the speaker's interlanguage system becomes more complex as do their utterances. It might be that only at this point does WM start to play a role. Here an important limitation of the study must be stated, that only filled pauses were measured and that pause duration was not taken into account. What is reported is only part of the total picture and learners with a tendency to pause silently are not represented here.

The positive relationship reported in section 4.1 for WM measures and morphosyntactic and other repairs were also found with the proficiency divide. Significant results were balanced across the higher-proficiency and lower-proficiency

groups with no clear pattern. Therefore it seems that if proficiency does affect the relationship between WM and repair fluency, it is a complex process whose intricacies were not detected in this study. This seems unlikely. Kormos (1999) found that self-repair frequency was unaffected by proficiency, only the type of repairs made are affected by proficiency.

#### *4.3. Points of discussion and limitations*

In order to probe the research questions more deeply groups were split into two or three based on task performance and proficiency scores using either k-means clustering or a median split. This could be considered a limitation of the study for three reasons. Firstly, the splits could be considered arbitrary in the case of the median split. Secondly, there is no reason to believe that the participants constitute a representative range of the attributes measured e.g. it is unlikely that the higher-proficiency group constituted high-proficiency in comparison to standardised norms. Finally, the power of the statistical analyses was considerably weakened by using fewer participants. However, the splitting of the groups can be justified by the insightful findings which resulted from it. Given the exploratory nature of the study and embryonic status of this particular research area, this was essential. Moreover, on splitting the groups the possible existence of non-linear relationships and thresholds arose, as in the case of restarts.

De Bot, Lowie and Verspoor (2007) describe a dynamic systems approach to language which states that any speaking situation that is thought to be a measure of competence will have a relationship to an individual's performance. This can be related to fluency where a dynamic relationship can be reasonably assumed for fluency and the context of a speech act - they affect each other. Segalowitz (2010) proposed L2 fluency is influenced by three components which interact; motivation to communicate, the larger social context and perceptual and cognitive experiences relating to fluency. This is an example of the numerous interesting avenues of investigation in relation to L2 fluency and show that the foci of research into fluency should not be limited to just cognitive variables such as WM.

In conclusion, it seems that WM capacity has affects repair fluency and the type of repair which a person makes. The difference between lexical repairs and the other types of repair measured here may reflect a distinction between the underlying processes mediating these repairs. Future research should further investigate repair frequency and type in relation to WM capacity as repairs are so important to L2 learning – they are indicative of attention to form, interlanguage restructuring and progress.

**APPENDIX***Appendix 1 – A table showing the descriptive statistics for the participants*

	<b>N</b>	<b>Minimum</b>	<b>Maximum</b>	<b>Mean</b>	<b>Std. Deviation</b>
<b>Rate A</b>	79	99,17	298,52	154,9101	34,99874
<b>Rate B</b>	79	86,78	256,23	138,6015	33,49441
<b>Filled pauses/min</b>	79	,50	14,24	5,0749	2,75451
<b>Inter-clausal filled pause/min</b>	79	,00	6,67	1,6507	1,22267
<b>Intra-clausal filled pauses/min</b>	79	,00	8,97	3,4241	1,97459
<b>Repairs/min</b>	79	,00	5,84	1,7993	1,03393
<b>Lexical repairs/min</b>	79	,00	4,54	,9636	,80468
<b>Morphosyntactic repairs/min</b>	79	,00	2,70	,6106	,56704
<b>Other repairs/min</b>	79	,00	1,53	,2460	,35529
<b>Repetitions/min</b>	79	,00	18,37	5,2190	3,40399
<b>Restarts/min</b>	79	,00	1,82	,3913	,45280
<b>Dysfluencies/min</b>	79	,27	20,82	7,4095	3,97473
<b>Reading span task</b>	79	7,00	75,00	35,9620	16,83589
<b>Letter span task</b>	65	5,00	93,00	66,9692	15,86365
<b>Attention span task</b>	76	-5,48	63,94	16,6897	15,16139
<b>Oxford Proficiency Test</b>	79	28,00	56,00	44,7089	6,40343

## REFERENCES

Arbuthnott, K. and Frank, J. (2000). Trail making test, part B as a measure of executive control: Validation using a set-switching paradigm. *Journal of Clinical and Experimental Neuropsychology*, 22, 518-528.

Atkinson, R.C.; Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In Spence, K.W.; Spence, J.T. (Eds.). *The psychology of learning and motivation (Volume 2)*. New York: Academic Press. pp. 89–195.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, 417-422.

Baddeley, A. (2001). The concept of episodic memory. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. 356(1413): 1345–1350.

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

Conway, A.R.A., Kane, M.J., Bunting, M.F. Hambrick, D.Z., Wilhelm, O. and Engle, R.W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*, 12, 769-786.

Dornyei, Z. and Kormos, J. (1998). Problem-solving mechanisms in L2 communication. *SSLA*, 20, 349-385.

Engle, R.W., Laughlin, J.E., Tuholski, S.W. and Conway, A.R.A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology*, 128, 309-331.

Iwashita, N., Brown, A., McNamara, T. and O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics*, 29, 24-49.

Juffs, A. and Harrington, M (2011). Aspects of working memory in L2 learning. *Language Teaching*, 44, 137-166.

Favreau, M., & Segalowitz, N. (1983). Automatic and controlled processes in the first and second language reading of fluent bilinguals. *Memory & Cognition*, 11, 565–574.

Fillmore, C.J. (1979). On fluency. In D. Kempler and W.S.Y. Wang (Eds), *Individual differences in language ability and language behaviour* (pp. 85-102). New York: Academic Press.

French, L.M. and O'Brien, I. (2008). Phonological memory and children's second language grammar learning. *Applied Linguistics*, 29, 463-487.

Gathercole, S.E., Service, E., Hitch, G.J., Adams, A. and Martin, A.J. (1999). Phonological short-term memory and vocabulary development: further evidence on the nature of the relationship. *Applied Cognitive Psychology*, 13, 65-77.

Steven Morris

Gilabert (2007). The simultaneous manipulation of task complexity along planning and +/- Here-and-Now: effects on L2 oral production. In García-Mayo, M.P (Ed.). *Investigating tasks in formal language learning*. Clevedon: Multilingual Matters.

Gilabert, R. and Muñoz, C. (2010). Differences in attainment and performance in a foreign language: The role of working memory. *International Journal of English Studies*, 10, 19-42.

Hummel, K.M. (2009). Aptitude, phonological memory and second language proficiency in non-novice adult learners. *Applied Linguistics*, 30, 225-249.

Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning*, 49 (2), 303-342.

Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah NJ: Lawrence Erlbaum.

Kormos, J. and Denes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32, 145-164.

Kormos, J. and Sáfár (2008). Phonological short-term memory and foreign language performance in intensive language learning. *Bilingualism, Language and Cognition*, 11, 261-271.

Kormos, J. & Trebits, A. (2011). Working memory capacity and narrative task performance. In Robinson, P. (Ed.). *Second language task complexity*. Amsterdam: John Benjamins p. 267-289.

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

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed), *Perspectives on fluency* (pp. 25-42). Ann Arbor: University of Michigan Press.

Mackey, A., Adams, R., Stafford, C, and Winke, P. (2010). Exploring the Relationship Between Modified Output and Working Memory Capacity. *Language Learning*, 60, 501-533.

Mizera, G.J. (2006). Working memory and L2 oral fluency. PhD Dissertation. University of Pittsburgh.

Mora, J. C. & Valls-Ferrer, M. (in press) Oral fluency, accuracy and complexity in formal instruction and study abroad learning contexts. *TESOL Quarterly*. doi: 10.1002/tesq.034

Mota, M.B. (2003). Working memory capacity and fluency, accuracy, complexity, and lexical density in L2 speech production. *Fragmentos*, 24, 69-104.

Steven Morris

O'Brien, I. Segalowitz, N., Collentine, J. and Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second language oral production by adult learners. *Applied Psycholinguistics*, 27, 377-402.

Sajavaara, K. (1987). Second language speech production: Factors affecting fluency. In H. W. Dechert and M. Raupach (Eds.) *Psycholinguistic models of production*. Norwood, NJ: Ablex. Pp. 137-174.

Sawyer, M. and L. Ranta. 2001: Aptitude, individual differences and instructional design, in P. Robinson, (Ed.) *Cognition and second language instruction*. Cambridge: Cambridge University Press. Pp. 319-353.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14, 357-385.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206-226.

Segalowitz, N. (2010). *The cognitive bases of second language fluency*. New York: Routledge.

Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers, *ELT Journal*. 65, 71-79.

Weissheimer, J. and Mota, M.B. (2009). Individual differences in working memory capacity and the development of L2 speech production. *Issues in Applied Linguistics*, 17, 93-112.