University of Southampton

FACULTY OF HUMANITIES

Department of Modern Languages and Linguistics

Exploring the Effects of Working Memory Capacity on Second Language Oral Fluency and the Acceptability of Object Resumptive Pronouns on Adult Learners

DOI: 10.5258/SOTON/D1533 doi:10.5258/SOTON/D1533

by

Deida Perea Irigoyen

ORCID ID **0000-0003-0424-7521**

Thesis for the degree of Doctor of Philosophy in Applied Linguistics

October 2020

University of Southampton

Abstract

FACULTY OF HUMANITIES

Modern Languages and Linguistics

Thesis for the degree of Doctor of Philosophy

EXPLORING THE EFFECTS OF WORKING MEMORY CAPACITY ON SECOND LANGUAGE ORAL FLUENCY AND THE ACCEPTABILITY OF OBJECT RESUMPTIVE PRONOUNS ON ADULT LEARNERS

Deida Perea Irigoyen

The purpose of this thesis is to investigate if working memory capacity (WMC) plays an important role in the acquisition of second language (L2) aspects that represent a challenge for adult learners in oral production and grammatical comprehension. With this in mind, the main goal is to explore the effects of WMC on the L2 oral fluency and the acceptability of object resumptive (R) pronouns among adult Spanish-speaking learners of English. Research in second language acquisition has demonstrated that learners who start their acquisition process after puberty attain the L2 in a variable manner (Schmid, 2011). Emergentist theory, on the one hand, sustains that late L2 learners are able to reach automatized, meaning fast and efficient (Segalowitz & Segalowitz, 1993), processing levels. On the other hand, research with a generative approach has demonstrated that these learners can acquire L2 features that are abstract and can only be comprehended through deep L2 grammatical knowledge (cf. Rothman & Slabakova, 2017; White, 2003, 2007; White & Juffs, 1998).

Therefore, it is important to explore if individual differences in cognition, particularly in working memory (WM), can explain the variability in the degrees of L2 attainment of these particular aspects amid adult L2 learners. Considering that WM serves to temporarily store information while underpinning higher-order skills (Baddeley, 2003a, 2003b, 2007, 2017), the proposal of this thesis is that a higher WMC allows adult learners to develop L2 automaticity and to cope with the parsing of certain L2 grammatical conditions.

In order to test the effects of WMC in these aspects of L2 acquisition, two studies were conducted. For these studies, a group of intermediate (*N = 22*) and a group of advanced (*N = 27*) adult Spanish-speaking learners of English were considered; as well as a group of adult native speakers of English (*N = 24*). The first study explored the relation between WMC and the temporal measures of oral fluency (Kormos, 2006; Skehan, 2003; Tavakoli & Skehan, 2005). WMC was measured using a listening and a reading span task (Conwell et al., 2005; Redick et al., 2012). For oral fluency, a speech generation task (Daneman, 1991; Segalowitz, 2010) was employed to measure speed, breakdown, and repair fluency (Skehan, 2003).

The results indicate that there is a significant equation found *(F (2,18) = 5.098, p < .003)* between WMC and the temporal measures of L2 oral fluency in the group of intermediate learners. The second study examines how WMC influences the acceptability of sentences with an object R pronoun (Leal-Mendez & Slabakova, 2012) condition. A grammaticality judgment task (Mackey & Gass, 2012; White, 2007) was designed to study the degree of acceptability of sentences with an object resumptive condition and with a gap in object resumptive position. It was observed that there is a significant equation *(F (6,15) = 3.504, p < .023)* between WMC and the sentences with a gap in object R position among the intermediate group of L2 learners.

The findings of this thesis indicate that adult L2 intermediate learners with a higher WMC have better measures of L2 oral fluency (higher measures of speed fluency and lower measures of breakdown and repair fluency); also, intermediate learners with a higher WMC are more accepting of sentences with a gap condition as they do not have to rely on the R pronoun to alleviate the processing load of the long-distance sentence structure (cf. Alexopoulou & Keller, 2007, 2013).

Table of Contents

Table of Tables

Table of Figures

Para mi mamá, Ma. Elva Irigoyen Silva

Research Thesis: Declaration of Authorship

Print name: Elva Deida Perea Irigoyen

Title of thesis: Exploring the effects of Working Memory Capacity in Second Language Oral Fluency and the Acceptability of Object Resumptive Pronouns on Adult Learners

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

None of this work has been published before submission

Signature: Date:

Acknowledgements

"We lived, as usual by ignoring. Ignoring isn't the same as ignorance, you have to work at it."

— Margaret Atwood, The Handmaid's Tale

I thought that the day in which I was to write this section was never going to come, but here I am. This is indeed a long journey, and a lonely one. However, as this thesis has taught me, nothing happens if one does not create deep connections; there is no learning if one does not pay attention to what is around.

Thus, the first person that I like to thank is my supervisor, Dr. Roumyana Slabakova, for her invaluable guidance; her advice and feedback helped me in numerous ways to come to this day. I would also like to thank the faculty and staff of the Distance programme in Applied Linguistics; specially to Dr. Adriana Patiño who always supported me to continue as well to Dr. Laura Dominguez, my advisor. I would like to express my gratitude to Dr. Tania Leal for kindly offering her guidance in a moment in which I really needed it. I truly appreciate the valuable time that she took to read my thesis document, give me advice, and ultimately me offering me her support.

I would also like to take the chance to thank Dr. Clare Wright for her insightful comments and suggestions; I would also like to thank Dr. Sarah Rule for the time she took to carefully read this thesis document and give me specific and pinpointed feedback. The recommendations that you gave me helped me not only to take my thesis document to the level that it required, but also to grow as an applied linguist and researcher. Also, I would like to thank you both for the invigorating academic discussion that we had during the VIVA; it was one of the most rewarding experiences of my life.

However, I have a powerful army of young scholars in-the-making who inspired me, helped me, and in more than a metaphorical sense held my hand through this process; especially in the second stage of this research. Special thanks to my angels, Brenda, Luisa, and Jacqueline; without your brilliance, willingness, and beautiful minds this research would have never been possible. Also, my deepest gratitude goes to Vianey, Ali, Diana Q., Michelle, Victor, Citlali, Eber, Iván R., Armando, Jaqueline S., Viridiana, and Javier C. All of you occupy a VIP spot in my procedural memory; your will, your selfless aid, and your remarkable minds made all the difference.

Also, I want to thank my former boss, Manuel Loera de la Rosa, who always supported me and believed in me; thank you for the confidence in my work, without it, I would not be accomplishing this goal. I cannot forget my friends who have become my family, and who always had a hug, a smile, a joke, a meal, and a drink to share when I needed it the most: Daniel, Saúl, Ercilia, Jezús, and Luisa; I've been fortunate to count on you. In the same manner, thank you to my dad, Alfredo Perea, and to my siblings Javier, Alfredo, and Abril; it has been your love and support what has given me the strength to come to this moment.

My eternal gratitude goes to my mom, to my Elvita: tú me has dado toda la fuerza para no rendirme, gracias por tu amor y tu luz.

A super special thank you goes to my best friend, Sussan. We started this academic journey together, and I could not have found a better partner in crime. Thank you for having my back, for the necessary advice, for the push, the comprehension, the encouragement, the support. This has been a road full of black holes, but your friendship always served as a lighthouse to find ground and direction.

A huge thank you also goes to my soul sister, Bea. Every time I think of the most significant moments of my life, you are there. We might live in different sides of an ocean, but we are always together. You always know what to tell me to continue moving forward; thanks to your strength and unique way of seeing the world, I get inspired and motivated to never give up. Thank you, Nenita!

Last, but never least, thank you, Jason. Your unconditional love, your tireless support, and your luminous presence in my life have been the best inspiration and the greatest motivation to continue. Thank you for always listening to my nerdiness, my rants, for hugging me in the moments of darkness, and for holding me when I thought I was not going to make it. Thank you for teaching me how to believe in myself. You are light and love, and this is for you. I love you; I love us.

Chapter 1 Introduction

The aim of the present thesis is to explore working memory capacity effects on second language oral fluency and the acceptability of English object resumptive pronouns among adult Spanish-speaking learners of English. Hence, the aim of this first chapter is to introduce the bases and overall contents that comprise this investigation.

To begin with, I present the rationale and justification of the current work on the grounds of second language acquisition (SLA) cognitive theory and research. Next, there is a presentation of the general and specific objectives, which includes the definition of second language (L2) oral fluency and resumptive (R) pronouns; and, how they are relevant objects in the SLA field. Then, the predictions and the general and specific research questions are provided. To conclude this chapter, I give an overview of the chapters included in the current work.

1.1 Second language acquisition and late learners

The present investigation originates in the search to comprehend why adult learners (L2Aers) arrive at different levels of second language proficiency, and specifically on the quest to identify what factors determine this variability in L2 attainment. The latter remains as an aspect that requires further exploration in current SLA theory and research.

One of the aspects that have brought attention to SLA cognitive research in terms of individual differences is working memory. This aspect of cognition has gathered attention owed to the role that it plays in the processing of information, but most importantly in learning. Hence, the work here presented focuses on investigating working memory as an aspect that might play a crucial role for late learners to acquire a second language, and on the degree of attainment of the L2. I attempt to elaborate on this point in the next section as I narrow down the purpose of the present research work.

A fundamental consideration for this thesis is the tenet that the learning of a second language is difficult, challenging, and overall demanding among late learners. In this regard, Schmid (2011) makes the remark that “whereas all normally developing children can attain full native language proficiency, there is considerable variability in ultimate attainment among older learners of a second language […]” (p. 51). Schmid (2011) adds that some

researchers attribute such variability to maturational constraints in processing (e.g. DeKeyser, 2010; Lenneberg, 1967; Penfield & Roberts, 1959; Johnson & Newport, 1989), while some others to the competition between the first language (L1) and the L2 during the acquisition stage (e.g. MacWhinney, 1997). Therefore, there is an agreement amid cognitive SLA theorists that L2 attainment widely varies among learners who began their acquisition process later in life; more specifically, after puberty (Schmid, 2011).

In this respect, this investigation does not seek to add to the controversy of whether maturational neurological conditions or competition between languages prevents post-puberty learners (referred to as "late learners" from now on) to reach nativelikeness. Rather, I want to focus the present work on the variability of L2 attainment; and, on the assumption that learning, and ultimately acquiring an L2, poses a major challenge for late learners' cognitive processing. This means that in order to overcome such challenge, late learners have to develop cognitive strategies and skills that can facilitate L2 learning and can lead them to a successful acquisition. In doing so, some late learners might be better at developing such skills and strategies in cognition. Consequently, the adequate development of the required cognitive strategies to overcome the challenge of L2 learning might be the reason behind the high proficiency of certain late learners; whereas the partial development, or the failure to develop these strategies and skills might explain why some L2 late learners are not as successful.

In light of the latter, working memory (WM) is a cognitive mechanism that has gathered attention in SLA. The reason is that WM, as Alan Baddeley (2003, 2007) has stated it, is "a temporary storage system under attentional control that underpins our capacity for complex thought" (p.31). In other words, what makes WM an essential component of L2 learning for late learners is its integration of cognitive tasks related to retention in short term memory (STM), while using higher skills and knowledge from long term memory (LTM) for the processing of information (Baddeley, 1983, 2007; Juffs & Harrington, 2011; Mackey, et al., 2010). In this regard, WM might serve as a central mechanism that aids late learners to succeed in the acquisition process; which requires them to learn considerable loads of new linguistic information, while they also need to cope with the myriad of tasks that acquiring an L2 requires. This might imply that second language late learners (L2Aers) need to rely more on the functions of WM given the processing constraints caused by brain maturation.

However, WM varies in capacity from individual to individual (cf. Cowan, 2005). The latter is an important aspect for the present research. These differences in capacity are

obviated in the pioneering work of Daneman & Carpenter (1980). In their research, they demonstrated that the functions of WM (Baddeley, 2007), are not only measurable, but also that individuals operate the functions of WM in different capacities. Having a higher capacity in WM has been correlated with a better performance in general knowledge tests (Daneman & Carpenter, 1980; Just & Carpenter, 1992). In terms of SLA, WM has been approached from multiple perspectives, but the implications of the individual differences (IDs) in this mechanism on late learners' L2 acquisition process need to be further explored.

The need for further research is evident on the limitations in terms of methodology (e.g. Juffs, 2004, 2005, 2006), contexts of L2 learning (e.g. Dussias & Piñar, 2010), L2 features considered for study (e.g. Felser & Roberts, 2007; Fortkamp, 1999; Mizera, 2006; Rodríguez, 2008; Sagarra, 2017), and a narrowed approach to the components of WM (e.g. Ellis & Sinclair, 1996). Overall, there is a lack of conclusive empirical results that can determine the repercussion that IDs in WM capacity (WMC) have on the learning and acquisition process of L2Aers (Juffs & Harrington, 2011; Wen, 2015, 2016).

Therefore, the research here presented attempts to broaden the constrained empirical knowledge on WM in relation to SLA by implementing two automated instruments (to measure WMC) and guarantee more data validity for methodological purposes. Moreover, this investigation analyses WMC effects on two aspects that pose a challenge for L2Aers to attain: the development of oral fluency (Segalowitz, 2010) and the acceptability of certain morphosyntactic features (cf. White, 2003, 2007; White & Juffs, 1998) such as English resumptive pronouns (cf. Alexopoulou & Keller, 2007, 2013). In this manner, the novelty of the present research is that it encompasses two dimensions of L2 cognitive processing: oral production (utterance fluency) and grammaticality (acceptability of resumptive pronouns) and; thus, it seeks to emphasize how IDs in WMC influence the degree of success in the attainment of adult L2 learners.

1.2 General and specific objectives

Considering the prior justification, the general objective of the present research is to explore WMC effects on L2 oral fluency and the acceptability of object resumptive (R) pronouns among adult Spanish-speaking learners of English. In order to achieve this general objective, I pursue two specific objectives:

1) analyse the effects of WMC in the development of L2 oral fluency amid L2Aers, and

2) analyse the effects of WMC in the degree of acceptability of object R pronouns.

The reason to examine the effects of WMC on L2 oral fluency is that it is a very costly aspect to develop in the L2, and not attained by all late learners. The explanation for its difficulty in L2 development and acquisition might be related to its characterization; for Segalowitz (2010) oral fluency is "an observable characteristic of real time speech behaviour (p.6)", which can be completed with Fortkampt's (1999) observation that it is "one component of oral proficiency and is basically related to speech rapidity, to the flow of speech without this being impeded by hesitations (p.7)".

The former definitions indicate that for L2 learners to reach oral fluency, it is necessary to efficiently retrieve grammatical knowledge as they convey and adapt meaning and lexicon (as observed in De Bot's, 1992 adaptation of Levelt's 1972 model to L2 speech) to real-life contexts, which requires a rapid flow of speech delivery (Skehan, 2009). In this sense, DeKeyser (2001) mentions that oral fluency has received attention, mainly because it is an aspect of L2 acquisition that derives from the development of automaticity (DeKeyser, 2001; Schmidt, 1992) provided that it is the result of automatizing grammar and vocabulary (p.141).

Regarding this, SLA researchers and theorists with an emergentist focus have observed that L2Aers reach automaticity in the L2 by detecting linguistic cues and patterns, which require to be frequently reinforced in the L2 input (Ellis, 2008; Ellis & Sinclair, 1996; McLaughlin & Heredia, 1996). Nevertheless, empirical data in SLA emergentist research shows that even when a significant number of learners achieve to automatize L2 advanced phonology and syntax, there are still some L2Aers that do not get to the same result. In this respect, studies focusing solely in L2 fluency display similar results: there is variability in attainment of L2 oral fluency amidst late learners. Moreover, empirical research on L2 automaticity and oral output processing is scarce (Kormos, 2009) to make any assumptions as to how L2 fluency is developed.

Furthermore, the second specific objective of the current research is to examine the WMC effects on the acceptability of object resumptive pronouns, given that this grammatical property in English as an L2: a) is not frequently used (might be more used in spoken English) (cf. Leal-Mendez & Slabakova, 2012), b) is used in very specific grammatical contexts of English (cf. Rouveret, 2007; Tsimpli & Dimitrakapolou), c) does not have an exact grammatical correspondence in Spanish (cf. Leal-Mendez & Slabakova, 2012), and d) might serve as an aiding device to process long-distance Ā-bound dependencies (cf. Alexopoulou, 2009;

Alexopoulou & Keller, 2007, 2013). An example of English resumptive pronouns can be observed in 1a, taken from Ross (1967) (the resumptive pronoun is marked in bold).

1 (a) I just saw the girl who Long John's claim that **she** was a Venusian made all the headlines.

Some SLA generative theorists consider that this type of linguistic features posit a major difficulty in L2 acquisition, given that they require a deep, underlying knowledge of the grammar (Leal-Mendez & Slabakova, 2012; White, 2003, 2007; White & Juffs, 1998). However, it has been observed that L2 late learners are able to accurately distinguish L2 grammatical features such as wh-movement (White & Juffs, 1998), dative clitics (de Garavito, 2006), resumptive pronouns (Leal-Mendez & Slabakova, 2012), among others. Nevertheless, as discussed before, the results in these studies indicate that not all the learners were able to determine the grammaticality of these L2 grammatical items with accuracy (Juffs & Rodriguez, 2014). Consequently, these empirical findings in sentence evaluation bring up the question of whether IDs in WMC, among late learners, are a determining factor to successfully develop the configuration of the L2 grammar.

In sum, given that WM (Baddeley & Hitch, 1974; Baddeley, 2000) is a mechanism that functions as a temporary storage of perceived input while processing higher-order skills from LTM (Baddeley, 2007), its involvement in L2 acquisition needs to be further explored as it might determine the L2Aer's degree of L2 attainment. Thus, the present work has as an objective to investigate the influence of IDs in WMC on aspects of the L2 that a) represent a challenge on both production and grammar, and b) are susceptible to varying results of L2 attainment among L2Aers. In this respect, the specific objectives of the present work are to explore WMC effects on 1) second language oral fluency, and 2) the acceptability of object resumptive pronouns.

1.3 Predictions and research questions

Considering that 1) late learners might need to rely more on mechanisms such as WM given their limitations in processing caused by age and brain maturation, and 2) there is an observed variability in the attainment of L2Aers, I predict that those learners with a higher WMC will show higher levels of oral fluency and will be less accepting of R pronouns as these learners do not need to rely on the object R pronoun, as an alleviating device, to cope with the processing load imposed by the L2 long-distance sentence.

With the purpose of testing this prediction, this study addresses the following general question: What is the role that WMC plays in the development of oral fluency and the acceptability of object resumptive pronouns among late L2 learners? To answer this question, this investigation seeks to respond the following specific questions:

1) **Does working memory capacity have an effect on the oral fluency of late L2 learners?**
2) **Does working memory capacity have an effect on the acceptability of object resumptives on late L2 learners?**

To answer these questions, the present research follows a quantitative research methodology. To extract the required data, three groups of participants were considered. One group served as a control group and it consists of 24 adult native speakers of English with at least two years of college education experience. Two groups were comprised of L2 late learners. The first group of learners is comprised of 27 Spanish-speaking learners of English about to conclude or who have concluded the English language teaching (ELT) bachelor's program. The other group of L2Aers involves 22 Spanish-speaking learners of English enrolled in an English as a Foreign Language (EFL) program at the time of the study; the ages of the participants ranges between 22 and 35. An English Oral Proficiency Interactive test by ACTFL (American Council for the Teaching of Foreign Languages) was administered to confirm the level of proficiency of the learners. Thus, the 22 EFL students were considered as the intermediate group of learners, whereas the 27 students or graduates of the B.A. in ELT were considered as the advanced group of learners.

Additionally, two studies were conducted to answer the specific questions of the current research. The first study was designed to analyse the effects of WMC on measures of oral fluency in order to answer the research question 1 of this investigation. With this in mind, two tasks were employed to measure WMC: an automated reading span task (RST) and an automated listening span task (LST) (Daneman & Carpenter, 1980; Conway et al., 2005; Unsworth et al., 2005); the two tasks were applied to the three groups of participants. In addition, a speech generation task (SGT) (Daneman, 1991; Fortkamp, 1999) was designed to obtain temporal measures of utterance fluency (Kormos, 2006) following Skehan's (2003) approach; in this approach speed, breakdown, and repair fluency are considered as the three aspects that comprise oral fluency. The temporal measures for speed, breakdown and repair fluency were calculated using voice analysis software (PRAAT) (De Jong & Wempe, 2009).

The second study contemplates the analysis of working memory capacity with respect to degree of acceptance of sentences containing resumptive pronouns. Thus, the measures obtained in the span tasks to measure WMC in the first study are considered. Moreover, two grammaticality judgment tasks (GJT) (Mackey & Gass, 2012; White, 2007) were designed (one in English and one in Spanish) to determine the acceptability of sentences containing R pronouns in object position.

Accordingly, the data extracted from each task was saved and organized in a general data base using Microsoft Excel. The statistical analysis of the data was performed in SPSS. For Study 1, the measures for WMC were approached descriptively and inferentially by means of a One-Way ANOVA. Similarly, the measures of oral fluency were statistically analysed per group of participants and reported in terms of speed, breakdown and repair fluency (Skehan, 2003). A multiple regression analysis was applied to study the effects of WMC in the temporal measures of oral fluency; these results were reported per group of participants.

For Study 2, the responses for the sentences comprising the GJT were analysed descriptively, first; the reactions of each of the three groups of participants to each of the sentences included in the GJT were analysed using descriptive statistics information (means and standard deviations). Additionally, the responses per sentence condition included in the GJT (resumption in object position, gap in object resumptive position) were compared with a chi-square analysis to determine if there were any significant differences among the three group of participants. Also, a Pearson covariance analysis was performed to find if the L2Aers' reactions to the sentence conditions in the GJT in English significantly correlated to their responses to these sentence conditions in Spanish, their L1. Finally, a multiple regression analysis was performed to explore the influence of WMC on the acceptability of the sentence conditions in the GJT.

1.4 Overview of the contents in this thesis

In Chapter 2, I offer a review of the literature that comprehends the grounding theory and research of SLA guiding the present investigation, together with a discussion of relevant theory and research concerning working memory, second language oral fluency, and resumptive pronouns. Chapter 3 describes the methodology that was adopted to explore the effects of working memory capacity on the acceptability of R pronouns and L2 oral fluency in adult Spanish-speaking learners of English. In Chapter 4, I present the results obtained in

Study 1; Chapter 5 is intended to report the results observed in Study 2. In Chapter 6 I discuss the results obtained in Study 1 and Study 2 as I answer to the specific research questions. Finally, Chapter 7 approaches the limitations, pedagogical implications, and conclusions of the current investigation.

Chapter 2 Review of the literature

The following chapter addresses the theory and empirical research around the concepts and objective of the current study. Thus, the purpose of this chapter is to provide a thorough description of the theoretical frameworks as well as of the research studies that led to the investigation of working memory capacity and its effects on oral fluency and the acceptability of resumptive pronouns of adult second language learners of English. In this manner, not only will the concepts that comprise this study be better comprehended, but also the empirical research that has led to its development will frame its relevance and justification.

2.1 Cognitive perspectives on SLA that focus on late learners' L2 developmental processes

The following section is intended to provide the background theory that accompanies the research work here developed. Given that the purpose of this study is to analyse the effects of working memory capacity on two aspects of second language acquisition: 1) oral fluency, and 2) the acceptability of resumptive pronouns, it is necessary to provide a description of the frameworks that have further developed and justified the need for their study. One of the concerns of emergentist and generative approaches to SLA is the analysis of aspects that indicate that late L2 learners are able to reach high and complex levels of attainment in the L2. Of particular interest for this research study is the background that emergentist foci have offered on the development of automatization, and that generative approaches provide in terms of second language grammars.

Consequently, the findings in emergentist theory and research have been key to understand that second language late learners (L2Aers) may become highly proficient, which represents L2 automaticity according to authors such as DeKeyser, Ellis, and MacWhinney. In addition, researchers within a generative approach have shed light on how L2Aers configure their L2 grammatical knowledge; thus, their empirical findings have made it possible to understand what linguistic properties are available for this subset of learners and how L2 grammars are configured. Given the prior, this section gives a brief account of emergentist and generative approaches to SLA as these frameworks set the scenario for the relevance of studying L2 oral fluency and the acceptability of resumptive pronouns.

2.1.1 The emergentist perspective of SLA

As a cognitive SLA framework, emergentist theory aims at uncovering the cognitive mechanisms involved in the attainment of a second language on late learners. Thus, the description of this framework is a crucial aiding guide to inform and understand why working memory capacity as a cognitive mechanism that underpins higher-order skills (e.g. reasoning, problem-solving, decision-making, etc.), while it retains information (Baddeley, 2007) might be involved in the development of automaticity of L2 output in the form of oral fluency; which is the main goal of this study.

To begin with, proponents of emergentist perspectives to SLA assert that learning a second language is like learning any other complex skill; for example, learning how to play an instrument. Furthermore, emergentist theory, places linguistic input at the core of its theory, analysis, and research. Therefore, authors who advocate for emergentist explanations to SLA propose that L2 learners are able to become proficient in the L2 by means of perceiving input, internalizing such input and making a series of mental associations with it. According to authors under this framework, the latter cognitive processes lead the learner not only to understand the target language, but also to be functional in it; a level that they have identified as L2 automaticity (which will be further discussed later under this section).

Hence, authors that support an emergentist approach, such as McLaughlin and Heredia (1996), propose that learning an L2 would require mechanisms that include a "pattern recognition, output systems, memory systems (e.g. short-term memory and long-term memory), and systems for intrinsic reasoning (p. 213)". These mechanisms are further explained by the same authors as part of a perspective under the name of information processing (IP). These elements, or mechanisms, serve as the principles that comprise L2 learning as viewed by those who advocate for an input-mental processes view. Briefly explained, in terms of IP, L2 learning is a process that can be divided into the development of simpler skills, which are then hierarchically organized and will be initially learned by means of "controlled processing". The latter is a central concept to the IP perspective, as well as "automatic processing".

Under information processing theory, the beginning stage of second language learning starts when the learner perceives the input, which triggers the activation of memory nodes temporarily followed by the restructuring or reorganization of information; the latter is described by IP proponents as a cognitive process that demands effort or difficulty for the

second language learner. This effort and/or difficulty is the result of mainly two limiting aspects of cognition: 1) information, or that the L2 input has not been learned yet, and 2) the learner is constrained by a limited capacity to process information. Namely, the limits in cognitive processing are known as controlled processing under the IP perspective to SLA.

In this vein, IP theorists and researchers propose that the L2 learner will go through a controlled processing stage to later retrieve information in a coordinated, integrated and efficient manner. This theoretical approach to SLA assumes that by means of accessing and processing the L2 through controlled-processing (which involves constant rehearsal in WM, frequency of exposure to the L2 input, noticing salient items, etc.), the L2Aer will be able to reach the stage of automatic-processing in the L2.Thus, the fact that the L2 learner uses the L2 in an automatic manner ultimately means that the learner has reached high levels of proficiency under the information processing theory.

All in all, under an IP perspective, the notion that L2 learning occurs from the change of "controlled to automatic processing" is based on the principle that L2 learning implies a shift from "declarative to procedural" memory or knowledge; in which declarative is associated with central processing, and procedural to automatic processing. The latter concepts are illustrated by Ullman (2004) who associates lexical memory to the declarative memory system and aspects of grammar to the procedural memory system. In this regard, the findings in Ullman (2004) are used to support the claims of SLA approaches such as the IP framework described above.

Ullman (2004) defines the declarative memory system as the result of many brain regions (temporal lobe structures, the hippocampal region, entorhinal cortex, etc.) in which the usage of facts, events, rapid learning, as well as the learning of arbitrarily related information occurs. In addition, Ullman (2004) defines procedural memory as the result of brain regions (basal ganglia, cerebellum, Broca's area, portions of parietal cortex, etc.) to which sensory-motor, cognitive 'habits', 'skills', and other procedures are attributed. Ullman (2004) mentions that language learning is gradual, sub served by the declarative memory and has relations with rule-governed, encapsulated and rigid non-modifiable knowledge (implicit knowledge).

Given neurological evidence, Ullman (2004) points out that the declarative system is in charge of lexical acquisition and states that 'the brain structures that sub serve declarative memory play analogous roles in lexical memory' (p. 245); however, the declarative memory

system is also associated with the learning of new information, which is said to be processed mainly by working memory. The procedural memory system, on the other hand, is in charge of aspects such as mental grammar, and it is associated with the fast and efficient retrieval of information that is mostly stored in long-term memory. In lieu of what IP accounts for, it can be understood that the assumptions made about second language learning are rooted on the findings of the D/P model, in the sense that the L2Aer has to engage in declarative memory (or controlled-processing) in their developmental L2 stages to be able to later operate the necessary aspects of the L2 through the procedural memory system (automatic-processing).

However, the foundations of IP as an emergentist perspective to SLA is key to the current study; particularly, the interrelation between IP and the D/P model. The study of working memory capacity in L2 learning contributes to the IP proposal that automatic processing in the L2, or high L2 proficiency, results from the demand on cognitive mechanisms that utilize higher-order skills, while information is perceived (or what IP proponents conceptualize as central processing). Consequently, L2 oral fluency, as an aspect of language that results from automatic processing, should be associated with the involvement of WM in developmental stages of L2 acquisition.

Nevertheless, it is necessary to point out that there are slight discrepancies on what advocates of emergentist perspectives propose as the principles that explain second language acquisition. As a result, in the SLA literature, the emergentist perspective can be interpreted in multiple ways, as it can be introduced and labelled with a wide array of titles, subdivisions and associations, which do not necessarily converge with the information processing theory. Thus, McLaughlin (1990) characterizes the emergentist framework to SLA as a derivation of cognitive psychology that is concerned with: 1) shedding light on mental processes involved in the acquisition and use of knowledge; 2) emphasizing mental structure, or organization; and 3) acknowledging that human cognition requires strategies, analysis, understanding, remembering and producing language.

Considering the former, there are emergentist theorists that lean towards a view of L2 learning that emerges from the connection between linguistic representations, patterns, extensive linguistic exposure, and mental mechanisms. To this end, Ellis (1998) explains that very complex language representations can “emerge” from the interaction of simple learning mechanisms when exposed to complex language data; therefore, L2 learning is a consequence of the interaction between the information provided by the L2 input, the

extensive exposure to it, and the learner's learning mechanisms. However, it can still be concluded that in spite of slight differences in theory, proponents of emergentist perspectives argue that L2 acquisition "emerges" from the input, and its synergy with general mechanisms of learning in cognition.

Up to this point, it has been discussed that one of the main theoretical frameworks of emergentist theory is information processing. Of particular interest for this study here developed is the principle that L2 learning occurs when processing shifts from controlled to automatic. The latter is intimately connected with the declarative/procedural model, which attributes language learning to the activation and demand allotted on cognitive mechanisms of general learning. Hence, the involvement of working memory as a mechanism that temporarily stores information while underpinning higher-order skills should be key in the reaching of L2 automaticity as claimed by the information processing framework within the emergentist perspective.

Nevertheless, it has already been discussed how there are slight differences among the proponents that support the emergentist view of SLA. However, there seems to be an agreement on portraying information processing as an approach that avails the principles of the emergentist theory, which ultimately sustains that L2 acquisition is triggered from the interaction between the input and the general mechanisms of learning in cognition.

Given the relative mismatches among emergentist approaches in the literature, it is necessary to give an account of what they entail, how they differ from IP, and how they encompass the emergentist theory. In this way, the study of WMC and its association with L2 oral fluency is better justified and comprehended as a necessary step towards shedding light on the cognitive mechanisms and individual differences (IDs) that intervene in the degree of attainment that learners accomplish. For that reason, approaches such as the competition model and connectionism will be generally defined in the next subsection.

2.1.1.1 The competition model within the emergentist framework

One of the main approaches developed under the emergentist view to SLA is the competition model. Firstly, the competition model (CM) as proposed by MacWhinney (1986) might be an alternative to generative, empiricist, formalist and functionalist theories for language acquisition. MacWhinney (1986) bases his findings on a Darwinian perspective, in which, just like species in nature tend to adapt and learn from their environment, so do linguistic items in the brain adapt to the circumstances they are exposed to. MacWhinney

(1986) states that L2 comprehension starts with auditory processing and the storage of lexical units; which entails lexical segmentation and the attachment or creation of a 'commitment' of cues to these lexical units. Nevertheless, comprehension also depends on the competition and control of role attachment to grammatical functions, or case roles.

MacWhinney (1986) highlights the following grammatical roles as being recognized by the competition model: 1) subject, 2) object, 3) indirect, 4) final, 5) head, 6) complement, 7) coordinate, 8) topic, 9) focus, and 10) antecedent. These roles might be separated by case, thus allowing certain cues or arguments to be assigned to them. Still, MacWhinney (1986) points out that the assignment of cues to cases or roles is subject to competition; this entails that competition of cue assignment gives strength to certain cues over others to pertain to certain grammatical roles and to the lexical items that might fit in them.

Therefore, the parsing of lexical items in terms of argument-filling or attachment on the CM begins with the lexical item that has a given grammatical role (depending largely on the recognition of its preverbal or post-verbal cue). These grammatical roles have arguments attached to them, which allows a link between lexical items to form a resulting speech phrase. The latter implies that the L2 input contains the necessary information to provoke acquisition; provided that the learner activates the prime cognitive mechanisms that can detect these cues in the L2.

More focused on L2 acquisition, further work by MacWhinney (2002), emphasizes that the CM takes on a constructive, data-driven process perspective to account for L1 and L2 acquisition; which is clearly different to universal linguistic perspectives based on principles and parameters as proposed by Chomsky (1965). In lieu of the latter, MacWhinney (2002) describes L2 acquisition in terms of the role of the input and the detection of its cues in the L2 learner's cognitive processing. Thus, the author explains that the primary and first step is input being presented as it is crucial for language comprehension. Once input is presented, it offers the possibility to detect patterns, or "cues".

Evidence to support the view on cues offered in the input is a cross-linguistic study on at least twelve languages. MacWhinney (2002) finds that languages have saliency on certain "cues" such as case-marking or syntactic structure. Thus, the learning of a given language will depend on strengthening cues from the input. The strengthening of cues is achieved by having "task frequency" (relating cues to certain linguistic tasks), "availability" (the cue is available given a certain linguistic task), "simple reliability" (the cue is reliable if it is functional

for the task), and “conflict reliability” (the cue is still suitable in the face of conflicting choices for tasks or cues) (MacWhinney, 2002, p. 5).

Moreover, the learner is seen from a cognitive neurological perspective. According to the author, on L2 acquisition, first, the learner has a great deal of transfer; second, these types of transfer get corrected; and third, these errors from transfer get minimized. Therefore, if the learner creates associations between phonological strings and semantic loads in the L2, which will imply a restructuring of underlying neuronal maps, the learner is said to increase automaticity in L2 lexical access. Such automaticity is regarded by the author as a firewall that will work against interference or transfer effects from the L1 to the L2.

Thus, the competition model is an emergentist approach that explains L2 acquisition by means of a full reliability on the “cues” or patterns offered in the linguistic input. Regardless of the differences with the information processing approach, the competition model acknowledges that late learners are able to reach L2 automaticity by means of cue detection which demands a considerable use of cognitive mechanisms. In such manner, this approach emphasizes the need for investigating individual differences in working memory capacity since it is an approach that claims that the L2 learner’s activation of cognitive mechanisms such as working memory is key to attain a high level of proficiency, or L2 automaticity.

This reinforces the notion that WM must play a crucial role on the late learner’s L2 acquisition process as it is necessary for the storing of the novel L2 linguistic input, while it serves to detect L2 “cues” or patterns. Consequently, the degree of activation of WM might determine the development of L2 automaticity and might explain why some L2Aers are able to reach higher levels of proficiency; the latter is demonstrated by the degree in which WMC correlates with the learners’ measures of L2 oral fluency in the present research work. The CM would be complemented by the findings of the present study since they expand the scope of this approach in terms of how the development of L2 automaticity is achieved.

2.1.1.2 **The connectionist approach within the emergentist framework**

One more approach stemming from the emergentist view to SLA is connectionism. Firstly, connectionism has its roots in psychology that dates back to the 1940s and is based on the notion that learning takes place when mental associations between the stimuli and the responses to such stimuli increase. Therefore, connectionist approaches are based on what is known about the human brain in terms of functioning. Accordingly, Ellis (1998)

suggests that proponents of connectionism are concerned with investigating "how simple learning mechanisms in artificial neural networks are able to acquire the associations between, for example, forms and meanings, along with their respective reliabilities and validities, and then use these associations to produce novel responses by "online" generalization (p. 638)".

One of the main assumptions sustained in connectionism is that language behaviour is rule-like, but not necessarily rule-governed by an innate mechanism or system (Ellis, 1998). Although some literature in SLA does not directly link connectionism to emergentist approaches formerly mentioned in this section, Ellis (1998) considers that emergentist views to SLA do involve connectionism given that it serves as a source of computational tools, that allow one to explore the conditions by which language emerges or comes to be (p. 645). Hence, proponents of connectionism base their research on computer models that simulate a vast number of neurons, connected in a parallel manner, as a way to emulate how a language is acquired (Ellis, 1998).

One of the most prominent models derived from connectionism is based on the parallel arrangement neurons and their function in the parallel distributed processing (PDP) approach, mainly studied by Rumelhart and McClelland (1986). Namely, PDP has studied the pathways by which networks of nodes are connected in the brain when processing language, or information in general. As a result, the associations between these nodes, in such networks, are called connections strengths or patterns of activation. To that end, PDP proponents relate the strength of associations to the frequency of linguistic input and the nature of the feedback that the computer network receives, in an attempt to compare what occurs in the computational network to how the human brain acquires a language. Hence, under a PDP perspective both frequency and factors such as noticeable patterns in the input are crucial causal factors for L2 acquisition.

Thus, broadly speaking, connectionism is an emergentist approach that relates language acquisition to the creation of neurological connections resulting from the exposure to L2 input. These connections strengthen with the frequent encounter of L2 items, and the noticing of the saliency of L2 features. Connectionist researchers study how computational models acquire language by training them to detect patterns from linguistic input and compare these outcomes to how learners, and L2 learners acquire language in similar ways (Ellis & Schmidt, 1998; Matesa & Anderson, 2000).

Connectionist theory also addresses the relevance of the input to trigger language acquisition in cognition. Similar to the competition model, connectionism considers that the strengthening of neural associations for the L2 deriving from the identification of patterns and linguistic saliency poses a considerable demand in cognitive mechanisms for the acquisition of a second language because it requires the use of complex thought, or complex cognitive processing.

Therefore, connectionism also justifies the study of working memory capacity as an aspect that might be directly involved with the efficient acquisition of a second language. As with findings born in the competition model, the establishing of WMC as a relevant influence on high levels of L2 oral fluency, and even on the acceptability of resumptive pronouns among L2 late learners will reinforce the notion that L2 acquisition results from the activation and "programming" of cognitive mechanisms as proposed in connectionism. Consequently, the present study might expand the knowledge of this approach in terms of identifying the specific cognitive mechanisms that are in charge of the processing of L2 input as well as the programming of neural networks that lead to second language acquisition.

2.1.1.3 The convergence in emergentist approaches to SLA

However, research based on the competition model and connectionism does not stay separated from one another. Some studies rely on both of these approaches to set their hypothesis and develop their methodologies; for example, the study performed by MacWhinney, et. al. (1989) in which the acquisition of gender, number, and case paradigm for the German definite article are studied simultaneously. Even when three computer-based models are used, this study is aiming at exploring acquisition of these three grammatical features based on one of the main constructs of the competition model, namely cue strength.

In order to do so, three models were used; each model works on a series of networks composed of several internal layers of 'nodes', plus an input and an output layer. Every time there is a trial, each connection on the network assumes a value, which is the product of the current activation and the node on the input side of the connection and the strength or 'weight' associated with that connection. All models were presented with German nouns; each node on the input layer described a single cue. If the cue was present, an input node was fully activated, and if it was not then the node remained inactive.

The first model was supposed to mimic the acquisition of a young native German learner. Words presented to this model were extracted from a German corpus of 80,000

words. Results show that on the training session the network was able to proceed at a level of mastery of items and it was able to select one of the six forms of the definite article. Two items remained unlearned. The second model differed from the first in that nouns were not disambiguated. It was possible for two different nouns to have identical identifications. Results on this second model show that performance on the network dropped significantly with this experiment. Learning occurred after 50 epochs, missing still 560 tokens.

The authors relate this to the network simply memorizing a map between the nouns and the articles. The third model relies on a phonological representation of each noun and it made no use of the 11 arbitrary disambiguating units as in the first model. 130 units were used distributed over 13 slots with 10 features (+labial, +coronal, +voice, +high, etc.). Results on this third model show that performance improved at both 50 and 100 epochs. The authors interpret these results as the networks being better guides for the extraction of correct higher-level cues when the units are not hand-crafted for simulation.

According to MacWhinney et. al. (1989), the model matches the data on the acquisition of the declension of the definite article in German. It also matches the omission that has already been reported in articles on early acquisition stages, together with the tendency to overgeneralize the feminine. Even though, there is mention of their results accounting for information-processing views on language acquisition as the processes of rote, combination, analogy, and paradigm application are represented as patterns of associations between cues (p. 275). However, this is a paper that demonstrates the fine line that divides the notions of connectionism, through the development of computer-based models that simulate the human brain and its acquisition of language, and the notions of CM that state a cue's strength of neurological mental maps given input and its frequency.

The latter is useful for the development of the present study in that computer models simulate an "ideal" cognitive processor of the language; however, MacWhinney et. al. (1989) do not specify what aspects of the processing in the model create a failure to identify the "cues" presented to it. Also, the study cannot state that all those actual learners of German, in other studies, had the exact results in acquiring the declension of the definite article. Thus, the margins of error on the computational model, and the lack of general results on acquisition on human learners opens the window to study what individual differences in cognition lead to the successful detection of salient cues from the L2 input. Given this, the present study contributes to emergentist research as it is the case of this early work by

MacWhinney et. al. (1989) by shedding light on the specific mechanisms that either allow or deter L2 acquisition.

Additionally, the work by MacWhinney (2004) also adds to the "overlapping" or connection that exists between, the competition model, connectionism, and information processing as approaches that derive and/or comprise emergentist theory. In his study, MacWhinney (2004) proposes a unified model based on the competition model, which aims at explaining processes and mechanisms of late L2 language learning. First, the unified model is based on the idea that linguistic forms are being constructed and ultimately stored in what is denominated as self-organizing maps (p. 342). These self-organizing maps are claimed to be integrated by means of engaging in processes such as buffering, chunking and resonance.

To further understand the concept of self-organizing maps, MacWhinney (2004) explains that they function as sheets where neurons or units are connected to other neighbouring sets of neurons and start being programmed with the L2 in a series of trials. For each trial a given set of neurons will be activated, while other sets will not. The process of activation (or learning), within sets of neurons in self-organizing maps, consists of three phases. The author points out that these self-organizing maps have "a correspondence with the nature of cortical regions which possess a 'spatial organization […] to the physical structure of the input space (p. 344)". He uses this as evidence to support his view that self-organizing maps operate in similar ways at the syllable, lexical, and construction levels.

The author concludes that his findings imply that late L2 language learners will need to apply further strategies to develop the process that these self-organizing maps were capable of carrying out; such as optimization of input, promotion of L2 resonance, and avoidance of processes that deter the internalization of chunks from the input. In addition, MacWhinney (2004) points out that a unified model is needed to account for some holes found on the competition model, and to provide a better understanding for the functioning of self-organizing maps in relation with competition and cue strength.

In this sense, this study is related to the investigation here presented since, as in the findings of MacWhinney et al. (1989), the role of complex mechanisms is emphasized. Moreover, the findings in studies such as this reinforce the proposal that L2 learning is not a goal to be achieved effortlessly and by all late learners, but that it is dependent on the use and/or activation of certain cognitive mechanisms. Thus, mechanisms such as WM make it

possible for L2 learners to optimize input, promote L2 resonance, and internalize L2 chunks; elements that MacWhinney et al. (1989) deem necessary to achieve L2 acquisition.

One more iconic study under the light of emergentist approaches is the one by Hernandez et al. (2005), which also agrees in that the emergentist view to SLA involves knowledge from the competition model, connectionism, and even neurology. These authors depart from the notion that emergentist theory explains both L1 and L2 acquisition based on the evidence found in studies concerned with neuronal plasticity, competition, and transfer. Based on what was theorized by Elizabeth Bates (1999), who defied a critical period existence for language acquisition, the authors expand on this perspective by adding aspects such as completion and entrenchment as crucial for L2 language learning. These two concepts are further defined by their theorization of how simultaneous bilinguals acquire language.

First, they point out that it takes bilingual children a few months to pick up the specific segmental differences between the two languages and that by the time children start producing their first sentences there is almost no interaction between the two languages on the children's production. As an explanation for this low interaction between languages, the concept of competition is introduced based on the competition model; later, the authors indicate resonance, parasitism (erroneous parsing of sentences), and entrenchment also play a role on L2 learning. Consequently, competition is defined as the leverage between L1 and L2 items that the bilingual undergoes; and, entrenchment as the intertwining of L1 to L2 items or vice versa. Thus, bilingual children go through a process of continual practice that strengthens activation or inhibition as they continue using their languages, in which there is a competition between the L1 and the L2, and very little entrenchment takes place.

Furthermore, Hernandez et al. (2005) support their views by highlighting the findings of a neural network model that can be trained on the development of lexicon called the DevLex model. The authors indicate that the fact that such a network model gets to develop modular representations for different languages is evidence that modules are created, not innate. Given this, the authors discuss acquisition on late bilinguals, and challenge the notion of a critical period since late learners can indeed achieve high levels of L2 proficiency.

Thus, based on what their DevLex model is able to achieve, the prediction made in this study is that given that late learners have less plasticity, young late bilinguals will need to recur to more explicit metacognitive procedures that involve rehearsal, recoding and imagery

so that acquisition can be induced. Hernandez et al. (2005) also state that evidence based on neurolinguistics offers further support to their stand on L2 acquisition.

In this vein, they begin by presenting event-related potential (ERP) evidence that shows that low proficient speakers tend to show a neuronal activation located on one area as their L2 is parasitic to their L1; second, recent neuro-imaging studies show that less proficiency means more widespread activation than the one observed for processing a strong L1; third, functional magnetic resonance imaging studies also show that similar lexical categories are related to certain brain area activation in the L1 and these same lexical categories differ in area of activation on the L2. These discoveries might be interpreted as indicative of the process of competition and entrenchment that lead to L2 acquisition since these patterns of activation show that late learners process linguistic loads on very specific cognitive mechanisms to process L2 items; these mechanisms are not necessarily the same that are activated for the processing of the L1.

The study and proposals in Hernandez et al. (2005) further the emergentist view of L2 acquisition in that a considerable load of cognitive processing is required for the learner to make the "competition" and diminish "entrenchment" between the L1 and the L2. They even take a further step than in the study by MacWhinney (2004) as they touch on short-term and long-term memory cognitive models as necessary mechanisms intervening in the goal of L2 acquisition and how they can be explained by neurological studies. The latter not only resembles what is discussed in information processing theory to explain L2 acquisition, but also relates to the main element under research in the present study.

However, the view of short-term memory (STM), and long-term memory (LTM) under Hernandez et al. (2005) is limited to a dichotomy in which LTM is the result of the information stored in STM. The concept and research in this study complements this view in that storing information is not the only cognitive process required to locate information, or L2 input in LTM. Although, this distinction is discussed in detail later in this chapter, the proposal in this investigation is that the utilization of both storing and processing functions of WM are necessary to acquire a second language, and achieve L2 automaticity; however, as explained later, the capacity in WM varies from individual to individual and this has effects on the level of L2 attainment among late learners.

To sum up, this section has reviewed how information processing, the competition model, and connectionism converge in how a second language is acquired. Most importantly,

it was discussed how outstanding research studies, in the concurrence of emergentist approaches, set a justification for the relevance of this study. In this way, whether the competition model and connectionism derive from information processing or not, these three approaches enhance the emergentist vision on L2 acquisition: to achieve high levels of L2 automaticity, a considerable demand on the late L2 learner's cognition is imperative. However, emergentist approaches are limited in how they explain the development of automaticity on late L2 learners as well as on considering all the components of working memory in their acquisition. Such limitation will be discussed later to comprehend the relevance of the present research.

Nevertheless, before moving to the following section in this chapter, it is paramount to mention that emergentists prioritize input as the main element that triggers L2 acquisition. However, the stand I take in this research is that the learner's span in working memory capacity plays an essential role at determining the degree of proficiency, or automaticity of the L2 learner as it is necessary to engage in efficient processing of L2 information.

Therefore, although the main principles of the emergentist perspective to L2 acquisition are: 1) a focus on linguistic frequency (O'Grady, 2007; MacWhinney, 2006), 2) ability to detect/internalize patterns from the input (Ellis, 1998, 1999, 2007, 2008, 2009; Ellis & Larsen-Freeman, 2006; O'Grady, 2008; MacWhinney, 2006), and 3) the formation of mental networks in charge of linguistic use that increase its activation through usage/practice (O'Grady, 2008; MacWhinney, 2006), the present study intends to prove that working memory capacity is key in the learner's engagement of this procedure. What is more, if WMC plays a role in the former procedure that emergentists propose to explain L2 acquisition, then it should be intrinsically related to the achievement of automatization. Therefore, this study concentrates in finding the connection between WMC and L2 automaticity; being automaticity a referent of emergentist theory to describe learners' high L2 proficiency. Being so, the following subsection describes automaticity in light of emergentist theory.

2.1.1.4 The concept of second language automaticity within emergentist theory

As it has been previously discussed, emergentist research focuses on exploring the development of automaticity in the L2. Also as mentioned before, such a concept is key for the present study as I propose that working memory capacity has an effect in the goal of L2 automaticity. As it will explained in more detail in sections 2.2.1 and 2.2.2, working memory and WM capacity need to be related to L2 automaticity since WM is crucially involved in

learning processes such as new vocabulary acquisition, and even L2 novel words. What is more, WMC highly correlates with the individuals' degree of success in tasks that require higher-order cognitive skills such as reading comprehension and the parsing of ambiguous sentences.

In such manner, given that achieving L2 automaticity is the result of successfully placing L2 knowledge in long-term memory (or the "learning" of an L2), then it can be assumed that the individual's WMC determines the degree of success in attaining L2 automaticity. However, the present work makes said assumption considering the premise that WMC is involved in the development of L2 automaticity from beginning to intermediate stages since WMC is more present in developmental stages of the L2 learning process. Once learning, or in this case, L2 automaticity is achieved, the load on WMC should be less or load-free to be used on any other type of situation that requires the processing of new information and the involvement of higher-order cognitive skills. These premises lead to the discussion of L2 automaticity in the emergentist literature in the current section; and, overall to its association with the research objectives of the present investigation which seek to explore the effect that WMC has in the development of L2 oral fluency; an aspect of the L2 that has been studied to measure L2 automaticity.

To begin with, "automaticity" is central for both emergentist theory and research; but its conceptualization is intricate, and therefore, is not easy to dissect when reviewing the emergentist literature. Given this, the concept of "automaticity" and/or "automatization" is challenging to define all at once. For this reason, it is necessary to approach it from various angles. Perhaps one of the most attested definitions of automaticity is the one provided by DeKeyser (2001) and also discussed in Kormos (2006) in which, rather than being defined, automaticity is characterized as "fast, parallel, effortless, capacity-free, unintentional, result of consistent practice, little interference from and with other processes, unconscious, always based on memory retrieval, does not benefit from further practice, error-free and flexible, strong production rule, no interference from working memory, and no correlation between the mean and the standard deviation in performance measures" (p. 39).

In addition, DeKeyser (1997) reports that the theory that best supports what automaticity entails is based on Anderson's (1983) ACT1 model of the human cognitive architecture, in which knowledge starts out as explicit to further become more automatized (p. 196-197). The latter is used to explain, in a general manner, how automaticity is developed. In spite of this, in terms of SLA, DeKeyser (2001) discusses that automaticity has

been limitedly researched, and that a special focus has been made to L2 fluency as phenomena that results from automatizing grammar and vocabulary (p.141). The latter is tremendously important for the present investigation as it analyses L2 oral fluency as an aspect that denotes automaticity.

However, there are studies that aim at analysing L2 automaticity from a different focus. That is the case of DeKeyser's (1997) study on the automatization of explicitly learned rules of morphosyntax in an L2. For this study, subjects were taught the same rules and practiced them. As a result, DeKeyser finds that learning morphosyntactic rules requires highly specific skills and that such skills are gradually developed over an extensive period of time which matches the learning curve that other cognitive skills follow to be acquired. His findings illustrate how the theory of automatization, in the emergentist perspective, consists of describing L2 acquisition as the development of any other complex cognitive skill, which goes from an attentionally controlled, effortful process to a more efficient, easy-to-retrieve process (Anderson, 1992; Hulstijn, 1990; McLaughlin & Heredia, 1996).

The last mentioned is vital for the present study as it sheds light on the fundamental axel here researched: working memory capacity might be the mechanism in DeKeyser's (1997) findings that allows the automatization of the L2. Perhaps, the findings of the current study can attest that, if WMC does play a role in the development of L2 automaticity, then L2 acquisition is the result of the evolution in attentionally controlled process as stated by emergentists. Nevertheless, the view of automaticity by DeKeyser (2001) and how it has been undertaken by emergentist theorists in SLA needs to be reformulated.

Dekeyser's (2001) definition of automaticity having no interference from WM needs to be further addressed in light of more current views such as the one by Segalowitz and Hulstijn (2005) that stem from the empirical work of Segalowitz and Segalowitz (1993) (this definition will be further discussed in 2.3.3 as the concept that best adapts to the present investigation). Under this view, automaticity refers not only to speed in processing, but also to efficiency in retrieving and employing information accurately and/or appropriately. Taking the latter into account, the definition provided by DeKeyser (2001) requires to be expanded in at least three accounts: 1) automaticity does not only refer to speed effects on processing, 2) specifying that automaticity does not occur in a vacuum, and 3) the role that WMC plays to reach automaticity needs to be made more explicit.

First, the view that L2 automaticity only indicates a "fast" response has been contested by studies that have found that when the L2 learner becomes more proficient, their reaction times (RTs) in using the L2 and the variability in their responses decrease; but most importantly, their accuracy levels increase (cf. Segalowitz & Segalowitz, 1993). Therefore, it should be considered that automaticity also implies efficiency (demonstrated by faster response time and more accuracy in the L2), and not only rapid responses or immediate reactions using the L2. The latter opens the discussion for the second point on how Dekeyser's definition of automaticity has to be reshaped.

Noting that automaticity also involves more accurate responses and less variability in RTs indicates that central executive functions of a higher cognitive order are involved (e.g. reasoning, decision making, problem solving, etc.). Consequently, it is naïve to assume that L2 automaticity occurs without any preceding processes of a higher-order cognitive nature as the view by DeKeyser (2001) makes it appear. In this respect, Segalowitz and Hulstijn (2005) criticize the view that L2 automatization in processing occurs by default, or as a transformation from controlled to automatic processes as it has been interpreted by some cognitive psycholinguists such as DeKeyser himself (and, as it has been approached by emergentists).

Segalowitz and Hulstijn (2005) criticize this view based on the evidence from neurophysiological studies (e.g. Paradis, 1994; Squire & Knowlton, 2000; Ullman, 2001), which show that "explicit knowledge forms a prerequisite for implicit knowledge to come into existence rather than the claim that explicit knowledge transforms into explicit knowledge" (Segalowitz & Hulstijn, 2005, p. 378). Therefore, implicit knowledge or automatized knowledge (and skills) are preceded by the involvement of cognitive mechanisms that not only involve repetition and rehearsal of information, but that also require attentional effort, and access to higher-order skills stored in LTM; which takes me to the third point.

The role of WM and WMC in the development of L2 automaticity needs to be rephrased in DeKeyser's (2001) conceptualization, in which he addresses that there is "no interference" from WM. This claim seems to indicate that L2 automaticity occurs without the intervention of WM; or it is completely irrespective to WM. The latter needs to be re-worded in DeKeyser's (2001) definition given that there is evidence that demonstrates that WM functions are involved in the development of automaticity. As it will be properly addressed in section 2.2, the role of WMC has been proven to be crucial for learning because

it integrates higher-order skills, while it aids to retain information under attentional control (cf. Cowan, 2005); and, as it was mentioned, there are important connections between automatic processing and the integration of "higher level aspects of language" (e.g. using L2 grammatical structures in context, L2 reading comprehension, writing academic texts in the L2, etc.) (Hulstijn & Segalowitz, 2005; Gatbonton & Segalowitz, 1998, 2005). This leads to suppose that in order to achieve L2 automaticity, the L2Aer has to integrate cognitive resources that make it possible to keep L2 linguistic input under attentional control, while executing higher-order skills such as reasoning, pattern recognition, or any other that might be required to cope with the L2; the latter are functions of WM. Therefore, as a mechanism that encompasses storage and attentional control functions, WM should be addressed as the mechanism that precedes and leads to the development of automatic processing.

The latter is further supported by Segalowitz & Hulstijn (2005) who assert that "there is reason to believe that it is especially at the intermediate levels of syntactic, morphological, and phonological encoding/decoding, as well as at the lower levels of articulation and perception of acoustic or orthographic signals, that component processes can become automatic to a large extent" (p. 385). The latter combined with studies which prove that WM is paramount for L2 learning in ab initio and intermediate levels (e.g. Linck & Weiss, 2011, 2015; Santamarina & Sunderman, 2015; Serafini and Sanz, 2016; Sanz et al., 2014; Havik et al., 2009; Miyake & Friedman, 1998) are indications that automaticity is a subsequent result of the involvement of WM functions.

In sum, the definition of L2 automaticity needs to be understood in terms of the three notions that have been just explained so that the relation between WM and automaticity can be better established. Additionally, the relation between WM and automaticity is relevant for the present work as one of the research objectives here pursued is to understand how individual differences in WMC cause an effect on an aspect of the L2 that is directly linked to automatic processing among L2Aers; namely, L2 oral fluency. What follows is an account on how the concept of automaticity has been framed in the emergentist framework.

Emergentist theory and research aims at uncovering the trajectories on the process of second language acquisition that derive on L2 automaticity. Such a goal is reinforced by what Ellis (1998) accounts as one of the main principles for emergentists in SLA that "language is like the majority of complex systems which exists in nature and which empirically exhibit hierarchical structure […]. And as with these other systems, emergentists believe that the complexity of language emerges from relatively simple developmental processes being

exposed to a massive and complex environment (p. 644)". In other words, the interaction of these developmental processes and the exposure to linguistic information (and all the environmental aspects surrounding language) results, according to Ellis (1998), in the mastering of a language. These developmental processes can be interpreted as being composed of "simple learning mechanisms" (Ellis, 1998, p. 644) that are used to operate language data. It can be argued that, Ellis (1998) supports the notion that the attainment of language-specific modularity (or the acquisition of linguistic modules: phonology, phonology, morphology, syntax and semantics) comes from the automatization of the interaction between said learning mechanisms and the linguistic environment.

In terms of SLA, Ellis (1998) explains that for fluent language to take place, its users have to be exposed to several hours on the task of "language learning/acquisition" in order to extract a variety of utterances, which at the same time contain a wide variety of structures. In this line of thought, Ellis (1998) sustains that in order to acquire all this variety of L2 linguistic structures, language learners are required to apply simple learning processes (e.g. repetition, rehearsal, trial and error, etc.; all tasks related to working memory processing). Ellis' (1998) statements agree with MacWhinney's (2006) views in that L2 acquisition is the result of what emerges from language data and the social environment to which the learner is exposed; hence, the more L2 input and output, the higher the level of L2 attainment.

Broadly put, emergentism is a perspective to SLA that aims at explaining "how" language comes to be (Ellis, 1998; MacWhinney, 2006). In doing so, emergentism studies second language acquisition by exploring the effects of frequency of items on the input, by focusing on processes of perception, attention, and memory (Ellis, 1998, p. 656). Also, emergentism places special attention to computational models that imitate brain linguistic processes (MacWhinney, 2006, p.733). Although the latter descriptions of how emergentists conceive the language acquisition process provide a picture of the role that learning mechanisms play in the development of automaticity, the conceptualization, operation and relationship of these learning mechanisms with the automatization of linguistic processes needs to be specified.

Ellis' (1998) theorization of how language acquisition occurs barely mentions the role of working memory in the language learning process, while it seems to make the assumption that it is a "simple" learning mechanism that is used in the same manner by all individuals, in all tasks that have to do with language learning. Even when Ellis (1998) recognizes that attention and memory are important factors to process the linguistic input, he does not offer

a concrete explanation of how these aspects of cognition operate when the circumstance of language learning involves the acquisition of an L2 in adulthood. Ellis (1998) gives the idea that the frequency of exposure to the language combined with the employment of learning mechanisms (which as discussed, in emergentist theory, learning mechanisms are seen as "a given" and as mechanisms that function optimally) will result in the automatic processing of language, or automaticity; be it a native or a second language. The latter, particularly for learners who start the language process after puberty, is not necessarily a guaranteed outcome (cf. Schmidt, 2011).

In this sense, the view of L2 automaticity under the emergentist framework needs to be updated so that the role of WM can be better understood in the L2 acquisition process. As I mentioned before, there is empirical evidence that demonstrates that L2 automaticity does not only imply "fast" processing, but also qualitative (more efficient) processing (cf. Segalowitz & Segalowitz, 1993; Segalowitz & Hulstijn, 2005). One element that needs to be put under the lens for investigation is the claim by Ellis (1998) that rehearse and repetition lead to the development of L2 automaticity. As it will be further discussed in 2.2.2, rehearse and repetition only employ the storing functions of WM (which have been referred to as phonological working memory in the emergentist literature). Furthermore, as it has been pointed out, L2 automaticity cannot be reduced to the fast retrieval of L2 linguistic information; it also involves the retrieval of linguistic information that is adequate and accurate to the task at hand. Thus, emergentist views on rehearsing and repetition leading to automaticity are rather simplistic in the sense that they portray L2 automaticity as the result of a shift from phonological WM processing to the fast processing of L2 knowledge in LTM.

As it could be said that the functions of working memory are at the centre of emergentist theory, the view of WM and the individual differences in WMC among L2 learners is not well developed; it needs to be complemented with the findings of research as the one here conducted. As it has been mentioned, emergentists assume that L2 linguistic input and output need to be perceived, paid attention to and processed through memory to be acquired. This assumption corresponds directly to the functions performed by working memory as conceptualized by Baddeley and Hitch (1974) and Baddeley (2000, 2003, 2017).

Therefore, if second language learners acquire their L2 by means of working memory processing, the result should be a highly proficient L2 attainment; which, in emergentist terms is the result of the learner's automatic processing (Anderson, 1992; Hulstijn, 1990;

McLaughlin & Heredia, 1996) or reaching of L2 automaticity (DeKeyser, 1997, 2001; Schmidt, 1992). However, even when emergentists do place relevance on the role that the attentional (central executive) functions of WM play to develop L2 automaticity, their empirical efforts to support these claims have been limited to study only the storage functions of phonological WM (cf. Ellis & Sinclair, 1996); hence, leaving aside the functions of the central executive as well as the individual differences in WMC among L2Aers.

Thus, a careful consideration, treatment, and study of working memory is crucial to understand cognitive processes of L2 acquisition, and attest the claims made by advocates of emergentism. Of particular importance for the present investigation is to analyse to what extent WM and L2 automaticity are interrelated. What is more, there is a need to point out that L2 automaticity is not equally attained by all late second language learners (L2Aers). Consequently, working memory needs to be treated and studied as an individual difference, rather than as a default, obligatory, available-to-all, mechanism that allows L2 automaticity.

In the present research, L2 automaticity is measured in the L2 oral fluency of Spanish-speaking L2Aers of English, given that L2 oral fluency is an aspect that results from automaticity (Kormos, 2006). A detailed discussion on second language oral fluency as an aspect of L2 automaticity, and its connection with working memory capacity, however, are presented in detail in 2.2.4. in this chapter. Thus far though, this subsection has made the relevance of L2 automaticity explicit in emergentist research and theory. It has also served to have a better notion of how working memory and L2 automaticity are referred to by emergentist proponents; as it has aimed at justifying why the present investigation explores individual differences in WM among L2Aers and its effects on their levels of automaticity (measured through oral fluency).

Following the latter, the next subsection aims at framing the present study on the grounds of emergentist theory, and its contribution to understand L2 automaticity through the exploration of working memory capacity effects on oral fluency.

2.1.1.5 Emergentist research in aspects of L2 automaticity and its connection with working memory

From an emergentist perspective, from the bases of its theory to its current research, working memory has been approached as an influential aspect for L2 acquisition. In terms of theory, authors that advocate for an emergentist stance on SLA, such as Ellis (2008), acknowledge that L2 learning relies on “perception” to turn input into “intake”. In this

instance, “intake” is defined as “that subset of input that actually gets in and that the learner utilizes in some way” (Ellis, 2008, p. 238). Hence, the latter is said to depend on perception, which in itself works by means of higher-order cognitive mechanisms such as working memory (Ellis, 2008). The notion of input turning into intake in L2 processing is also developed by McLaughlin and Heredia (1996) who explain that input “goes through” a “controlled processing” mechanism in order to become internalized, or “automatized”.

This controlled-to-automatic processing is also highly related with how working memory operates. The notion of a controlled and an automatic processing in L2 acquisition can be traced back to Anderson (1992) who links controlled processing to the concept of declarative knowledge, which he defines as a cognitive mechanism responsible for converting information into a “rule form” type of L2 production. The latter would be characterized by Anderson (1992) as knowledge that has become procedural. Once declarative knowledge has been turned into procedural knowledge, it has become easier to activate, or to retrieve. Consequently, the dichotomy of declarative and procedural knowledge explains the dichotomy of controlled to automatic processing in emergentist views such as in McLaughlin and Heredia (1996).

As discussed previously under the information processing approach, the declarative and procedural (D/P) dichotomy has been illustrated previously in the work of Ullman (2004). The root concept of D/P knowledge in emergentist theory is key to understand its connection with the concept of working memory. This connection is made clearer in Hulstijn (1990) as he points out that two main stages of L2 acquisition are the declarative stage and the procedural stage. Hulstijn explains that on the former, a rehearsal of items occurs in working memory and this involves slow processes; on the latter, “knowledge is converted from declarative to procedural form” (p. 31).

Thus, Hulstijn states that there is a process of compilation which divides into “composition” and “proceduralization”. The first refers to “single” production with a certain “speed-up” effect, as the author describes it. The second refers to the creation of new “products” from old products, which do not need to be retrieved from working memory anymore. The illustration provided by Hulstijn (1990) in terms of central concepts such as controlled-automatic processing and D/P knowledge serves to disentangle the role that WM ultimately plays in emergentism since it determines in what exact part of the learning process WM intervenes.

The latter is a valuable assumption since it suggests that WM is crucial for acquisition from beginning to intermediate stages in which the lexical and morpho-syntactical stages are being grasped, formed, and developed by L2 learners. This will be furthered discussed in the analysis of the results obtained in the linear regression study performed on the data of the intermediate group of participants in the present research. Also, Hulstijn's (1990) modelling and description of WM in L2 acquisition reinforces the notion that variation in attainment might largely depend on how much L2 linguistic information is processed through the L2Aer's working memory capacity.

Furthermore, in terms of research, emergentists have conducted studies that directly explore the relation between working memory and L2 acquisition. An example is the paper by Ellis and Sinclair (1996) in which it is proposed that individuals who have a better phonological working memory, also have better results in retention of language in long-term memory. As a consequence, the authors suggest that phonological WM has an effect on the acquisition of syntax. The authors propose that a considerable body of evidence shows that phonological factors are involved in (particularly productive) vocabulary acquisition" (p.236). They support the latter provided that in their study, a group of L2 learners performing repetition of new L2 utterances scored significantly better than other two groups that did not use repetition. The authors used a vocabulary recall task to measure phonological WM, which does not involve central executive functions of WM.

For this study, the authors conclude that the rehearsing in phonological short-term memory "provides a wide range of language-learning advantages" (p. 243). They conclude that given their findings high proficiency is achieved by means of activating WM; in this case by means of phonological activation. It is clear that Ellis and Sinclair (1996) see WM as a cognitive mechanism that is intertwined with short term memory. At some points even, WM gets used as a synonym of STM. This interpretation of working memory is problematic because it reduces the functions of WM to the temporary storage of information as it ignores the scope that its central executive functions may have in L2 acquisition; the latter is one of the reasons why there is a need for the present research.

Although further details on the definition of working memory are provided later in this section, seeing working memory as a synonym of short-term memory capacity as in Ellis and Sinclair (1998) creates confusion and obscures the comprehension of second language processing at the cognitive level. First, the fact that participants were asked to repeat second language items poses a problem in terms of methodology as only short-term memory is being

tested. Second, it implies that any linguistic information that is repeated is subsequently proceduralized or automatized in LTM by default.

The present study makes a careful consideration on the measure of working memory capacity that does not limit it to a simple STM mechanism. In this sense, the study and findings by Ellis and Sinclair (1998) can be advanced and expanded to a notion of working memory that is more precise and can better expose why some learners do not automatize second language input.

To conclude the section dedicated to the emergentist perspective of SLA in this chapter, it is important to mention that the previous discussion raises an opportunity to gain a clearer ground on the examination of working memory capacity and its role on the development of L2 automaticity. This is attributable to what has been discussed in this section with respect to how, in spite of the recognition of WM as a mechanism of both storage and higher-order thought processing (Hulstijn, 1990) in emergentist theory, its role has not been analysed as such in empirical studies.

Therefore, in the first major segment of this section, it has been discussed how emergentist theory seeks to reveal the processes that lead to second language acquisition on late learners. It has been mentioned that although there are slight differences in the approaches that conform the emergentist perspective to SLA, they all coincide on investigating how L2 automaticity is reached. Furthermore, the concept of automaticity has been reviewed; and, its linkage to working memory capacity has been established. In such manner, this section serves to situate the objective of the present research at the core of specifying and extending the knowledge of WMC as a cognitive mechanism that might condition the degree of L2 automaticity reached by second language late learners.

2.1.2 The generative approach to SLA

One more objective of the present study is to analyse the effects of working memory capacity on the processing of lexicon and sentences of second language late learners (L2Aers). For this reason, I consider that to comprehend the relevance of exploring how L2Aers process L2 sentences and the possibility of working memory capacity as an influencing mechanism, it is fundamental to discuss the study of lexical and sentence processing within the scope of the generative perspective to SLA.

This cognitive perspective has extensively contributed to set the ground on conceiving how late learners reach high levels of second language acquisition. In particular, research with a generative approach has been key to comprehend how L2 grammars are constituted. As I will describe in the following section, research with a generative approach in SLA has proven that late learners are able to accurately distinguish morphosyntactic L2 features that are not necessarily available in their L1 grammars, and that are not accessible and/or comprehensible through the L2 input. However, the findings of generative research are still insufficient to explain why there is variability in the grammatical attainment of L2Aers. Therefore, the second objective of the present research is to explore if WMC has an effect on the L2 grammar of late learners.

However, in order to have a better understanding of the generative approach to SLA, the following sections are dedicated to provide 1) a general overview of the bases of the generative approach to SLA, 2) its scope in research concerning L2 late learners' acquisition, and 3) how research with a generative approach has explored WMC effects on L2Aers' grammars.

2.1.2.1 The bases of the generative perspective to second language acquisition

Firstly, the SLA generative perspective derives from Universal Grammar (UG) theory (White, 2007) as proposed by Noam Chomsky around the 1960s. Broadly explained, Chomsky's UG theory to language acquisition proposes that human beings are born with a biologically encrypted language faculty (White, 2003). This language faculty, or UG, is said to be "a genetically endowed blueprint" that "contains the linguistic information that is common to all human languages […]" (Rothman & Slabakova, 2017, p. 3).

In UG, the speaker's linguistic information is represented by what proponents of this perspective call "linguistic competence" (Juffs & Rodriguez, 2014; Tokowicz, 2014; White, 2003, 2007); which is a grammar that is composed of categories and features constrained by the types of operations that can be possible in a given language; namely, by principles and parameters by which grammars function (White, 2003). Hence, for generative proponents, grammars are determined by a series of principles, defined as a set of linguistic properties that are shared or "true" across all languages in the world (Rothman & Slabakova, 2017; White, 2003). Also, grammars are also constrained by parameters, defined as a set of language features that are specific to a language (Rothman & Slabakova, 2017; White, 2003) and are ultimately what make a language distinguishable from another language.

One of the strongest arguments to sustain UG theory is the logical problem of language acquisition or the poverty of stimulus argument (White, 2003, 2007; Rothman & Slabakova, 2017). As White (2007) points out, 'there is a mismatch between the input that children are exposed to and their ultimate attainment (p. 37)'. This is a relevant aspect for UG theorists as it is proof that language acquisition is indeed an innate ability since children are able to attain a fully developed grammar with relative ease, rapidity, without instruction, and at a fairly early age (White, 2003).

These UG bases provide the antecedent for the generative approach to L2 acquisition. This approach or perspective is also known in the literature as innatist (O'Grady, 2007), or nativist (Hawkins, 2004). In this paper, this framework will be referred to as generative (Juffs, 2005, White, 2003, 2007) since authors and researchers that are key to the research and literature presented here refer to it as such (Rothman & Slabakova, 2017). Additionally, as previously stated, the purpose of the present thesis is not to favour any stance as it is not to adhere to any controversy in terms of terminology.

Proponents of SLA generative theory, such as Lydia White (2003, 2007), argue that late L2 acquirers face the same task that L1 acquirers do: they aim at attaining a linguistic competence; in their particular case, an L2 linguistic competence. This second language competence has been referred to as interlanguage (IL) competence. According to its proponents, this IL competence might be constrained by the same underlying principles of L1 competence (White, 2003, 2007). This notion of UG being linked to L2 acquisition started to be discussed in the 1980s according to White (2003). Nonetheless, it could be argued that the pioneering work on IL by Selinker (1972), who coined the term, as well as that of Corder (1971) established the grounds for the study of L2 acquisition in relation to UG theory; ultimately, their work served to comprehend the systematicity and grammar-governed characteristics involved in L2 learners' languages. As a consequence, a series of concepts were developed in terms of the type of access that L2 acquirers have to UG.

Some theorists proposed that L2 acquirers no longer had the linguistic mechanisms that are available to the L1 acquirer in their initial state (White, 2003), and that they could only acquire their L2 via the grammar of their L1; this view was denominated the no access hypothesis (White, 2003). Some others proposed that L2 acquirers have access to UG and that their IL is constrained by UG principles (White, 2003), which is known as the direct access hypothesis. Furthermore, some researchers considered that both the L1 grammar and access to UG were involved in the L2 acquisition process. This means that these authors considered

that L2 acquirers started their process relying on their L1 to further engage in restructuring their IL, resetting their parameters, as they were more exposed to L2 input (White, 2003); this view is known as the indirect access hypothesis. However, all these concepts turned out to be problematic and direct and indirect were later replaced by full and partial access respectively (White, 2003).

Nevertheless, the focus on access has shifted towards examining the nature of IL grammar and exploring if these IL grammars show characteristics of L1, or natural languages (White, 2003). It is better explained when Lydia White (2003) points out that "if it turns out that the L2 learner acquires abstract properties that could not have been induced from the input, this is strongly indicative that principles of UG constrain IL grammars parallel to the situation in L1 acquisition (p. 22)." Hence, generative researchers studying L2 acquisition are seeking to know how IL grammars are constrained by UG. If indeed IL or L2 grammars obey UG principles, then L2Aers should display similar grammatical behaviours and patterns as native speakers of a language.

Further evidence for this is pointed out by White (2007), who highlights the fact that L2 learners of English get to know the restrictions on wh-movement and this is dissimilar to their L1 grammar. Therefore, there are UG constraints governing interlanguage competence as well. The author gives a further explanation on the basic elements that represent UG constraints: principles and parameters. First, principles refer to constraints that the individual possesses, such as wh-movement; second, parameters refer to the differences across languages around the realization of principles.

Furthermore, the author proposes that from a generative perspective to IL competence, evidence needs to be drawn from production, comprehension, and intuitional data. According to White (2007) common misunderstandings regarding generative SLA research involve a) the scope of the theory, b) lack of native like success in L2, c) transfer, and d) methodology. The author emphasizes, in this regard, that the theory is not meant to explain all aspects of L2 acquisition; rather it looks forward to describing and explaining IL competence, or the L2 learner's mental grammar.

Also relevant to the study of SLA from a generative perspective has been the fairly recent changes in the theory brought by the Minimalist Program (Chomsky, 1995). Even though, a fine-grained empirical analysis and description of grammatical properties has been a constant in UG based research, for both L1 and L2 acquisition, the Minimalist Program has

contributed to have a deeper and more complete understanding of the representations that underlie the features of the grammar (cf. Liceras et al., 2008; Rothman & Slabakova, 2017). For example, and of particular importance for the current investigation, is the tenet in the Minimalist Program that grammatical features can be categorised into interpretable and uninterpretable. Interpretable features, on the one hand, are those that are useful to convey the meaning of a sentence; an example of an interpretable feature is the morpheme -s for regular plural forms in English as it overtly indicates the number of the noun.

On the other hand, uninterpretable features are those that only possess a grammatical value and do not contribute to the sentence meaning; an instance of an uninterpretable feature can be the grammatical gender for inanimate nouns in Spanish since it only contributes to the agreement between the determiner and the noun at the noun phrase level. This take is relevant to the present investigation given that I have focused on the object resumptive pronoun as an uninterpretable feature to analyse if individual differences in working memory capacity are related to the L2 learners' reactions to features that are not comprehensible by the frequent encounter and/or the information provided by the L2 input, and that do not have the same configuration in the L1 grammar.

As stated by Rothman and Slabakova (2019), the changes in generative linguistic theory deriving from the Minimalist Program have resulted in further developments in terms of SLA. One of these developments has been the reframing of the views that sustain that L2 learners only have partial access to the UG such as the Interpretability Hypothesis (which will be discussed later in relation to the focus of the research here conducted) (cf. Hawkins & Hattori, 2006; Tsimpli & Dimitrakopoulou, 2007). Proponents of said hypothesis sustain that L2 learners' access to UG only allows them to either acquire and/or comprehend interpretable features; and thus, cannot access those features that are uninterpretable. However, this hypothesis has been contested by some empirical findings in research with a generative approach as the ones that will be described in the next subsection (2.1.2.2) (e.g. White & Juffs, 1998; de Garavito, 2006). The latter is useful to understand the lengths that research under a generative perspective has gone to comprehend what features of the target grammar are attainable and/or available for the L2 learner; what is more, to determine the relation between the L2 learners' competence and the access to those grammatical features that are not frequently encountered and/or comprehensible by the overt characteristics of the L2 input.

To sum up, in this section a general account has been given on the bases that comprise generative approaches in SLA. The idea that L2 acquisition processes are constrained by UG access remains a main focus of study in the SLA field. The focus of SLA generative approaches to explain and describe interlanguage or L2 competence is relevant to the present study given that their findings are germane to better understand the aspects that underlie L2 grammars.

2.1.2.2 SLA research with a generative approach on late learners

As it was formerly mentioned, research with an SLA generative approach has been concerned with studying how L2 grammars are constituted. Thus, this section has the purpose of describing the findings of some research with a generative approach that has concentrated on studying late learners' L2 grammars. These studies are relevant as they examine aspects of the grammar that are not accessible through the L2 input (thus not likely to be explained by the claims of emergentism). Furthermore, the studies here described set the ground to understand the need to explore L2 grammars in light of the L2Aer's individual differences (IDs) in WMC, which can contribute to shed light on the causes of grammatical variability among this subset of learners.

In this vein, the work by White and Juffs (1998) is worth paying attention for its focus on wh-movement and subjacency (see Chomsky, 1973, 1977) among late acquirers. White and Juffs (1998) recorded reaction times in a grammaticality judgment task containing grammatical and ungrammatical sentences with the wh-movement feature (e.g. "Who did the park ranger know__ followed the deer" extracted from White & Juffs, 1998). Their participants included Chinese-speaking learners of English and native speakers of English; among the Chinese-speaking participants, a group was denominated a "better" group since they were living in an English-speaking country.

The authors find that both groups of English learners were able to distinguish grammatical from ungrammatical wh-movement, and thus reacted to sentences in a similar manner to the native speaker group. With these findings, White and Juffs (1998) challenge the notion that if the L1 does not contain a feature such as wh-movement, then speakers of that language will not be able to acquire subjacency when learning an L2 that does have overt wh-movement ruled by subjacency (cf. Aoun & Li, 1993; Huang, 1982; Xu, 1990). Although White and Juffs (1998) suggest that their results imply that UG is available for learners since they are able to accurately distinguish a feature that is not part of their L1 repertoire (wh-movement and subjacency are absent in Chinese), the main contribution of these findings to

the present research work is that L2 learners are able to reach knowledge of the L2 that is not available in their L1, and that it is not necessarily acquirable and/or explicit in the L2 input.

However, these findings do not explain why certain second language learners are able to operate such complex L2 linguistic features such as subjacency and wh-movement, and though a minority, still a few learners do not seem to distinguish this feature. This poses an opportunity for the present research to further analyse the acceptability of resumptive pronouns among late Spanish-speaking learners of English and correlate it with their capacity in working memory, as IDs in WMC might explain the variation in grammatical knowledge on the subset of learners that started their acquisition process after puberty.

Another study with a generative approach to SLA is the work by de Garavito (2006), which analyses the acquisition of the clitic doubled noun phrase in Spanish. This study is especially relevant since, as in the present thesis work, it concentrates on late bilinguals as in the White and Juffs' (1998) research study. Given this, de Garavito (2006) proposes that if Spanish speakers in three different language circumstances (monolinguals, early bilinguals and late bilinguals) are able to detect ungrammaticality regarding the dative clitic doubling in Spanish, then an explanation for L2 acquisition based solely on patterns from input does not suffice. In this manner, de Garavito states that if late bilinguals, or late learners of Spanish are able to detect these L2 features, this would serve as proof for the access to an innate mechanism, which supports UG principles in SLA.

For this study, participants took a written grammatical judgment task test consisting of 82 sentences, 45 of which were grammatical and 37 were grammatical. Careful attention was paid on selecting nouns for the task in terms of [+human] and [-human] properties. The results indicate that all three groups of participants were able to make a distinction between grammatical and ungrammatical sentences. The author explains that this may be explained by the exposure of this group to formal instruction of Spanish as an L2. The results in de Garavito's (2006) study, in particular for the analysis here developed, indicate that late learners are able to accurately distinguish when complex linguistic features such as dative double clitics of Spanish as a second language are grammatical.

The findings in de Garavito (2006) serve as proof to demonstrate that late learners can successfully recognize properties of the grammar that would be classified as uninterpretable under the Minimalist Program (Chomsky, 1995), and thus confirm that the L2Aers' competence is not restricted to operate with grammatical features that are interpretable as

proponents of the Interpretability Hypothesis would claim (cf. Hawkins & Hattori, 2006; Tsimpli & Dimitrakopoulou, 2007) (a discussion on the Interpretability Hypothesis is included in 2.4.2). As in de Garavito's (2006) study, the grammatical property under investigation in the present work is an uninterpretable feature, namely, object resumptives in English as an L2. As it is important to understand how late learners react to this type of grammatical features to further attest if their L2 grammatical knowledge is limited to access features that are not overtly comprehensible by the L2 input, it is as crucial to understand if IDs in WMC cause for L2Aers to have variable reactions to these sort of L2 grammatical properties. De Garavito's (2006) findings do not explain why only some late learners are able to accurately distinguish double clitics in Spanish.

Therefore, neither does the de Garavito (2006) study explain the particular circumstances of L2 learning of these late learners (except for a general comment on their extensive exposure to the L2), nor does it explain why not all participants reached the same results. As de Garavito (2006) comments, a significant number of late learners were able to distinguish the complex feature under L2 processing analysis, but the fact that "most" not "all" participants accurately detected the feature raises the need for a study of IDs in WMC. The latter is the same observation that I made to the White and Juffs (1998) study.

In light of the latter, Juffs and Rodriguez (2014) conclude that early research on L2 competence and access to UG has led to findings that indicate that L2 learners are able to process uninterpretable features such as subjacency as illustrated in White and Juffs (1998) previously discussed. In this line, Juffs and Rodriguez (2014) point out that in spite of the claim that L2 learners have a major difficulty parsing subjacency "in very specific contexts when a wh-filler-gap relationship had to be reanalysed" (p.135), recent research shows that L2 learners are matching the knowledge shown by native speakers on "abstract grammatical properties of language and concomitant deeper structural processing" (p.137). Nonetheless, and without meaning to sound repetitive, an explanation for the causes of the variation in L2 grammars among L2Aers is necessary and not yet found in studies with a generative approach to SLA.

2.1.2.3 The relationship between the study of L2 grammar and working memory capacity within the SLA generative approach

From a generative perspective, working memory has been studied as an individual difference that influences L2 grammatical behaviours and outcomes (cf. Juffs & Rodriguez, 2014). The latter can be exemplified in the significant work by Juffs (2005) which explores the

connection between working memory and the differences in reading times of sentences containing wh-extraction among late second language learners. Over this study, Juffs (2005) does not find a significant correlation between WM and reading times among his subjects; however, he concludes that the design of the WM test might not have extracted measures of WMC per se.

The author points out that a careful analysis has to be made on the type of test that is employed to measure L2 learners' capacity in working memory. In his study, Juffs (2005) uses a word span task and a reading span task which consisted of cards with words presented to the participants to be recalled in order by the participants. Juffs (2005) attributes the lack of correlation between WM measures and grammaticality of wh-movement due to the employment of older methods that might not be as accurate to obtain measure of WMC (Juffs & Rodriguez, 2014, p. 134; see Conway et al., 2005). In spite of the apparent inadequacy in the Juffs (2005) study to find how WMC influences L2 grammatical parsing, his work emphasizes that there are variations in how L2Aers react to properties such as wh-movement (Juffs & Rodriguez, 2014, p. 133). Therefore, studies such as Juffs' (2005) reiterate the justification that has been divulged in this thesis work: variation in L2 grammars might be due to individual differences in working memory.

Another study worth noting is the one by Dussias and Piñar (2010) whose work focuses on testing and contrasting native speakers and L2 learners on how they parse long distance wh-questions, the type of information considered in the processing of these sentences, and how cognitive mechanisms may intervene in their parsing decisions. Dussias and Piñar pay special attention to WMC effects on recovering from subject - object misparse of long distance wh-structures that usually show garden path effects (p. 445).

In order to do so, the authors utilized a reading span task, and "varied the plausibility of the wh-word as a potential gap filler in [their] target wh-subject and wh-object extraction structures" (p. 464). Dussias and Piñar find that higher WMC L2 learners behave similarly to the monolingual participants in terms of subject-extraction sentences, which represents more difficulty than object-extraction sentences. Thus, this study indicates that individual differences, such as WMC, do have an effect in complex sentence processing in L2 learners. The Dussias and Piñar (2010) study is worth mentioning at this point as it illustrates the effects that variability in WM capacity has on the parsing of sentences that pose a considerable challenge on the L2Aers given the distance caused by the wh- islands and wh-movement in English. Therefore, it is an empirical finding that demonstrates in a more exact

manner what aspects of the L2 are affected by the learners' WMC; while at the same time, it uses more updated methods to measure WMC. Nevertheless, it is still necessary to offer more evidence of the effects of WMC on the parsing of L2 grammatical properties with learners at different levels of proficiency, and with aspects that have to do with L2 production as in the current investigation.

Given the latter, it is of paramount importance to study the effects of WMC on late learners' grammars; nevertheless as observed in the previous study, the examination of the effects of WMC on L2 grammars should include groups of L2Aers of different levels of proficiency (not only on a single level of L2 proficiency as in the Dussias and Piñar (2010) study); and with more updated tools to measure WMC to avoid the complications found in Juffs (2005) with respect to the instruments employed to obtain WMC measures.

2.2 Working memory capacity (WMC)

In the previous section, I discussed the positions of emergentist and generative theory to SLA in terms of how late learners develop L2 automaticity and how their L2 grammars are constituted. Broadly put, both cognitive theories validate that L2Aers are able to reach high levels of proficiency as proven by their empirical data. On the one hand, emergentist empiricists have proven that L2Aers are capable of automatizing phonological and morphosyntactic features through detecting frequent patterns from the L2 input, and extensive repetition practice of these items.

On the other hand, generative researchers have identified that L2Aers can acquire complex features of the language (e.g. subjacency as in White & Juffs, 1998) which demonstrates that learners' L2 competence (White, 1997, 2003, 2007) contains deep linguistic knowledge; this deep knowledge is acquired beyond the information provided in the L2 input and the frequency of encounter. However, these empirical findings also indicate that there is variation in the degree of attainment of these advanced linguistic aspects of the L2; id est, automatized grammar or L2 automaticity and acquisition of complex underlying linguistic features of the L2.

Therefore, the proposal I make is that this variation in acquisition on late second language learners depends on individual differences in working memory capacity. This proposal agrees with (as it has been discussed in the previous section in this chapter) emergentist and generative researchers who have acknowledged the influence of WMC to

account for variations in acquisition among L2Aers. Based on their views, it can be interpreted that emergentists consider WM as a crucial mechanism for the automatization of the L2, and generativists consider that it is a mechanism that conditions L2 sentence processing. Therefore, both of these SLA theoretical frameworks adjudicate a determining role in attaining high levels of acquisition to late learners' working memory capacity.

All in all, the main focus of the present research is to put working memory at the centre of the cognitive procedures that either halt or allow high proficiency levels in L2Aers. As mentioned before, there is no intention to add to the debate of L2 acquisition being the result of L2 input and frequency or access to UG; the proposed angle in the present work is that L2Aers can detect patterns from L2 input and achieve L2 automatization as well as access UG to acquire complex L2 linguistic features by possessing a higher working memory capacity.

The latter responds to the assumption that L2Aers need to employ higher-order cognitive strategies to engage in the effortful process of second language learning and accomplish the goal of becoming proficient. In this sense, the assumption is made that second language late learners attend to higher-order cognitive mechanisms to cope with the decay of plasticity, which has been presupposed responsible for the impediment to acquire the L2 as the L1 on late learners. Consequently, the endeavour of learning a second language poses a high demand in cognition, in particular to L2Aers who have to develop superior cognitive strategies to master a second language.

For this reason, the purpose of this section is to have an extended discussion on working memory and working memory capacity to better understand why the current research study seeks to analyse its effects on the acquisition of difficult and advanced linguistic aspects in late second language learners; in particular, on L2 oral fluency and acceptability of resumptive pronouns.

2.2.1 Defining working memory

The origins of working memory capacity can be found on the foundations of short-term memory (STM) that considered it as a "temporary memory store" (Baddeley, 2007, p. 31). Even further back in history a dichotomy in memory mechanisms was introduced by James (1890) who divided memory into primary and secondary; in which the first was in charge of immediate conscious knowledge, and the latter was in charge of permanent, long-held, remote knowledge (cf. Mizera, 2006).

However, this notion was altered by a more widely spread view in the midst of the twentieth century when the predominant perspective was that there was only one memory system. In this view of a single memory system, learning was considered as the result of association-forming and forgetting as an occurrence of interference between associations (cf. Baddeley, 2007, p. 32).

Nevertheless, this perspective shifted back to the retaking of a dichotomized view of memory that included a STM dependent on a long-term memory (LTM), which was able to hold longer-lasting neurochemical changes (Hebb, 1949). This view was sustained by Peterson and Peterson (1959) who added that loss of information was explained by a prevention of rehearsal; and thus, the role of interference was rejected (Peterson & Peterson, 1959).

Though, this assumption brought controversy among those that were interested in investigating memory and cognition since it was not clear if STM was a dependent or a separate component of LTM (Baddeley, 2007; Mizera, 2006). In this vein, studies such as the one by Melton (1963), employing digit span tasks, claimed that STM was a dependent component of LTM.

Further experimentation performed by Baddeley and Warrington (1970) and Shallice and Warrington (1970) led to confirm the position of two separate systems. In their studies with amnesic patients, these authors found that patients were able to perform well in STM tasks even when affected in LTM; and, patients with defective STM were able to perform well on LTM tasks (Baddeley, 2007, p. 34).

Nevertheless, more research needed to be performed to shed light on the confusion and controversy over these two memory systems (Baddeley, 2007). Therefore, the latter derived in the Atkinson and Shiffrin (1968) modal model that proposed three types of memory systems. The first was a sensory register that was in charge of receiving input from the environment; the second was a short-term store (STS) temporary working memory store that was responsible for controlling information in terms of rehearsal, coding, decision-making, as this information is retrieved from a long-term store (LTS). The third type of memory in the Atkinson and Shiffrin's model was a storage of permanent information. However, Baddeley (2007, p. 36) reports that two main problems were found in this model.

The first problem in the Atkinson and Shiffrin model was the proposal that information being held for a certain length of time in the STS was enough to pack the information, so that it could be held in LTS. Baddeley (2007) states that experiments did not point in that direction; particularly, those performed by Craik and Lockhart (1972). These authors found that factors such as depth of meaningful processing and coding were more critical for retention for information to be stored in long-term store than mere length of time (Baddeley, 1983; Baddeley, 2007; Mizera, 2006). A second problem with the modal model was what Shallice and Warrington (1970) found in amnesic patients; patients with a defective STS did not show problems with LTS tasks as proposed by Atkinson and Shiffrin (1968).

Consequently, this entanglement in views of the systems of memory led to the Baddeley and Hitch (1974) model that focused on multicomponents that shed light on how working memory was an interim mechanism of memory that was related to both short-term memory and long-term memory (Baddeley, 1983, 2007). Thus, working memory started off as a hypothetical mechanism that challenged what was proposed by Atkinson and Shiffrin (1968). The proposal was that short-term store was not only a system of limited capacity that maxed its limit by mere loads of information, but that there was a working memory system that would sustain memory loads while coping with other cognitive tasks that were more related to long-term store (Baddeley, 1983, 2007).

Henceforth, Baddeley and Lewis (1984) conducted a study that required subjects to remember sequences of one to eight numbers while they were performing a task that asked them to decide if sentences accurately described the order of letters with which they were visually presented. The authors found that although memory load interfered and caused an overlap of processing, the task that was related to reasoning did not affect their levels of accuracy (Baddeley & Lewis, 1984). These results were crucial for the advocacy of a working memory system developed by Baddeley and Hitch (1974) that "comprises a central controller together with a number of interrelated subsystems" (Baddeley, 1983, p. 315).

Therefore, Baddeley and Hitch (1974) (from now on referred to as "B&H" model) introduced the notion of a multicomponent model of memory, which they referred to as working memory. Such multicomponent modal mechanism was defined as limited-capacity information storage that pairs such temporarily stored information with higher-order cognitive activities that require complex thought (Baddeley, 2007, p. 41). In this regard, this system is composed of a central executive that is sub-served by two storage systems: a

phonological loop and a visuospatial sketchpad (Baddeley & Hitch, 1974; Baddeley, 1983; 2000; 2003b; Baddeley, 2007).

In this view, the phonological loop is "assumed to be capable of holding speech-based and possibly purely acoustic information in a temporary store" (Baddeley, 2007, p. 42). In this sense, this type of phonological storage depends on rehearsal that is either overtly or covertly vocalized. Secondly, the visuospatial sketchpad functions similarly to the phonological loop, but this subsystem focuses on visual and spatial information (Baddeley, 1983, 2000, 2003b, 2007; Baddeley & Hitch, 1974).

The decision in the B&H model of working memory to treat these systems separately stems from empirical evidence that reveals a little interference effect between these two channels. This was found in studies that involve repetition tasks for which subjects repeat lists of digits. Results in these studies indicate that participants are able to accurately recall items that are visually presented to them (Mizera, 2006). These findings indicate that visual input is equally recalled as auditory input. For this reason, Hitch and Baddeley considered that both the phonological loop and the visuo-spatial sketchpad operate in the same way, but avail to a major component of working memory system.

Such major component in the B&H model is the central executive; and, it is said to carry most of the workload in working memory. This is due to the functions that is in charge of performing; these functions include coordinating information between the two aforementioned subsystems while directing and transforming attentional focus, manipulating newly received information, as well as situating and retrieving long-stored information from long-term memory (Baddeley & Hitch, 1974; Baddeley, 1983, 2000, 2003b, 2007). Nevertheless, Baddeley (2000) added a new component to the central executive component because he observed that it did not account for the storage of larger chunks of information (Baddeley, 2007).

Said addition to the B&H model was an episodic buffer; introduced as a component that served as a bridge or a connector between working memory and long-term memory. Baddeley (2000) adds this episodic buffer as a subcomponent that holds conscious awareness as it temporarily stores information that requires larger storage capacity (p. 421). Baddeley (2007) explains that this episodic buffer differs from long-term memory in that it grants access to LTM. Also, Baddeley mentions that with the addition of an episodic buffer to the B&H model, there is a closer proximity to other approaches that describe how working

memory operates (Baddeley, 2007, p. 51). A graphic representation of Baddeley's (2000) WM model can be seen in Figure 2-1 below. In this model, the episodic buffer is included as a subsystem that serves to store long chunks of information that are processed in the central executive of the WM and retrieved from LTM.

Other models of working memory can be found in the literature (Mizera, 2006); an example is the model developed by Engle, Kane and Tuholski (1999) who see working memory as a single memory unit that is free in domain, which is not particularly related to the processing of information, but it is related to general intelligence. The model by Engle et al. (1999) proposes that individual differences are more relevant factors to learning than a phonological loop or a memory span capacity (Mizera, 2006, p. 12). Another relevant model of WM is that of Cowan (1988) in which there is an emphasis on the role that LTM plays in the use of WM. Cowan (1988) proposes a model of WM in which there is a subset of LTM in "activated state" as well as a "smaller subset" of LTM that serves to retain information under the focus of attention (Cowan, 2014, p. 202)

According to Cowan (2014) the B&H model seems to put forward the idea that the components of WM operate as "separate boxes", and this does not suffice to understand the taxonomy under which these components operate. In the Cowan (1988) model, the input that is received through working memory does not pass through "filters" (or what can be interpreted as the slave components of WM in the B&H model) first, to then be processed in the central executive and retained in the episodic buffer; this taxonomy in the B&H model, according to Cowan (2014), is hierarchical. For Cowan (2014) all information entering WM activates at least a portion of LTM and its retained under focus of attention (which in Cowan's view is a mechanism under the domain of LTM). Furthermore, the Cowan (1988) model also emphasizes the notion of "sensory memory". In this sense, Cowan (1988) expands the view of the type of input that is encoded in WM as it includes tactile, musical, and all other types of sensory stimuli; which are out of the scope of the phonological loop and the visuospatial sketchpad modules in the B&H model. In Cowan's (1988) vision of WM, sensorial stimuli are also processed by WM, but they require to temporarily activate a portion of LTM to be maintained under focus of attention and processed in WM.

One more relevant conceptualization of WM is that of Miyake and Shah (1996) who, similar to Cowan (1988), envisioned WM functions as more domain-specific rather than domain-general as in the B&H model of WM. For Miyake and Shah (1996), the functions of the central executive in the B&H are not the same for all kinds of information. More

specifically, Miyake and Shah (1998) observed that there are mechanisms of WM that serve exclusively to temporarily store and process linguistic information, while other separate mechanisms of WM are used for the storage and processing of spatial input.

Thus, based on empirical data, Miyake and Shah (1998) propose that there is a verbal working memory and a spatial working memory; and therefore, different and specific resources of the central executive are deployed in each of these working memories to process information under these two domains. With this in mind, Miyake and Shah (1996) state that WM does not necessarily operate as a "one size fits all" type of mechanism, but rather as a cognitive resource in which certain elements specialize in verbal information, and some others in spatial information; however, these "subsystems" of WM do not share the same tools and/or function in the same manner (Miyake & Shah, 1996, p. 22).

These other views of WM are relevant to better understand what WM represents at a cognitive level and its role on learning processes. However, at some level, these models expand the notion of WM as a modular cognitive system, which as in the B&H model, is comprised of specific components that are useful to store information and allow the processing or deployment of higher order cognitive skills. The models by Cowan (1998) and Miyake and Shah (1996) are more useful, in my view, to further understand what constitutes the individual differences in WM capacity. Hence, I will use their findings and conceptualizations to define what aspects of WM are variable, and thus cause individual differences in WM capacity.

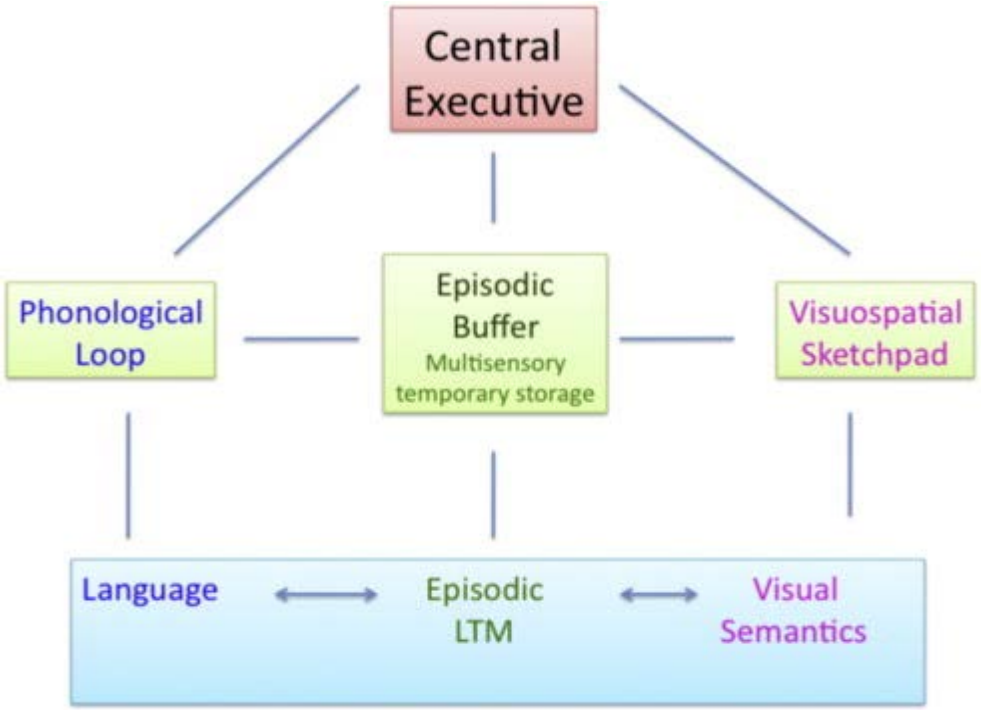


Figure 2-1 Baddeley's (2000) WM model.

In addition, there are some critics that consider that Baddeley's model leaves out syntactic processing (Caplan and Waters, 1999; Mizera, 2006). In a similar vein, Roberts and Gibson (2002) argue that more elements need to be added to explain language comprehension as this is a process that is not necessarily related to spans or loops. Other views, however, lean towards connectionist perspectives on the operation of WM.

In this line, Schneider (1999) states that WM is made up by a series of patterns of activation that consist of neuron networks, or units. In Schneider's view, capacity limits may be lessened as the user becomes more skilful or masterful at whatever activity he or she performs. In a similar vein, O'Reilly and Frank (2006) discuss that working memory is a system widely distributed in the brain; particularly, in the prefrontal cortex and the hippocampus, and it is an activated component that stems from long-term memory.

These views are similar to Ericsson and Delaney's (1999) model called "long-term working memory". In their view, working memory is not separated from long-term memory, and it is a set of mechanisms that permit the retrieval of information from LTM itself. However, as reported by Mizera (2006) these views and models are at least partially related to Baddeley's model of working memory. Perhaps, Schneider's, O'Reilly's, and Ericsson's views on a working memory that operates in long-term memory are the most differing views to Baddeley's model. Nevertheless, as Mizera (2006) discusses, if working memory is not

separated from LTM, its relevance to the aim of learning is diminished in comparison to Baddeley's model that highlights the role of working memory in learning (p. 14).

The emphasis put on learning in the B&H model with the addition of the episodic buffer (Baddeley, 2000) define the concept of working memory in the present research work. In light of the assumption, previously discussed, that acquiring a second language represents a major challenge for second language late learners, the role of working memory as described by Baddeley and Hitch (1974) and Baddeley (2000) might be central to deal with such a strenuous learning task. The identification of working memory as a mechanism that temporarily stores information as it processes such information with higher-order mechanisms of thought might accurately illustrate the required cognitive processing for the acquisition of linguistic information.

In other words, acquiring a second language involves developing a grammatical (from the phonetic to the syntactic level) and meaning knowledge that L2Aers can apply efficiently when necessary. This entails the processing of a very large corpus of information that requires to become a part of the L2 learner's long-term memory knowledge. Thus, if linguistic information of a second language is not processed by a mechanism that transforms it into LTM knowledge, then such information will not be processed, retrieved, and applied efficiently.

In this manner, the present subsection has pointed out the background theory that gave rise to the understanding and definition of working memory. Also, I briefly mentioned prominent research that led to the definition of working memory used in the research work that I am presenting. Finally, I presented a general rationale for the investigation of working memory capacity in the grounds of second language acquisition; and hence, in terms of the major role that it plays in the current study.

2.2.2 Defining working memory capacity

The previous section concentrated on how the conceptualization of working memory has evolved, from its origin to date. The conclusion that can be made is that given the characteristics of the aspects that comprise WM, its role in cognition and learning are crucial; in particular the model by Baddeley and Hitch (1974) with the integration of the episodic buffer by Baddeley (2000) sheds light on how WM operates in general learning processes. With this in mind, the present work seeks to shed light on how WM affects the learning of a

second language among adult learners (L2Aers). Nevertheless, it is necessary to emphasize that not all individuals use working memory in the same capacity; meaning, that while WM might be a determinant factor to learn new and complex information, not all individuals exploit the functions of this cognitive mechanism in the same manner. Thus, this section focuses on understanding what the capacity in WM is, and on the aspects of WM that create a difference in capacity from individual to individual.

Broadly speaking, Cowan (2005) explains that individuals' capacity in WM depends on their abilities to store and process information altogether when presented with new stimuli (p. 48). In this sense, the capacity in WM varies among individuals. Not only does capacity vary in how much information individuals can hold in phonological and or visual-spatial form (in the phonological loop and visuo-spatial sketchpad components of the BH-B model components), but also in how they are able to retain such information under attentional control, while retrieving the required knowledge from LTM. Although the model of WM suggested by Nelson Cowan (1998) slightly differs from that of the BH-B WM model, the author asserts that the main element of variation in WM capacity (WMC) is found on the focus of attention; the focus of attention is defined as a subcomponent of "[…] long-term memory, a subset of which was in activated state, and within that, a smaller subset of which was in the focus of attention" (Cowan, 2014, p. 202).

In this vein, the work of Cowan (2005, 2014, 2016) is useful to understand how the variability in WMC lies on functions that have to do with both the central executive and the episodic buffer. As mentioned before, for Cowan (2014), the degree of variability in the components of WM stem from the focus of attention as he sustains that the number of chunks of information under focus of attention in WM are variable; thus, those individuals that are able to retain more chunks of information under focus of attention are considered to have a higher WMC. Nevertheless, retaining more chunks of information under attentional focus goes beyond the function of storage in WM; it implies the temporary storage of informational items, while at the same time being able to retrieve knowledge from LTM to process the chunks of information. The latter corresponds to the functions of the central executive and the episodic buffer in WM.

Cowan (2014) also mentions that the capacity limits in WM have to do with age; the author explains that cognitive strategies and knowledge in LTM increase with age, and thus, adult learners may have a higher capacity in WM compared to that of children. This view is also shared by Gathercole and Alloway (2007) who explain that as children grow older, their

capacity in WM memory becomes “better”. Cowan (2014) adds that the latter view is controversial as age does not determine the differences in capacity in WM; and, as it can be observed in the present and other research involving the measurement of WM, there are differences in capacity even among individuals who are the same age (cf. Gathercole & Alloway, 2007; Daneman & Carpenter, 1980). There are, nonetheless, agreed differences in the characterization of what constitutes a higher working memory, or what Halford et al. (1998) calls a “good working memory”; said characteristics are, in broad terms, that the individual a) can hold enough information in mind, b) make relations among these items of information to solve a problem or accomplish a goal while simultaneously demonstrating speed in processing, c) retrieving knowledge, and d) executing their available cognitive strategies.

Moreover, Cowan (2005) adds that those individuals with a higher WMC are those who are able to “stay in task”, which means that learning as a result of processing stimuli through WM correlates with the individual’s “ability” to suppress distractions and who, in other words, are able to keep their focus on the goal of the task in which they are involved. Gathercole and Alloway (2007) also agree with this view and explain that “[individuals] need to continue to pay attention to what is being held in working memory if it is to persist over even short periods of time” (p. 6). Gathercole and Alloway (2007) also consider that a lower functioning and/or failure in WM are caused by distractions, trying to hold too much information at once, and engaging in demanding tasks. The latter circumstances are commonly encountered in learning contexts, daily routines, etc. Nonetheless, individuals with a higher WMC have been found to inhibit distractors, retain relevant information and have more efficient processing responses in spite of the complexity of the tasks or the load of information to be carried out. On the other hand, individuals with a lower WMC struggle to retain and process information under similar circumstances.

An example of how variability in WMC depends on both storage and central executive functions is the study by Conway, Cowan and Bunting (2001), in which they are trying to investigate the cocktail party phenomenon (cf. Moray, 1959) with regards to IDs in WMC. In their research, an operation span task (a discussion on how the distinct WMC span tasks work is offered in 2.2.3) was applied to a large number of undergraduate college students to measure their WMC. After this, 40 participants were classified into the categories of lower and higher WMC using a quartile selection process. The participants listened to two recorded messages simultaneously using headphones; one of the messages contained 330 monosyllabic words and was played 30 seconds prior to the second message that contained

300 monosyllabic words plus the participant's first name. In order to engage on the task, participants were instructed to repeat the words that they were listening on their right ear as accurately as possible and ignore the words in the message played on their left ear (the recording containing their first name).

The results show that most individuals with a higher WMC were able to report the words on the "relevant" message (the recording that was played on their right ear and were asked to focus and report on) with more accuracy. The lower WMC participants, on the other hand, reported to listen to their name and showed more difficulty to report the words from the recording played on their right ear. Conway et al. (2001) conclude that the results demonstrate that IDs in WMC reside on the individuals' ability to retain information under attentional focus; meaning, that the variability in WMC has to do with both storage and higher order cognitive processing. The latter is necessary to inhibit distractions, stay focused on task, and retain relevant information.

In addition, in an account of various studies involving individual differences in WMC, Engle (2002) emphasizes that the variability in WMC has to do with both storage and attentional control functions (central executive, episodic buffer, and storage components of WM). For Engle (2002), the empirical data and results obtained in studies using an antisaccade task (Kane et al., 2001), a stroop task (Kane & Engle, 2001), and the dichotic-listening task (Conway et al., 2001) serve as proof "[…] that performance on WM-capacity tasks is related to performance on other cognitive tasks primarily because of individual differences in executive attention" (p. 21).

In the study conducted by Kane et al., (2001), subjects with low and high WMC are asked to fix their sight in the centre of a visual display, while responding to information that is randomly presented to them on one side or the other of the display. Before they respond to the target information, a distraction is shown on the opposite side from where the target information appears; the participants are tested on the times that they move their eyes (saccade) to see the "cue" that is presented as a distraction. The results demonstrate that the individuals with low WMC tended to move their eyes in the direction of the distraction, while the high WMC individuals did not, and therefore, could provide more accurate responses when they were asked to report on the target information.

In the same vein, the Kane and Engle (2001) study consisted of having low and high WMC individuals to report on the ink colour of words displayed to them in a screen. The

words were actually names of colours and could or could not be congruent with their colour of ink; for example, the word “red” could have been presented in red ink (congruent), but the word “blue” had a red colour of ink (incongruent). The authors created three of types of stroop tasks; one with 0% incongruencies, one with 50% of incongruency, and finally one with 75% of incongruencies. The results indicate that whereas there were no significant differences between low and high span WM capacity participants on the Stroop tasks with a 0% and 50% congruency conditions, the low span WMC individuals made twice as many errors on the task containing 75% of incongruencies compared to those individuals with a high WMC.

Basing on these three studies, Engle (2002) offers empirical evidence to demonstrate that individual differences in WMC are not merely in terms of storage (on the items of information that an individual can hold in mind), but also in terms of central executive (or attentional control). The latter is demonstrated in the results shown in the studies above since they show that high span individuals outperform low span ones in tasks that require inhibition of distractors, focus on high-demand task goals, and ultimately accuracy as well as efficient cognitive processing. Thus, variability in WMC is found on both storage and processing functions of working memory. Engle adds that the memory (storage) and central executive functions of an individual’s WMC highly correlate with their performance on higher-order cognitive skills such as reading comprehension, verbal abilities, computational coding, etc. These higher-order cognitive skills include learning and attaining a high proficiency in a second language in adulthood. However, the individual differences in WMC might also explain why there is variability in levels of ultimate attainment among L2Aers.

The capacity limits as envisioned by Conway, Kane, & Engle (2003) and by Engle (2002) are characterized by the following main aspects as accounted in Cowan’s (2016) “Working Memory Capacity” shown in Figure 2-2 below.

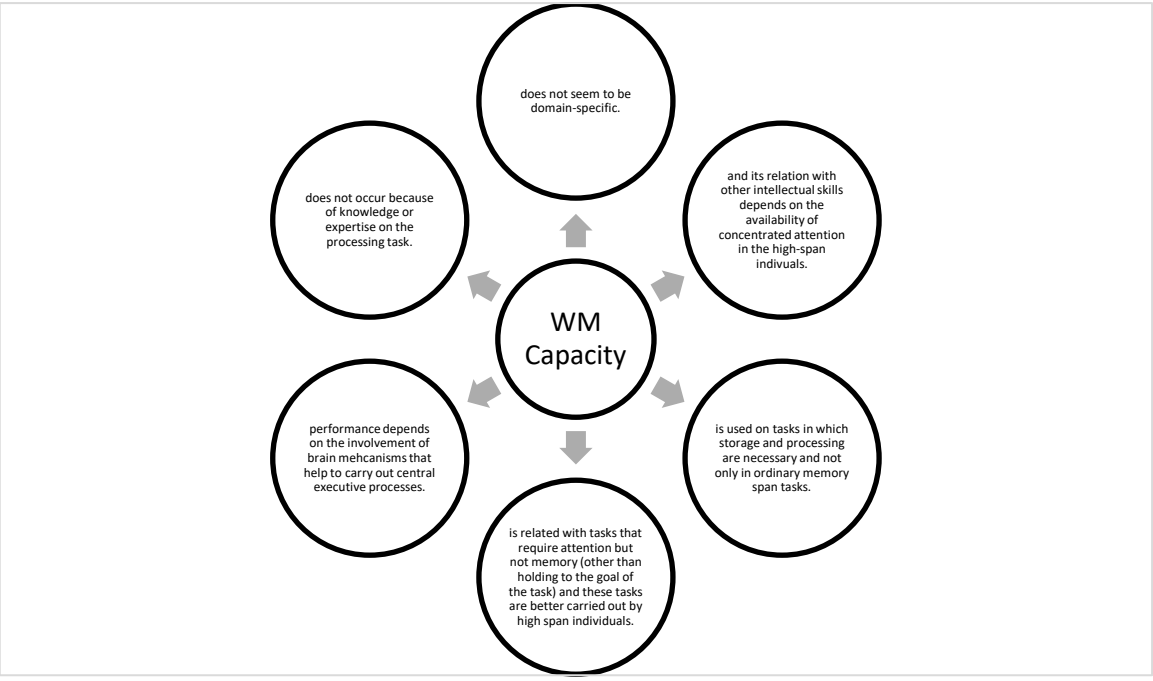


Figure 2-2 The characterization of WMC (based on the account provided by Cowan, 2016, pp. 51-52)

It is important to emphasize that the model by B&H-B (with the addition of the episodic buffer) is still relevant to understand how WM operates. However, the characteristics of what constitutes the capacity of WM are more thoroughly established by Cowan (2005), and Conway, Engle and Keane (2003); their description is more helpful to understand that the capacity in working memory depends on how individuals utilize both processing and storing demands of the stimuli with which they are presented.

Other views of limits in WMC include the views of Miyake (2001). In an account of multiple empirical findings, Miyake describes that capacity varies according to differences in both storage and processing demands, which work in a domain-specific manner. Nonetheless, the storage and processing functions of WM can extend to general-domain tasks. In terms of capacity limits, Miyake (2001) explains that there are views of WMC that are based on "the trade-offs between processing and storage" (p. 164). Under this model, the capacity in WM varies depending on the individual's skilfulness at certain processing tasks (e.g. reading academic articles) as this determines the amount of information that can be temporarily stored and maintained in WM; thus, the more processing ability at a certain task, the more opportunity there is to allocate a larger amount of resources for the retention of the target information.

Other model of WMC, according to Miyake (2001), "equals the capacity of the slave system (e.g., the phonological loop in the case of verbally oriented span tasks) plus the

efficiency of central executive functioning" (p.165) (cf. Engle, Kane & Tuholski, 1999). One more proposal of WMC has to do with "the task-switching hypothesis" (cf. Towse, Hitch, & Hutton, 1998, 2000). On this model of WMC, storage and processing are in an indirect relationship; meaning that the efficiency with which processing is executed determines the duration in which information items can be maintained in storage, and the degree up to which said items are likely to be forgotten. Lastly, Miyake (2001) mentions that one more important proposal on WMC is a model based on inhibition (cf. Hasher & Zachs, 1988), in which performance on WMC depends on "one's ability to inhibit irrelevant information" (Miyake, 2001, p. 165). The latter view is similar to that of Cowan (1988, 2005) as it also places capacity limits on the individual's ability to control their attention; measured by the degree in which they are able to inhibit interruptions and/or select relevant information.

For the purposes of the present work, it is important to emphasize that capacity limits in WM do not solely vary in terms of how many information items can be held in the memory systems (on either phonological, visual, spatial or any other type of stimulus/input), but also on the efficiency of the processing that individuals possess. As Miyake and Friedman (1998) point out "when WM is viewed as a computational arena where task-relevant processing and storage activities dynamically take place, it becomes clear that its capacity should be construed as operational capacity, taking into account both the nature of the information being maintained and the nature of the operations being applied" (p.342).

Therefore, understanding that individual differences in WMC are caused by variations in both storage and processing limits can lead to understand why certain learners are able to reach higher levels of L2 attainment, while others cannot. The acquisition of an L2 needs for adult learners to constantly retain verbal information, while they execute higher-order skills since they need to convey meanings, make communicative decisions, etc. Hence, the limitations that individuals may have on the storage and processing functions of WM may lead "to not just quantitative but also qualitative differences among learners" (Miyake & Friedman, 1998, p. 344).

With this in mind, the instruments that should be utilized to have a more accurate measure of the individuals' WMC should include tasks that pose a challenge in terms of both storage and processing; and, measurement of WMC should be made with tasks that are carefully designed to test how individuals deploy both of these cognitive functions either simultaneously or in closed proximity to one another. Thus, the next subsection is dedicated

to describing the relevance of utilizing WM span tasks that can reliably and validly provide a measure for the individuals' WMC.

In the case of the present work, the WMC of L2Aers needed to be explained in order to have a more concise picture of what aspects of this cognitive mechanism vary and to what extent these variations in WMC affect their L2 oral production and L2 grammatical abilities; namely, what characteristics in WMC shape the L2Aers' L2 oral fluency and acceptability of sentences with or without a resumptive condition.

2.2.3 Methods to measure working memory capacity

In terms of measuring working memory, the literature is constant and specific. First, the term "capacity" added to working memory refers to the "span" that individuals exhibit when engaging in a task that requires storage of information and its interaction with complex thought (Baddeley, 2003b) (or, higher-order processing of information that is stored in LTM). The literature usually refers to those who score better in working memory tasks as "higher WMC" individuals, whereas those who score lower are considered "lower WMC individuals" (Daneman & Carpenter, 1980; Conway, et al., 2005; Engle et al., 1999).

However, the question is: how do these scores or determinations come to be? How is it that they determine an individual's WMC? In order to answer these questions, a brief history of how WM tasks were designed are discussed in this subsection followed by an account of the current WMC tasks applied in research. In addition, I provide a brief account on the validity and reliability of the current tasks measuring working memory capacity. This discussion attempts to have a wider panorama of how working memory capacity has come to be more efficiently and accurately measured; and, how such instruments lead to have more reliable WMC data to analyse in the extent of acquiring L2 advanced features and skills.

To begin with, authors such as Baddeley (2003b) recognize the influence of the WM span tasks designed by Daneman and Carpenter (1980). In an attempt to obtain a more truthful measure of WM, Daneman and Carpenter (1980), proposed an alternate way to traditional measures, such as digit spans or word spans as they were more likely to measure STM. In such quest, the authors came up with a task that could correlate with individual differences in reading comprehension since as they observed; this was a task that involved WM usage.

Their rationale was based on assuming what is proposed in Baddeley and Hitch's (1974) model. The proposal in this model is that working memory has both processing and storage functions and that it serves "as a site for exciting processes and for storing the products of these processes" (Daneman & Carpenter, 1980, p. 450). Therefore, Daneman and Carpenter (1980) considered that a task on WM should correlate with measures of reading comprehension as this latter is a task that requires the storing of discourse, meaning, and structural information from preceding text to further interpreting the text that follows (p. 450).

With this in mind, the authors designed an instrument that could measure both processing and storage functions of WM. First, in order to measure the processing and storage components of the test, they included a component that measured sentence comprehension. Second, in order to measure storage, they included the requirement of maintaining and retrieving the final words of the sentences. Daneman and Carpenter (1980) denominated this measure instrument: reading span task (RST). The authors hypothesized that the scores obtained by individuals in their RST should correlate with reading comprehension scores, which they tested with the Verbal Scholastic Aptitude Test (SAT) (a standardized test commonly used in the United States to measure high school students' aptitudes prior their entering to university level).

Consequently, they ran two experiments; in their first experiment, their subjects were asked to undergo the RST, a reading comprehension test, and a word span test. The results for this first experiment show that the span test significantly correlates with the subject's Verbal SAT scores for reading comprehension ($r\ (18) = .59,\ p < .01$). For their second experiment, they incorporated a listening comprehension test which aimed at measuring spoken as well as written verbal material (p. 457).

For the design of their Listening Span Task (LST) they report to have modified their RST so that it could include both listening and reading elements. In addition, they included a modified span test for which subjects had to state if sentences were true or false. Daneman and Carpenter (1980) report that their second experiment confirms the findings of their first experiment in that the LST was very similar to RST at predicting reading comprehension. The authors add that subjects with higher listening spans were more accurate at spotting pronominal references ($r\ (19) = .67$ and $.72,\ p < .01$). Daneman and Carpenter conclude that span tasks, such as the RST and LST, are better reflectors of working memory capacity and

that there are individual variations in this “capacity” as there are also in reading comprehension (p.463).

Hence, the work of Daneman and Carpenter (1980) was crucial to continue investigating WMC in psychology and other scientific fields (Conway et al., 2005). Nevertheless, WM span tasks have evolved in both their design and their scoring measures (Conway, et al., 2005; Unsworth, et al., 2005). In this vein, Conway et al. (2005) gives an account on WM span tasks; their design, aims, optimal administration, validity and reliability, and scoring procedures.

First, they depart from the design introduced by Daneman and Carpenter (1980), passing through more adaptations such as the one by Turner and Engle (1989), which included fewer items than Daneman and Carpenter’s test, as it tested semantic and syntactic accuracy. Further adaptations on Turner and Engle’s test were the administration of span tasks to smaller groups to which input items were projected using overhead transparencies. This span task also allowed to keep a certain pace by presenting auditory items with a cassette recorder (Conway, et al., 2005, p. 772). Conway et al. (2005) report that current versions of the RST include modifications to the word to be remembered. It is more common for words to be unrelated to the sentence, so that this process is not a mere episodic recall (p. 772).

Later versions to the RST also include that instead of remembering a word, subjects remember a random isolated letter right after being presented with the sentence item (Conway et al., 2005; Kane et al., 2004). Other WM span tests used currently by researchers are the operation span task (OST) and the counting span task (CST) (Conway et al., 2005). In Turner and Engle’s (1989) span task, sentences are replaced by mathematical operations that include correct or incorrect operations such as (10/2) - 3 = 2 and (10/2) - 3 = 4 respectively (Conway et al., 2005, p. 772). Late versions of the OST also include a randomized presentation order of stimuli. This makes it differ from previous designs in which these stimuli were presented in an ascendant fashion (Conway, et al., 2005; Engle et al., 1992).

One more working memory span task, the counting span task consisted of counting geometrical shapes and remembering this counting number on a later recall (Conway et al., 2005). Some versions of this CST include the one by Engle et al. (1999) that added more complexity to the task by putting the target shapes among a variety of distractions, such as the same shape or colour. Additionally, this span task included a counting component of

items, which were presented in a randomized arrangement, with the next item display presented immediately after (Conway et al., 2005, p. 773).

In terms of the general objective that WM span tasks should fulfil, Conway et al. (2005) states that "the tasks are designed to force WM storage in the face of processing (or distraction), in order to engage executive attention processes (p. 773)". Given this, the application of these tests should follow certain recommendations to obtain the most optimal results possible (Conway et al., 2005). The first recommendation is to be attentive of how stimuli is presented and making sure that subjects do not rely on rehearsal so that the WM span task does not become a simple STM task. In addition, it is highly recommended that the span tasks are applied individually rather than in large groups; provided that large groups make it more difficult for the researcher to attest if subjects are indeed engaged in a processing task (Conway et al. 2005).

Lastly, Conway et al. (2005) recommend that the item size of the span tests should be sufficient or else there is a probability of ceiling effects. Regarding reliability and validity, WM span tasks, such as the ones formerly described, have proven to be coherent, stable, and effective. In this sense, Conway et al. (2005) mention that coefficient alphas and split-half correlations obtained from these tests show a consistency on the participants' responses in a range of .70 to .90; in which "0" indicates no reliability, and "1" indicates perfect reliability (p. 776).

Furthermore, the validity of these tests relies in their correlation with a varied repertoire of higher-order cognitive tasks such as language comprehension, reasoning, etc. (Conway et al., 2005). All in all, compared to other tasks that attempt to measure WM, span tasks accomplish their objective by conveying instruments that measure the processing of new information while performing complex cognitive skills.

It is important to mention that more adaptations and modernized versions of WM spans have been developed and are currently used in research. One milestone to WM span tasks is the adaptation to automated versions, such as the one developed by Unsworth et al. (2005) for the common operation span task. This version permitted subjects to answer the test with a computer mouse, and it also reported scores automatically. This version of the test was reported to have a reliability on consistency of answers of .78 in alpha measures (Unsworth et al., 2005).

Other advancements to span tasks were introduced by Redick et al. (2012) who developed automated versions of more span tasks such as the RST of Daneman and Carpenter (1980). The authors report that these automated versions allow to obtain scores of a vast number of subjects with an immediate report of their scorings. By the same token, Oswald et al. (2014) not only analysed the advantages of automated span tasks introduced by Unsworth et al. (2005) and Redick et al. (2012), but also designed reduced versions of these tasks.

Regarding this, Oswald et al. (2014) report that these shortened automated versions save time for the subject, and further the easiness for the researcher in terms of administration time and scoring. As a matter of fact, the versions of the WM tasks that were administered for the study of this thesis are the automated reading span task and listening span task designed by Redick et al. (2012) and Oswald et al. (2014). As it can be seen these span tasks have been carefully designed and tested for reliability making it possible to test WMC following the recommendations of Conway et al. (2005) for optimal administration.

More concretely, the advantages of using shortened automated versions are the following:

1) They can be applied remotely as they are accessible through the internet;
2) They are answered using a computer mouse (or digital pad), which makes it more convenient to respond for the subjects;
3) They are significantly more reliable (cf. Unsworth et al., 2005) than manual non-automated versions (cf. Miyake, 2001 who reports that previous versions of WM span tasks are not highly reliable);
4) They save time for the test taker and the test applier at the same time that they reduce human error and administration times.
5) They facilitate the calculation of scores and the storage of the results (cf. Juffs & Rodriguez, 2014 who report problems using manual outdated WM span tasks on WM and L2 processing studies such as the one by Juffs, 2005 and Felser & Roberts, 2007).

In sum, this subsection has reviewed the considerations taken to design instruments that can validly and reliably measure working memory capacity as defined by Baddeley and Hitch (1974) and reinforced by Baddeley (2007). It has been discussed how these instruments have evolved from fairly rudimentary and time-consuming (Daneman and Carpenter, 1980)

to digitalized, time-efficient, highly valid and reliable span tasks (Oswald et al., 2014; Reddick et al., 2012).

The availability of these WMC tests has a substantial effect for the current research study as it creates the opportunity to obtain WMC measures of L2Aers in a context where this type of research has not taken place. Not only do these online automated WMC tests make it possible to test L2Aers in an English as a Foreign Language (EFL) context, but they also allow for larger groups of different proficiency levels to be considered as subjects of study. These two aspects are crucial for the advancement of SLA cognitive research in terms of individual differences in working memory capacity because studies in this niche have been constrained to relatively small groups of participants in mostly English as a Second Language (ESL) contexts.

Therefore, the automated designs (Oswald et al., 2014; Reddick et al., 2012) of the reading and listening span tasks (Daneman & Carpenter, 1980) have made it possible to have a group of 22 intermediate and 27 advanced late Spanish-speaking learners of English participants, and a group of 24 late native speakers of English tested for WMC in the current study. The possibility of measuring participants' WMC with two tests assures that the objective of exploring the effects of this individual difference on these specific participants' attainment of L2 complex aspects is accomplished. Furthermore, it contributes to expand the knowledge of WMC and SLA to EFL contexts, and to late L2 learners of different proficiency levels.

In this manner it amends the inaccuracy on WMC measures due to outdated test instruments (Juffs & Rodriguez, 2014), and it reinforces the findings of studies that have found a relation between WMC and L2 processing of complex properties (Dussias & Pinar, 2010) to contexts of limited L2 access among learners with distinct proficiency levels. Lastly, since automated RST and LST deliver highly accurate, and precise WMC measures, it has been plausible to explore the effects of this individual difference in the L2 automaticity of these same participants.

2.2.4 Working memory capacity and language

Nevertheless, one of the goals of the current section is to situate the study of working memory capacity (WMC) in the current research of second language acquisition focusing on late learners. Thus, the following subsection addresses the relation between WMC and

language in a semi historical mode in an attempt to establish the significance of the study of working memory as a mechanism that is intimately correlated with the processing and production of language in general.

Given the latter, Baddeley (2017) mentions that the study of working memory and its relation to language became obvious with the work by Baddeley, Papagno, and Vallar (1988) in which they were testing the role of the phonological loop (PL) subsystem on people that had an impairment in short term memory. As an antecedent, Vallar and Baddeley (1987) noted that the patient in their study had a normal language production as well as a normal comprehension under circumstances that did not require complex sentences.

However, in a later study this same patient was asked to learn eight vocabulary items pertaining to Russian; her learning of Russian words was to be compared with her learning of unrelated words in her native language. Results for this study pointed to an obvious difficulty in learning the Russian words, and an easy learning of native words. Baddeley et al. (1998) concluded that the PL was helpful in learning new words. This is of particular importance for the SLA field and had repercussions in the development of theory and research with an emergentist approach who studied the phonological loop as a separate phonological memory (PM) (cf. Juffs & Harrington, 2012) (e.g. Ellis, 2001; Ellis & Sinclair, 1996).

Albeit, Baddeley et al. (1998) furthered their discoveries on the phonological loop and the learning of new vocabulary, particularly in terms of foreign language (FL) words. In this vein, Papagno, Valentine, and Baddeley (1991) studied the learning of FL words and compared it to the learning of native words on normal late subjects. The authors found that requesting participants to suppress rehearsal by means of articulating an irrelevant sound affected the learning of foreign words but not that of the native words. This meant that rehearsal in short-term memory had a great impact to acquire novel words, and this directly resonated in the methodology of research that started considering the role of working memory in SLA (Juffs, 2004). Nonetheless, this conception limited working memory to PM as a component of STM as I have pointed out previously.

The fact that Baddeley (2003a, 2017) attests that the phonological loop in working memory intervenes in the acquisition of new and even foreign vocabulary does need to be considered in cognitive SLA studies, though careful considerations need to be made for the development of methodology to measure WM. As mentioned before, some research in SLA

has concentrated solely on the phonological loop or on phonological memory, and thus working memory has been measured with repetition tasks (e.g. Ellis & Sinclair, 1996). This is problematic not only with respect to methodology, but also in regard to implications for SLA research on late learners because it reduces WM to a simple memory processor that any late learner can access by means of a vocal or sub-vocal repetition of L2 input. Although, the latter might be sufficient to explain acquisition of L2 lexicon, it does not suffice to explain acquisition of much more complex L2 grammatical features (e.g. resumptive pronouns), or L2 oral fluency; also, the view of WM as phonological memory is insufficient to explain variation in late learners' L2 attainment.

In spite of the latter, the work by Just and Carpenter (1992) has widened the range of empirical studies in SLA and working memory. In their study, Just and Carpenter (J&C) (1992) propose that working memory capacity is related to the syntactic and semantic processing of language comprehension. In order to prove this, J&C tested adult college students using a reading span task (RST); as previously discussed, the RST marked a difference in measuring WMC as this task involved use of complex thought -processing through the central executive and episodic buffer (Baddeley & Hitch, 1974; Baddeley, 2000).

Hence, J&C (1992) compared the RST scores, with the participants' comprehension and parsing of sentences that pose syntactic and pragmatic ambiguity. Their findings suggest that there is a high correlation between working memory spans and individuals' comprehension and parsing of sentences containing syntactic and pragmatic complexities. The findings in J&C are greatly significant for the study of WM (measured as a mechanism of complex thought, not only of STM storage) in SLA; especially for individual differences research on late L2 learners. Consequently, these findings indicate that 1) adults have different levels of capacity in working memory, and 2) this capacity directly affects the processing of linguistic intricacies. This justifies, more precisely, the need to conduct more research on individual differences in working memory capacity in SLA since it offers more bases to explain a) variability in attainment, as well as b) how late learners attain high L2 proficiency.

Worth mentioning is that J&C (1992) associate their findings to theories of connectionism that are based on Anderson's (1992) ACT model and the language acquisition principles proposed by Rumelhart and McClelland (1986). Over these connectionist frameworks, Just and Carpenter (1992) discuss that "the computations that are involved in language comprehension also can be expressed as manipulations of activation, as they typically are in connectionist models of comprehension" (p. 123). The latter implies that since

working memory is a mechanism involved in linguistic processing, it might be due to be activated and altered by the frequency of input and strengthening of neuronal association as proposed by connectionist theorists (Ellis, 1998; Ellis & Schmidt, 1998; Matesa & Anderson, 2000; Rumelhart & McClelland, 1986). This assertion in J&C is twofold in SLA; it might be advantageous for L2 pedagogy to consider working memory as a mechanism that can be "trained" through frequency and patterns found in the input, but it is disadvantageous because if WM is trainable then: why is there proficiency variation in late learners of a same group that is exposed to L2 input, and taught to identify patterns?

The assumption of working memory being subject of manipulation in J&C (1992) through frequency and patterns provided by the input is contradicted in the findings of Ellis and Sinclair (E&S) (1996) extensively discussed here. As pointed out, E&S (1996) tested WM as a component of short-term memory. However, they followed a connectionist principle in their study, which is akin to most emergentist approaches: their subjects of study received training that consisted in input exposure, rehearsal, and controlled-for patterns (cues) provided in the input. E&S suggest that "an involvement of phonological WM in syntax acquisition as a) phonological short-term memory predicts native grammatical ability, b) individuals with short-term memory deficits show restricted acquisition of syntax both in native and foreign languages, c) the more chunks in syntactic marker, the more difficult it is to acquire, and d) children's short-term memory capacity determines their success in learning the syntax of an artificial language" (Ellis & Sinclair, 1996, p. 238).

The assumptions in E&S agree with Just and Carpenter (1992) in that working memory is a crucial mechanism for language processing, but E&S go further and attest that it is crucial for language acquisition and that is activated through rehearsal in phonological memory. However, E&S treated PM as merely a memory storage mechanism, which refers WM as a component of short-term memory. Therefore, it cannot be proved that working memory, as considered in J&C (adding functions of higher-order processing retrieved from long-term memory), is activated and/or manipulated through frequency of pattern encounters through the input. If anything, the findings in E&S (1996) indicate that only the memory storage function of working memory can be "trained" and maybe modified by means of phonological repetition in order to learn certain aspects of the language.

All in all, the study of J&C impacted research in language and working memory as it evidenced the role of WMC in grammar processing as well as that set the precedent for considering that this mechanism is present in adult stages. For these reasons alone, the J&C

(1992) study serves as a guide to second language acquisition research. Yet, many considerations have to be made before starting to consider that working memory is susceptible to training, and thus improvement. First, as mentioned in Juffs and Harrington (2011), more data needs to be extracted and analysed to be certain of the role of WMC in L2 learning and attainment amid late learners. For the moment, the present investigation focuses on precisely that goal.

To sum up, this sub-section was intended to describe the association between working memory and language in general. For this purpose, the contents were organized chronologically to show how the study of language and working memory started to develop. Most importantly though, in depicting such development, it was demonstrated that working memory is highly significant for language processing, and that it is present in language related aspects that require complex processing in adult stages. Additionally, I added a brief commentary on the implications that every cornerstone study had on the present investigation, and ergo, to the field of SLA cognitive research on late learners. The topics covered in this sub-section precede the discussion of the next subsection.

2.2.5 Working memory capacity and second language acquisition

As stated in the previous section, the relation between working memory and language is intimately close, distinctively for the function that WM has on processing linguistic information and underlying linguistic knowledge. In this way, it has been inevitable for researchers following the Hitch and Baddeley (1974) model, and to Baddeley himself to acknowledge that if working memory influences L1's processing of higher and complex features and skills, then it must have a central role in second language learning (Baddeley, 2003a; 2017).

Nonetheless, as mentioned in section 2.1 under emergentist and generative approaches to SLA, the concept of working memory does not abide to a consensus. Regarding this, Wen (2016) discusses that some WM studies in SLA "used the same broad term 'WM' to denote several different things or, to be more exact, to implicate different factors or components of the same construct (e.g. its phonological component)" (p. 73). Consequently, the difference in definitions within the SLA field creates confusion to understand how exactly working memory operates in L2 processing, or skill development; there is even a misunderstanding in determining if all or just certain elements of WM have an effect in L2 learning.

As I emphasized before, some of this confusion regarding working memory can be observed in emergentist theory and research on SLA studies. Though for the present research, emergentist research guides the notion of what learners are capable of acquiring; it is also an approach that contributes to understand the role of WM in SLA. More specifically, the functions corresponding to working memory comprise the theoretical grounds for approaches such as information processing (IP) (McLaughlin et al., 1983; McLaughlin & Heredia, 1996), as discussed in the first section of this chapter. In fact, Skehan (1998) considers that processing information in working memory is necessary for learners to start their L2 acquisition development; in information processing terms, working memory is the central processing that turns L2 input into "intake" (cf. Skehan, 1998), which leads to L2 output (Wen, 2016).

What is more, in Skehan (2016) the hypothesis is made that higher capacity in working memory provides an advantage for L2 learners since it causes them to unfold a deeper analysis of the L2 input. This same hypothesis extends for "noticing" and "pattern recognition", which are cornerstone factors for IP theory; in this regard, working memory is said to be the mechanism where "noticing" takes place given that it allots L2 input for longer periods of time (Skehan, 2016). Nevertheless, as in Ellis and Sinclair (1996), Skehan's (2016) definition of working memory does not agree with the B&H model, adopted in the present work, in that WM is limited to merely a memory storage mechanism.

The problem with Skehan (2016) and Ellis and Sinclair (1996) when disregarding the functions of the central executive and the episodic buffer attributed to WM in the B&H-B model is that there is an omission of complex thought involvement when information is temporarily stored; ergo, there is no interaction with knowledge and skills stored in long-term memory. In other words, confining working memory to serve as a memory unit that holds information would imply that any L2 learner who retains or memorizes L2 input for a certain amount of time will be able to store such input in LTM, which can be further retrieved efficiently or automatically. In this sense, this approach to working memory is not consistent with the research conducted by Baddeley et al. (1988) mentioned in 2.2.3 above.

In addition, as it was mentioned in the previous section, proponents of emergentism such as Nick Ellis (1998, 1999, 2007, 2008) link WM to L1 and L2 acquisition. Additionally, the proposal of WM extends to connectionist views in SLA. Also, as mentioned in chapter 1, section 1.2.2, Ellis and Sinclair's (1996) work is key to define the role that WM plays in SLA.

In the same vein, other researchers focusing on the parsing of L2 grammar at the sentence level have taken into account the role of WMC in the SLA process. An example is the work of Juffs and Harrington (2011) who not only give an account of the current methods to measure WMC, but also provide a larger picture of the SLA research that has included WM. In lieu of this, Juffs and Harrington (2011) make the following final remarks on their WM current research account: 1) effects of individual differences on memory span is not detectable and when it is, there seems to be a relation with factors such as schemata or pragmatic inferences; 2) listening span tests might predict output, but there are few studies to confirm such findings; 3) WM is assumed by many researchers as a form of attentional control on later learning which needs to be controlled in terms of resources and suppression of competing sources as a key to succeed in L2 learning.

On the line of emergentism, and aiming at exploring factors of automaticity and WMC, studies such as the one by Fortkamp (1999) are central for the purpose of study in this thesis. Fortkamp tests the influence of WMC over the fluency of L2 speech production at the discourse and articulatory level. The author bases on Daneman's (1991) study in which a significant correlation was found between WM and L1 fluency. In order to do so Fortkamp runs a variety of tests, such as a speaking span test in English and in Portuguese as well as an RST to measure WMC during language comprehension. Additionally, to make sure that subjects were applying comprehension processes, a grammaticality judgment test was applied. A noticeable test that was applied was a speech generation task (SGT) to measure L2 fluency, in which subjects were asked to describe a picture with as much information as they could provide.

Finally, an oral reading task and an oral slip task were applied to measure L2 fluency as well. The author finds a correlation between WMC and speech generation. In addition, Fortkamp finds that those individuals with more efficient processing skills for the L2 performed better on the SGT; whereas subjects with less efficient processes, might have less storage to remember words in their exact form. Fortkamp also finds that there is a significant correlation between the number of words on the SGT and the ratings of the two native assessors in the study. This study's results support the assistance provided by WM on task-specific situations. The author makes a call for further research since there is a limitation on findings related to speech generation tasks. Finally, Fortkamp points out that the topic of individual differences in WMC is still an unresolved issue in the literature and research of SLA.

Other work on L2 fluency and working memory capacity includes that of Mizera (2006) who integrated different tasks to measure L2 adult learner's fluency in their L2, as he used more advanced statistical methods to explore any correlation with WMC. The use of statistical models, such as a linear or a multiple regression analysis, as in the work of Mizera, and in this research, allows to understand if WMC has a significant size effect; as well to have a more specific notion of the aspects of L2 acquisition that WMC influences. However, the results in Mizera's investigation show that there is not a significant effect between WMC and oral fluency in the group of advanced learners that he tested. Mizera explains that his results are in line with views that sustain that advanced learners produce speech in automatic processing, and thus WM does not have a role in this stage of L2 acquisition (cf. Temple, 2000, 2007).

Thus, research approaching WMC in SLA needs to include groups of participants of different levels of proficiency as it is the case of this investigation. Some recent work has paid attention to levels of proficiency, such as Sagarra (2017) who conducted a longitudinal study on the effects of WMC on the grammatical knowledge of late learners. The author finds that WM only significantly correlates with grammatical knowledge when the learners are in a lower stage of L2 acquisition. Nonetheless, Sagarra's work approaches WM as a mechanism that does not vary in capacity from learner to learner; which might be problematic because it suggests that all late learners undergo the L2 acquisition process in a similar manner. Moreover, Sagarra (2017) does not report what specific aspects of grammar are significantly related to WM in a low level of acquisition.

Furthermore, the latter emphasizes an important issue with the study of WM in SLA: the lack of a defined approach. As Wen (2016) states:

> "[…] due to the many controversies and debates surrounding the WM construct and the daunting number of WM measures and assessment procedures in the source discipline of cognitive psychology, some of the WM-SLA studies are fraught with limitations and caveats that have led to severe pitfalls in their research designs and methodologies" (p. 76).

Therefore, the study of WM in SLA needs to first consider the multimodality of WM (Baddeley, 2017). In other words, the methodology to measure WM must be based on tasks that measure all the components involved in WM; including the functions of the executive control (Juffs, 2011; Wen, 2015, 2016). Secondly, individual differences in the capacity to operate WM cannot be minimized and assumed to be the same for all L2 learners as assumed in emergentist theories to SLA (see 2.1.1.5 for a detailed discussion of the role of WM in emergentism). Finally, the effects of WMC need to be explored in different levels of L2 proficiency, and on specific aspects that are challenging to attain by L2Aers; in terms of both

grammatical comprehension and oral production. All of these aspects are considered in the current research work.

2.3 Second language oral fluency

As discussed, the current research work aims at studying how working memory capacity (WMC) influences second language (L2) oral fluency and on late learners (L2Aers) of English. The reason for which L2 oral fluency is a targeted aspect of study in this research is its connection to L2 automaticity (DeKeyser, 2001; Kormos, 2006; Schmidt, 1992; Segalowitz, 2010) (see section 2.1.1.4. for a detailed explanation on the concept of automaticity in the theory of cognitive second language acquisition). Thus, it is paramount to understand what second language oral fluency entails, the challenge it represents to be developed by L2Aers, why it is an aspect that denotes L2 automaticity and how it has been approached in L2 research.

However, L2 oral fluency as an aspect of language that derives from reaching L2 automaticity is a level of acquisition that varies from L2Aer to L2Aer. As it has been previously pointed out, it is precisely in these variable aspects and expected outcomes of the L2 where WM effects should be assessed. Because in spite of the treatment that emergentists have given to WM as a necessary catapult to L2 automaticity (e.g. Ellis & Sinclair, 1998; McLaughlin & Heredia, 1996), not all late language learners arrive to this level with the same characteristics (see 2.1.1.4 and 2.1.1.5 in this chapter for a further discussion on second language emergentism, L2 automaticity and working memory capacity respectively). That is why, the present study seeks to analyse working memory capacity and its effects on the variation of L2 oral fluency on L2Aers given that this cognitive mechanism needs to be empirically explored as an individual difference on measurable aspects of L2 automaticity in SLA cognitive research. Considering this, the present section is intended to present a theoretical ground and scope of L2 oral fluency, why it is a measurable aspect of L2 automaticity, and how it has been studied in terms of WM.

2.3.1 Defining second language fluency

As it was pointed out in the previous section the concept of fluency is not stably defined and it takes multiple approaches, studies, and fields to get a grasp of what it entails. As it was mentioned, the present study deems the definition of fluency by Lennon (2000) as "a rapid, smooth, accurate, lucid, and efficient translation of thought into language under temporal constraints of on-line processing" (p.26) as the most complete and accurate. The

reason for this is that it resembles Schmidt's (1992) affirmation that fluency results from an automatic procedural processing of information. The association between fluency and automatic processing is germane for the research objective of the current work since I attempt to explore second language (L2) automaticity through oral fluency and its relationship with working memory capacity. Thus, the previous definitions of oral fluency corroborate that it is a product of L2 automaticity. So far though, fluency has been defined from a general perspective that is limited to the native language (L1) circumstance.

Consequently, it is necessary to describe oral fluency from a second language lens. With this mind, Segalowitz (2010) argues that second language oral fluency requires the grasp of multiple frameworks to "start thinking about L2 fluency" (p.7). Then, a starting point is interpreting L2 fluency from an adaptation to Levelt's (1989, 1999) blueprint model provided by De Bot (1992). On this adaptation, De Bot (1992) presents a model that accommodates more to the L2 speaker; this model is denominated "Blueprint of the L2 speaker" (Segalowitz, 2010). On this blueprint of the L2 speaker model, De Bot attempts to underscore the aspects in the blueprint model by Levelt that 1) might have to do with an intersection of knowledge of the two languages of the L2 speaker, and that 2) might cause an overlapping relationship between the speaker's knowledge of the L1 and the L2. The latter is referred to with dotted and dashed circles in De Bot's blueprint of the L2 speaker model (Segalowitz, 2010, p. 9); these dotted and dashed circles indicate the difference between the L1 and the L2 linguistic systems in terms of speech fluency (referred to as oral fluency in the present study) (see Figure 2-3 for a visual representation of the Blueprint of the L2 speaker by De Bot).

Furthermore, De Bot (1992) adds a symbol for fluency, {f} that indicates potential aspects where the L2 speaker might encounter vulnerability to carry out speech in the L2. In this regard, the L2 speaker might find their fluency compromised at the following levels according to De Bot's model: micro planning, grammatical encoding (pre- and post-processing), morphology-phonological encoding, phonetic encoding, articulation and self-perception.

There is much to discuss about how the L2 speaker's fluency finds a burden on these processing levels, but the relevant point to be made is that L2 oral fluency cannot be accounted as a synonym of L1 oral fluency since at the very levels of underlying processing, its generation undergoes different circumstances as pointed out by De Bot's L2 Blueprint model on a comparison with Levelt's (1989, 1999) blueprint model for L1 speech.

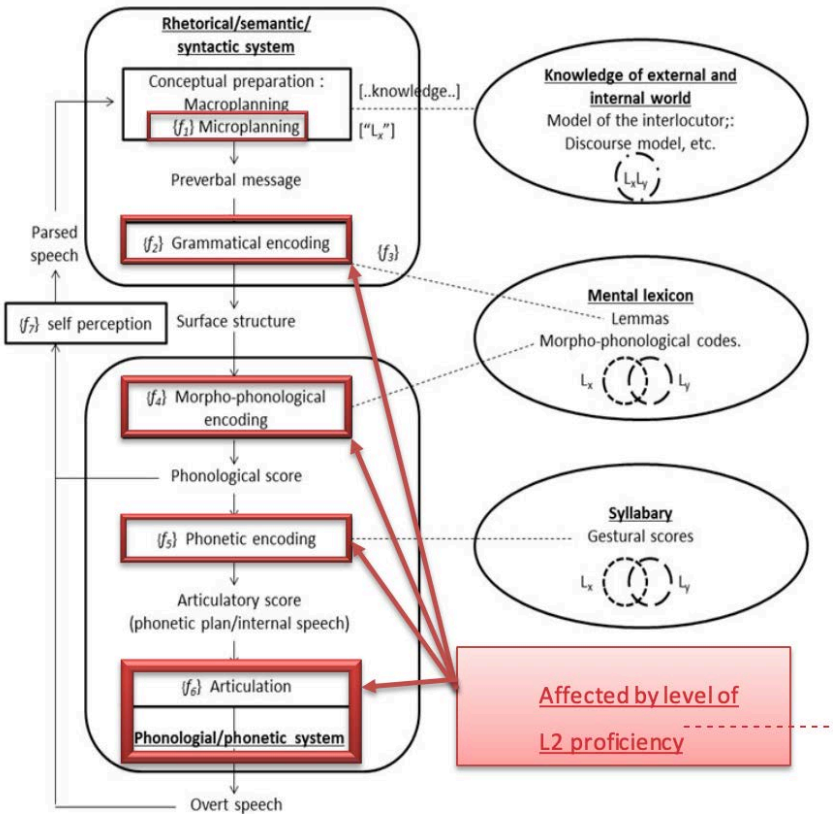


Figure 2-3 De Bot's (1992) Blueprint for the L2 speaker (retrieved from Segalowitz, 2010, p.9) with highlights of where the L2 speaker finds challenges in the speech production process in red.

By the same token, Segalowitz (2010) proposes to analyse L2 fluency from a dynamic systems theory approach as this might compensate for the lack of explanation on how L2 proficiency develops. Also, a systems theory approach to L2 fluency might encompass the impact of the environmental effects on De Bot's (1992) L2 blueprint model. In this vein, Segalowitz positions L2 fluency on Larsen-Freeman and Cameron's (2008) proposal of a dynamic system to L2 acquisition. Henceforth, L2 fluency is approached from the following features of the dynamical systems theory: heterogeneity of elements and agents, system dynamics, non-linearity, openness, and adaptation (Segalowitz, 2010). Figure 2-4 below shows Segalowitz's visual representation of the effects that the features on the dynamic systems theory have on L2 speech production.

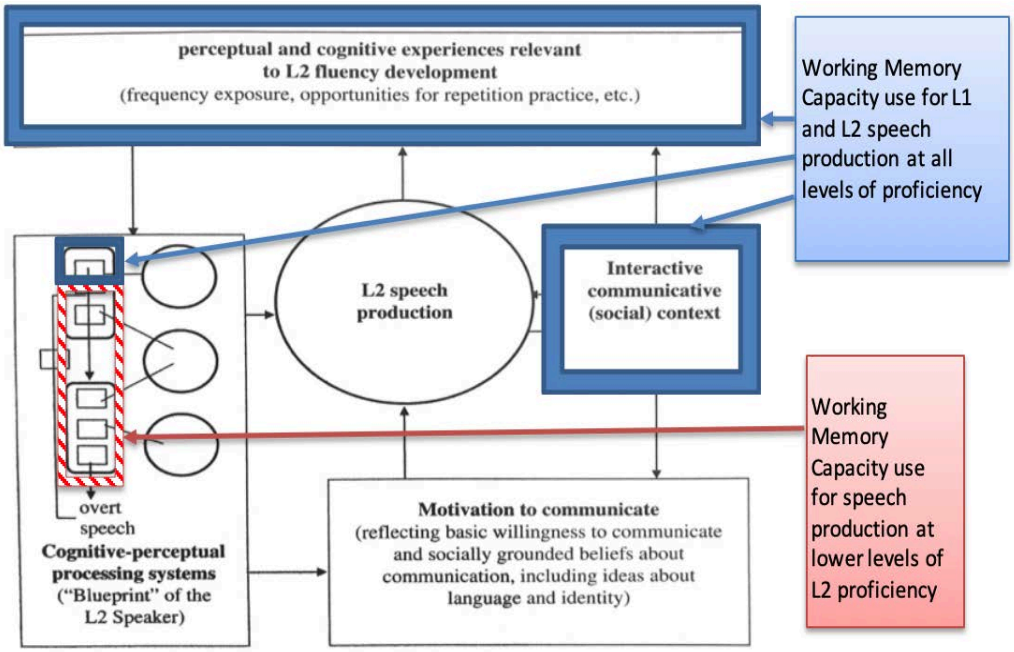


Figure 2-4 Segalowitz' (2010) provisional framework for the dynamic relationships that influence L2 fluency (p.21) with an indication of the influence of WMC.

According to Segalowitz, the five features that comprise Larsen-Freeman and Cameron's theory of dynamic systems serve as a spinal guide to examine the different aspects that might influence the fluency of the L2 speaker. As Segalowitz (2010) points out, the core of L2 fluency rests on how the L2 speaker manages, and with time, masters these dynamic system features. It is important to mention that while a native speaker's oral fluency is influenced by similar features, the L2 speaker needs to deal with the managing and mastering of these features at the same time that they are acquiring the second language linguistic knowledge (e.g. phonotactic constraints, intonation, morpho-syntactic information, etc.). Moreover, these five features provide an insight on L2 learner's oral fluency as they shed light on how complex this aspect of L2 acquisition is, and ultimately, they serve to comprehend how much more challenging it is for adult L2 learners to develop it.

Nonetheless, the features of the dynamic systems are not being tested in the present research. It is important to consider, however, that in order to have an accurate sample of L2Aers' oral fluency, these five features must inform the methodological aspects employed to extract the data concerning this variable. Therefore, the implication that L2 oral fluency is a fluctuating aspect of language and cognition that is altered by social and contextual elements guided the design of the elicitation task used to extract L2 fluency data in the present research work.

However, the adaptation of the Levelt (1989, 1999) model for speech production by De Bot (1992) and its inclusion in the framework of dynamic relationships influencing the L2 fluency system by Segalowitz (2010) are important to discuss in light of the role that WMC plays in terms of L2 oral fluency. In order to make this point is necessary to refer to the Levelt (1989, 1999) model of speech production for the L1 (see Figure 2-5 below). First, as it was mentioned before, De Bot (1992) suggests that the L2 speech production process starts to differ (from the L1 speech production process) from the Conceptualizer (at the microplanning level), all the way to the phonetic encoding in the Formulator. In the original model by Levelt (1989, 1999), the Conceptualizer is "where the selection and ordering of relevant information takes place and where the intentions the speaker wishes to realize are adapted in such a way that they can be converted into language." (De Bot, 1992, p. 4).

However, De Bot (1992) suggests that parts of the speech production processing in the Conceptualizer, which include Macro and Micro planning stages, are language specific and language independent; the former (the macroplanning) is language independent and the latter (microplanning) is language specific. Thus, for both monolingual and multilingual speakers these two stages in the conceptualizer would necessitate the engagement of WMC as central executive functions might be required to assess the type of discourse and the external influences mentioned by Segalowitz (2010) in the framework for the dynamic forces that influence L2 oral fluency (see Figure 2-4 above where the use of WMC has been marked in blue for L1 and L2 speakers in the framework).

Nonetheless, for monolinguals and/or for L1 speech production, the microplanning stage in the Conceptualizer might not represent a considerable load in processing as it would for an L2 learner. This stage, according to De Bot (1992), requires for the L2 learner to engage in the decision making of choosing the linguistic information that is appropriate for the communicative goal; this decision making process at the microplanning stage might be more costly for L2 learners who have to deal with the competing linguistic information available to them in their L1 as well. Therefore, even when the macroplanning (which involves discourse and context) and the microplanning at the Conceptualizer stage require the involvement of the speaker's WMC, the load in capacity for L1 production (for monolinguals) might be less since the linguistic information for the goal is available. At this stage, L1 speech production does not require the inhibition, control, and decision-making processes that are required when processing the linguistic repertoire of an L2; which is possibly not fully acquired, not yet stored in LTM, and not fully automatized in some instances.

Moreover, WMC also needs to be employed by L2 learners at the Formulator and Phonetic Encoding stages, which precede the Articulator (the final stage of speech production in the Levelt (1989, 1999) model). At the Formulator, "the preverbal message is converted into speech plan (phonetic plan) by selecting the right words or lexical units and applying grammatical and phonological rules" (De Bot, 1992, p. 4). For monolingual speakers or for L1 speech production, the type of linguistic processing that takes place at the Formulator level are completely automatized. That might not be the case for an L2 learner as their grammatical knowledge, both at the morphosyntactic and phonetic level, could be incomplete and in competition with their L1 grammatical knowledge.

Therefore, it can be hypothesized that WM functions are necessary for L2 learners at the Formulator stage as well since they need to inhibit L1 grammatical distractors, while accessing and retrieving the L2 morphosyntactic and phonetic information that is needed for the task at hand. The latter causes a great demand in the L2 learners' WMC as they might still be in the process of automatizing the grammatical and the lexical knowledge of the language. This in return implies that WMC needs to be operated in full capacity as storage and higher order skills are required to produce the speech that is required for the given situation.

The role of WMC on the processing stages of speech production in the Levelt (1989, 1999) and De Bot (1992) models has been used to describe speech production and oral fluency among L2 learners as in the study performed by Temple (2000). In her research, Liz Temple investigated the differences in speech production between language learners and native speakers. Temple (2000) calculated temporal measures of fluency to compare the speech samples of 30 learners of French and 20 French native speakers.

The results in this study show that the L2 speech varies significantly from L1 speech in fluency measures related to error, pause and repair rates. Temple (2000) explains that the significant differences in measures of fluency have to do with the limits in WMC of the language learners. The author highlights that while native speakers do have to engage in working memory at the Conceptualizer level, where the preverbal message is planned (cf. Levelt, 1989; De Bot, 1992), they do not need to place a load on WM at the Formulator stage as their lexical and grammatical knowledge is automatized; thus, out of the domain of WMC processing. The L2 learners' speech was characterized by more dysfluencies (cf. Kormos, 2006; Segalowitz, 2010; Skehan, 2009, 2014), according to Temple (2000), given that they have to put a greater demand on WM at the Conceptualizer and Formulator stages of speech

production. As a consequence, the learners' L2 fluency is affected and is "lower" compared to that of the L1 speakers.

Nonetheless, Temple (2000) did not measure the leaners' WMC; neither did the study treat WM as an individual difference (ID) among language learners. Other studies involving L2 speech production, as the one by Fortkamp and Bergsleithner (2007), do treat WM as an ID as well as they allude to the De Bot (1992) L2 speech production framework. Although this study does not measure L2 oral fluency per se, it does attempt to explain the effects that IDs in WMC have on the L2 learners' speech production. In their research, 18 learners of English (classified as pre-intermediate) were asked to perform a speaking span task to measure their WMC; they were also asked to perform in a speech generation task in which they had to use a particular grammatical structure. Finally, they were involved in a task in which they had to do an oral report on the grammatical rules that they had been exposed to.

Fortkamp and Bergsleithner (2007) found that there was a significant correlation between the L2 learners' WMC and their speech production (no significant correlations were found between WMC and the learners' performance on the noticing task). The authors find that L2 learners with a higher WMC were able to produce more grammatically accurate sentences with regards to the L2 learners with a lower span in WMC. The finding in the Fortkamp and Bergsleithner (2007) study is relevant to the current discussion as they also highlight that WM is involved in the processing of L2 oral production. Furthermore, the authors theorization of the underlying processes of L2 speech production are in line with De Bot's (1992) framework and emphasize that it is at the Conceptualizer and Formulator stages where the L2 learner encounters considerable challenges in processing that have an effect on their speech production.

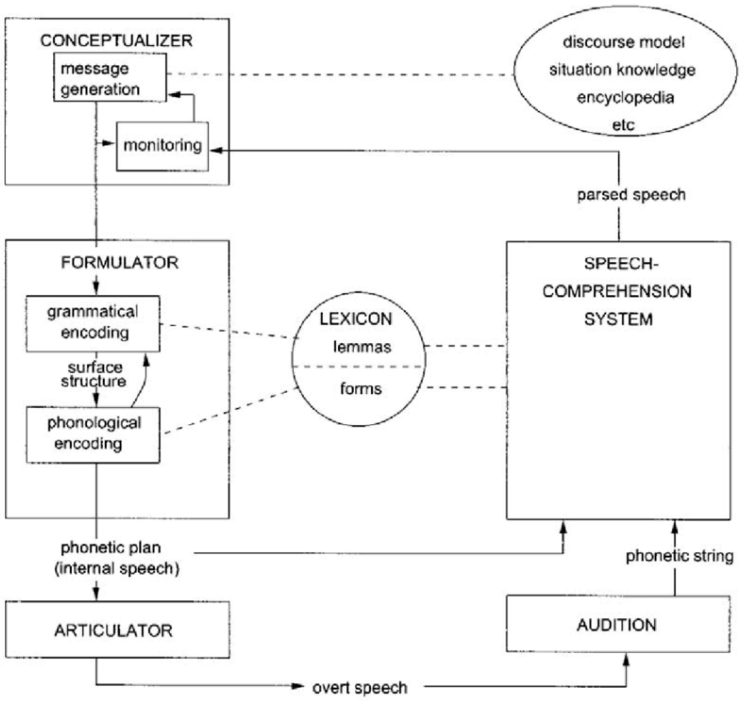


Figure 2-5 Levelt (1989) model of (L1) speech production

Although the descriptions and interpretations made up to this point are highly useful to have an in-depth knowledge of L2 fluency, I would like to close this section by stating the assumptions and conceptualizations taken in the current research work. The first crucial assumption is that L2 oral fluency "is an observable characteristic of real time speech behaviours" (Segalowitz, 2010, p.6). The second aspect for the conceptualization of L2 oral fluency here adopted is that it is "one component of oral proficiency and is basically related to speech rapidity, to the flow of speech without this being impeded by hesitations" (Fortkamp, 1999, p.7). About the latter though, it can be argued that hesitations are a part of the flow of speech (Goldman-Eisler, 1968). Hence, a pertinent amendment to the latter adopted definition is that L2 fluency does involve hesitations, but these are not major or noticeably significant (cf. Kormos, 2006; Segalowitz, 2010).

Finally, as pointed out by Segalowitz, L2 fluency is a real-time behaviour; so, in this respect, speed is insufficient to define L2 oral fluency. Therefore, the role of "flow" in Fortkamp's (1999) definition alleviates this constant misunderstanding since it summarizes the multiple factors that intervene on its composition. As a result, these assumptions and definitions comprise the take on L2 oral fluency in the present work since it serves better to accomplish one of the main objectives: explore L2 oral fluency as an aspect of automaticity with respect to working memory capacity. This objective can only be fulfilled by focusing on

the measurable components of L2 oral fluency, which lie at the cognitive processing level and are mainly defined and identified with the assertions stated above.

In sum, this section has broadly described L2 oral fluency in the light of cognitive theory relevant for both linguistics and second language acquisition. Nevertheless, it has been emphasized that the current research is based on the views of L2 oral fluency as an observable aspect of speech (Segalowitz, 2010) that is associated with its flow (Fortkamp, 1999). The latter is germane for the current study since it allows to analyse this complex aspect of the L2 from a cognitive perspective; also, this stand permits to analyse L2 oral fluency in a systematic, quantifiable way with respect to WMC on L2Aers. With this mind, the following section is intended to portray the measurable characteristics of L2 oral fluency, from a purely cognitive perspective, that allowed its quantitative analysis in the current investigation.

2.3.2 Second language utterance fluency

The previous section provided fundamental frameworks and angles on L2 oral fluency that are paramount for its comprehension and study in SLA or language teaching. However, Segalowitz (2010) poses that, in order to have a systematic understanding of second language oral fluency, it is necessary to analyse it from a strictly cognitive perspective.

With that aim in mind, Segalowitz categorizes oral fluency in three types: cognitive fluency, utterance fluency, and perceived fluency. Segalowitz (2010) defines cognitive fluency as what "has to do with the speaker's ability to efficiently mobilize and integrate the underlying cognitive processes responsible for producing utterances with the characteristics that they have" (p.48). In other words, cognitive fluency is the result of the efficient functioning of the underlying processing systems that carry out a fluent speech production. This definition can be traced back to Levelt's (1989) blueprint model (see Figure 2-2 for an illustration on this model) in which fluency is described with an optimal functioning of the underlying mechanisms that comprise the rhetorical/semantic/syntactic, and the phonological/phonetic systems of speech.

The second category, utterance fluency, has to do with the characteristics or features that are contained in an utterance of speech (ibid). Utterance fluency is the part of L2 oral fluency that has been targeted to meet the first research objective of the study here

presented. For this reason, a more thorough description of this category is discussed later in this section.

The last category proposed by Segalowitz, perceived fluency, is related with the interpretations that listeners create about a speaker's cognitive fluency. Although the label of this aspect of fluency is somehow self-explanatory, it can be added that perceived fluency has to do with the psychological perceptions and assessment that the listener assigns to the speech of the L2 speaker; although latent, this is an aspect of fluency that is highly subjective and complex to measure. An understanding of the dominions of these three aspects of fluency can be better understood in Figure 2-6 below, in which Segalowitz graphically represents the relation and dominion of the three.

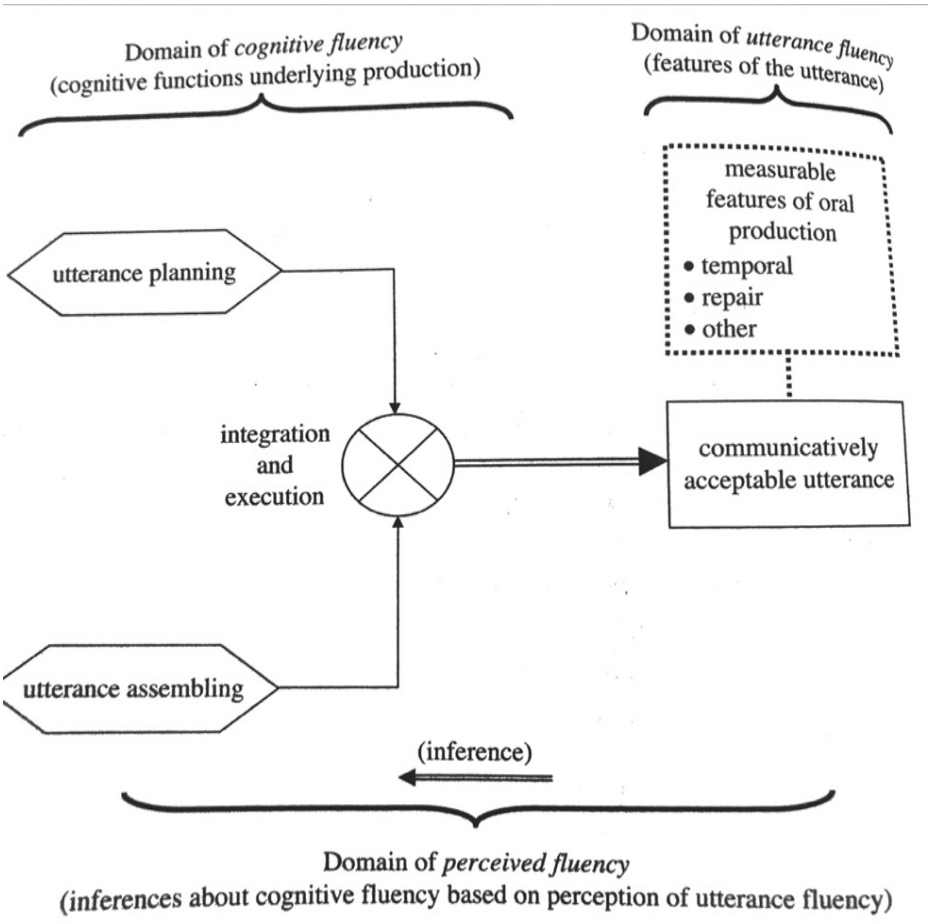


Figure 2-6 The dominions of the three types of fluency (Segalowitz, 2010, p. 50).

The reason for the focus on utterance fluency is that it is composed of measurable features that result from oral production. Since these features can be quantified, they make it possible to have a systematic analysis of L2 oral fluency in terms of cognition (Segalowitz, 2010, 2016). Furthermore, this dimension of fluency was also selected for the present study given that it allows to obtain and analyse speech data with tasks and tools that are somewhat

easily accessible for novice researchers as they do not represent economic challenges or difficult-to-operate software tools to obtain quantifiable measures of oral fluency. In the present research, the task to elicit speech only required a couple of drawings obtained from the internet (see Appendix H) and a free recording software to elicit the speech samples; to analyse the speech samples and obtain measures of utterance fluency, a free software called PRAAT (cf. De Jong, 2009a) was used.

Moreover, there is substantive theory and empirical data that points in the direction of using utterance fluency measures in L2 research if the goal is to have a more complete picture of what oral fluency entails in the L2 acquisition process. Segalowitz (2016), for instance, emphasizes the importance of using utterance fluency measures to understand the length of pauses in L2 speaker's speech and relate it to specific aspects of L2 proficiency such as vocabulary knowledge; as in the research by De Jong and Bosker (2013). In their findings, De Jong and Bosker (2013) observe that the length of silent pauses in the speech of L2 speakers of Dutch correlates with their vocabulary size (taken as a general measure of L2 proficiency), when the thresholds for silent pauses are adjusted within minimum and maximum limits of duration. Segalowitz (2016) mentions that "this finding moves the discussion of how utterance fluency reflects cognition from speculation to empirical test" (p. 82).

Also, Skehan (2014) mentions that the area of measurement of fluency "is far from straightforward, not at least because of complexity of fluency (or dysfluency) itself" (p. 18). Nonetheless, in a proposal to analyse L2 language performance using measures of complexity, accuracy, and fluency (known as the CAF framework), Peter Skehan (2009) suggests to use temporal measures such as length and number of pauses, repairs, and length of run to gain more understanding of what aspects of oral fluency are significantly different between native and non-native speakers. An example is the work by Skehan and Foster (2008) who compared the duration of the pauses made by native and non-native speakers using a task that involved the narration of a story (which the participants had to create based on different pictures that were presented to them); the narration of the story also involved making connections between the pictures.

Skehan and Foster (2008) found that native speakers tended to have longer pauses at clause boundaries and shorter pauses within clause boundaries; the non-native speakers showed longer pauses even within clause boundaries. The authors interpreted the results as an indication that non-native speakers need to pause more within clause boundaries because

they need to cope with lexical choices. Skehan (2009) asserts that research such as the one by Skehan and Foster (2008) shows improvement in terms of measurement of oral fluency; and, in order to "make sense" of the complexity, accuracy, and fluency framework for the analysis of language performance, " the characterization of fluency needs to become subtler and deeper" (p. 514). Skehan (2009) also mentions that aspects like fluency depend on underlying cognitive mechanisms such as WM; and, that limits in capacity in WM have an effect on L2 performance aspects. However, the role of WM is not fully detailed in the CAF framework by Skehan (2009); and thus, the study here developed should contribute to understand what aspects of L2 utterance fluency are related to the L2Aer's WMC; and, at what level of proficiency, WMC has more effects in the L2 developmental process of this subset of learners.

In a more recent approach, Tavakoli (2019) emphasizes the importance of seeing oral fluency as a "complex and multifaceted construct that is difficult to define, measure and operationalize" (p. 40). Tavakoli (2019) expresses that the measures of fluency proposed by Skehan (2003) provide a way to have a "finer-grained analysis of fluency so that different and what he referred to as 'separate measures of fluency' can be examined" (p. 46). In this vein, Tavakoli (2019) mentions that the integration of temporal measures, or the measures that constitute what Segalowitz (2010) labels as utterance fluency, make it possible to relate fluency to other aspects of L2 proficiency; or, as I mentioned before, calculating temporal measures of utterance fluency allow to relate L2 oral fluency to general cognitive mechanisms such as WMC, as in the present work.

Tavakoli (2019) adds that thorough research has been conducted to understand what measures of utterance fluency are more significant to the study of SLA. In a study performed by Tavakoli and Skehan (2005), for instance, 19 temporal measures were used to evaluate the different aspects that characterize oral fluency. In their research, 140 language learners' speech samples were analysed in terms of speed, breakdown and repair measures. The results confirm that the L2 speech samples displayed reliable and distinct characteristics of fluency, which supported the Skehan (2003) three-dimensional approach to study oral fluency. However, recent empirical research has shown that these measures can be reduced in number as high correlations were found among them (cf. De Jong, 2018; Segalowitz et al., 2017; Bosker et al., 2012). This finding will be further discussed in light of the design and results of Study 2 in this investigation.

Although the study of oral fluency basing on temporal measures of utterance fluency brings great advantages to the second language acquisition field in terms of having more reliable quantitative methods, there are other aspects of fluency that should be further investigated; namely, cognitive fluency. In this respect, Segalowitz (2016) mentions that research on cognitive fluency would signify the study of "[…] the speed and efficiency of semantic retrieval, the handling of the attention-focusing demands inherent in utterance construction, operations in working memory, among others" (p. 82). Therefore, adding this perspective to the present work would be highly beneficial as it would allow to have a deeper understanding of the effect of WMC, as an underlying mechanism, in L2 utterance fluency.

Segalowitz (2016) mentions that certain studies have been able to approach aspects of cognitive fluency such as the one by Segalowitz and Freed (2004). In their research, reaction times (RTs) and coefficients of variability (CVs) were obtained from a task that included a visual word semantic categorization (categorizing living from non-living objects) in the participants' L1 and L2; the measures for cognitive fluency were taken as L2-specific since they were residualized against the L1 measures. The results show significant correlations between this measure of cognitive fluency and L2 fluency.

Another approach to examine cognitive fluency is using tasks that measure the "flexibility in the control of linguistic attention" (Segalowitz, 2016, p. 84). For example, in a study conducted by Taube-Schiff and Segalowitz (2005), L2 speakers were asked to perform two tasks that demanded attentional shifts. One task included the judgment of the verticality meaning (above/below) of pieces of sentences such as "over the spot"; the other task required for participants to judge the proximity in meaning (close/distant) of parts of sentences such as "near the place". The tasks were presented in an intertwined manner to the participants so that they were forced to shift the focus of attention to a different linguistic activity. The results show that there is a strong correlation between the indexes of flexibility in focus of attention (as a measure of cognitive fluency), L2 utterance fluency, and L2 proficiency.

As advantageous as it is to study L2 oral fluency with measures of cognitive fluency, the tasks to extract these measures are not as accessible as those to obtain measures of utterance fluency. As a matter of fact, Segalowitz (2016) comments that "[…] RT speed and stability measures of cognitive fluency have always been obtained from visual, receptive tasks […]. Thus, the tasks require special software that displays a certain type of linguistic input, arranged in a specific manner, which can measure the participants' RTs in real time so

that CVs can be calculated. Unfortunately, these types of tasks are not attainable for me at the moment since they represent support from computational experts that can design the tasks at hand; the latter might imply using extra economical and human resources that were not available for me as a distance student with a limited budget at the time that I conducted the studies for this research. In spite of this, the study here developed, using measures of utterance fluency, should offer a valid and noteworthy perspective on the relationship between L2 oral fluency and WMC from a quantitative and cognitive perspective. One can only hope that the results here obtained can be further complemented with research that integrates measures of cognitive fluency. As Tavakoli (2019) mentions, "[…] research in fluency still needs to develop further if a thorough understanding of the complex nature of fluency is to be arrived at" (p. 49).

Nevertheless, as Tavakoli (2019) asserts, the research conducted on both utterance and cognitive fluency is highly valid and valuable; in particular, after the contribution of Skehan (2003) to understand and classify the temporal measures that integrate the three dimensions of utterance fluency. One of the most important benefits of using temporal measures of utterance fluency to investigate oral fluency is that the temporal measures resemble the perception of fluency of external raters. This is attested in the study of Derwing et al. (2004) who calculated the temporal measures of 20 speech samples produced by beginner Mandarin learners of English and compared them with the ratings of 28 untrained and 3 trained examiners. The raters had to evaluate the L2 learner's speech samples in terms of fluency, comprehensibility, and accentedness. The results show that the temporal measures and the examiners' ratings of fluency are significantly correlated.

Derwing et al. (2004) comment that "the finding that fluency ratings can be predicted from measurable characteristics of speech further supports the claim that rating data from even untrained listeners reflect properties inherent in the stimuli and are therefore useful in the evaluation of speech samples" (p. 672). Therefore, although there is a need to complement the research on measures of utterance fluency with measures of cognitive fluency, the value of using temporal measures is that they provide a less subjective and more qualitatively accessible route to assess L2 oral fluency; what is more, the temporal measures of utterance fluency allow to have a reliable characterization of the L2 learners' oral fluency.

What follows is an account of the temporal measures of utterance fluency. Said temporal measures, however, vary in the research and literature of L2 fluency. In spite of the lack of convention in measures of utterance fluency, Kormos (2006) has put together a

compilation the most common measures that have been used in L2 oral fluency research; they are described in Table 2-1 below. In the table created by Kormos (2006, p. 163) here below, it can be observed that the converging measures of L2 fluency include speech rate, articulation rate, phonation-time ratio, mean length of runs, the number of silent pauses per minute, the mean length of pauses, the number of filled pauses per minute, the number of disfluencies per minute, pace, and space. These measures are estimated following certain definitions and formulae; these definitions and formulae are described in table 1 as follows.

Table 2-1 An overview of Measures of Fluency (Kormos, 2006, p. 163)

Temporal Measure	Definition and formula
Speech rate	The total number of syllables produced in a given speech sample divided by the amount of total time required to produce the sample (including pause time), expressed in seconds. This figure is then multiplied by sixty to give a figure expressed in syllables per minute. Riggenbach (1991) suggested that unfilled pauses under 3 seconds should not be included in the calculation of speech rate.
Articulation rate	The total number of syllables produced in a given speech sample divided by the amount of time taken to produce them in seconds, which is then multiplied by sixty. Unlike in the calculation of speech rate, pause time is excluded. Articulation rate is expressed as the mean number of syllables produced per minute over the total amount of time spent speaking when producing the speech sample
Phonation-time ratio	The percentage of time spent speaking as a percentage proportion of the time taken to produce the speech sample (Towell, Hawkins, & Bazergui, 1996).
Mean length of runs	An average number of syllables produced in utterances between pauses of .25 seconds and above.
The number of silent pauses per minute	The total number of pauses over 0.2 seconds divided by the total amount spent speaking expressed in seconds and is multiplied by 60.
The mean length of pauses	The total length of pauses above 0.2 seconds divided by the total number of pauses above 0.2 seconds.
The number of filled pauses per minute	The total number of filled pauses such as uhm, er, mm divided by the total amount of time expressed in seconds and multiplied by 60.
The number of disfluencies per minute	The total number of disfluencies such as repetitions, restarts and repairs are divided by the total amount of time expressed in seconds and multiplied by 60.
Pace	The number of stressed words per minute (Vanderplank, 1993)
Space	The proportion of stressed words to the total number of words (Vanderplank, 1993).

In regard to these measures, Skehan (2003, p.8) points out that these can be classified in three major descriptors of fluency: breakdown fluency, speed fluency, and repair fluency. While breakdown fluency refers to the silent pauses involved in speech, speed fluency indicates speed rate and automatization in the delivery of oral output; repair fluency is related to repetitions, false starts, replacements and reformation in speech delivery (Tavakoli & Skehan, 2005). These three divisions allow a more precise study of utterance fluency, as they reaffirm that that this type of fluency is "a multidimensional construct" (Lahmann, et al., 2015). The measures of utterance fluency associated with these three dimensions can be seen in the table below.

Table 2-2 Measures comprising the three dimensions of utterance fluency.

Dimension	**Measures**
Speed fluency	Speech rate (syllables divided per total time) Articulation rate (syllables divided by phonation time) Mean syllable duration (phonation time divided by number of syllables) Mean length of utterance (in syllables) (syllables divided by silent pauses +1) Mean length of utterance (in seconds) (phonation time divided by silent pauses + 1)
Breakdown fluency	Mean pause duration (total length of silent pauses). Number of pauses per minute (total time). Number of pauses per minute (speaking time)
Repair fluency	Total number of Repairs/Repetitions Total length of Repairs/Repetitions

The measures of utterance fluency that were calculated in the present study are: speech rate (syllables divided per total time), articulation rate (syllables divided by phonation time), mean syllable duration (phonation time divided by number of syllables), mean length of utterance (in syllables) (syllables divided by silent pauses +1), mean length of utterance (in seconds) (phonation time divided by silent pauses +1) for speed fluency; number of pauses per minute (total time), number of pauses per minute (speaking time), mean pause duration (total length of silent pauses) for breakdown fluency; and, total length of repairs/repetitions for repair fluency.

The formulae to estimate the three dimensions of utterance fluency made it possible to explore with more accuracy and precision the relation between L2 oral fluency and working memory capacity on L2Aers in this research. The latter, though, is presented and discussed in 3.4 in Chapter 3.

To sum up, this segment aimed at describing L2 oral fluency from a cognitive stand. Regardless of the different conceptualizations, utterance fluency is a multidimensional aspect of oral fluency that leads to a more systematic analysis and understanding of this complex aspect of L2 acquisition (Kormos, 2006; Segalowitz, 2010). In addition, measures of utterance fluency, as well as the formulae to estimate its three dimensions make it possible to quantifiably and reliably compare and contrast data on L2 oral fluency and working memory capacity under study in this investigation. What follows is an examination of how L2 oral fluency, namely utterance L2 fluency is connected to automaticity.

2.3.3 Second language oral fluency as an aspect of automaticity

As explained in Chapter 1, the main reason for the study of L2 fluency is owed to its connection with automaticity. The development of "automaticity" and/or "automatization" is intricately encapsulated on the theory of L2 emergentism (see 2.1.1.4 for a discussion on L2 automaticity in emergentist theory). Although it is challenging to define automaticity and what it entails for L2 acquisition, DeKeyser's (2001) definition of automaticity encompasses the meaning that it conveys in emergentist theory and research. Thus, DeKeyser (ibid) describes automaticity as

"[the] fast, parallel, effortless, capacity-free, unintentional, result of consistent practice, little interference from and with other processes, unconscious, always based on memory retrieval, does not benefit from further practice, error-free and flexible, strong production rule, no interference from working memory, and no correlation between the mean and the standard deviation in performance measures" (p. 39).

In this line of thought, DeKeyser (1997) had formerly reported that the theory that best supports what automaticity entails is Anderson's ACT1 model of the human cognitive architecture. In this model, knowledge starts out as explicit, or in declarative form, to further become more automatized or procedural (p. 196-197). Regarding the development of automaticity in a second language, DeKeyser (2001) argues that research is limited. Nevertheless, he states that automaticity studies have concentrated on L2 fluency as it is an occurrence that results from automatizing grammar and vocabulary (ibid, p.141).

The previous annotation by DeKeyser is crucial for the current research, as it evidences that L2 oral fluency is a component of automaticity, and thus it serves as a vehicle to explore L2 automaticity. Further testimony to this is what Kormos (2006) discusses in terms of theories of automaticity and speech production. Accordingly, Kormos (2006) argues that automaticity theories attest for two processes that are causative for the attainment of L2

fluency: “automatization of encoding processes and the use of prefabricated language units called formulaic language” (p. 156).

As in DeKeyser’s (1997) formulation of L2 automaticity, Anderson’s (1983, 1995) theory for L2 acquisition serves to expand Kormos’ (2006) assertion on the close connection between L2 fluency and automaticity in L2 processing. In Anderson’s view, L2 speech requires two processes for which procedural or rule-based knowledge plays an important role: syntactic and phonological encoding. Kormos (2006) adds that “studies investigating the development of speech production processes have shown that is indeed in these two steps of encoding where major changes contributing to fluency development take place” (p.158). From this perspective it can be concluded that for L2 fluency to be achieved, the learner has to efficiently use complex grammatical knowledge (L2 phonology and syntax), which according to Anderson and Kormos requires procedural processing of the L2; namely, automaticity.

Though so far, it has been pointed out how L2 fluency is associated with automaticity, there are some gaps in the theory that connects these two aspects. One possible problem with assuming DeKeyser’s (1997, 2001) and Kormos’ (2006) conjectures on why L2 fluency results from L2 automaticity is that it needs further empirical support than just basing on Anderson’s (1983, 1995) declarative to procedural memory model for L2 acquisition. Moreover, as it was described in 2.3.1 and 2.3.2, L2 fluency cannot be simply defined as a unidimensional aspect of language and speech: it is a multimodal (Lahmann, et al., 2015; Skehan, 2003) and very complex characteristic of language (Segalowitz, 2010) (see Figure 2-3 for an illustration on the complexity of L2 fluency). Hence, relating L2 fluency to automaticity on the bases of the dichotomy between declarative and procedural knowledge (see 2.1.1.4 for a description of these theory and concepts) is limiting in scope of L2 fluency and also an abstract definition of L2 automaticity.

Furthermore, DeKeyser’s conceptualization of L2 automaticity needs to be empirically supported to understand up to what extent automaticity: 1) is based on memory retrieval, 2) cannot be modified with practice, 3) implies there are no errors, yet is flexible , 4) is rule-based, 5) is not influenced by from working memory, and 6) does not indicate a correlation between the mean and the standard deviation in L2 tasks involving production. It is important for the purposes of the present research not to take this definition of L2 automaticity for granted; mainly, as I mentioned, because more empirical data is required to attest that automaticity results from a retrieval mechanism. What is more, describing automaticity as

unmodifiable by L2 practice implies that once L2 processing is automatized the learner reaches some sort of glass ceiling or has arrived to a certain L2 destination. This notion is also challenged by the results in Segalowitz and Segalowitz (1993) (discussed below) and others who have observed that automaticity is modified as the L2 learner becomes more proficient (cf. Segalowitz, 2010).

DeKeyser's description of automaticity as error-free, flexible, and based on rule production seems contradictory and highly ambiguous. Additionally, the take that automaticity is not interfered by working memory is very problematic; particularly, because it contrasts with both the premises and the results of the present research. Finally, DeKeyser's point about the lack of correlation between the means and the standard deviations in L2 productions tasks needs to be exemplified with empirical data.

Contrastingly, although Segalowitz (2010) acknowledges the link between L2 oral fluency and automaticity, he argues that the concept of automaticity needs to be rekindled and reformulated. He suggests considering two stands of L2 automaticity that have been approached in cognitive research dealing with L2 fluency; said stands are not necessarily opponent views to automaticity. In one of these perspectives, automaticity is conceived as an inertial flow or motion in the processing; meaning that it is constant and unstoppable. The other perspective describes L2 automaticity as a stability in processing, which requires a re-configuration of underlying cognitive mechanisms to effectively function in the L2.

Segalowitz (ibid) points out that studies such as the one by Favreu and Segalowitz (1983) exemplify the treatment of L2 automaticity as a non-stoppable, constant automatization in processing, or what he calls "ballistic" processing (cf. Segalowitz, 2010, p. 80). In the Favreu and Segalowitz' research, it is found that highly proficient L2 learners with high levels of fluency on both their L1 and their L2 also demonstrated to have a ballistic (non-stoppable) automaticity in the processing of their L2; whereas those learners who were less fluent in their L2 did not show this type of automaticity in their L2.

The authors were able to test ballistic automaticity by exposing both groups of learners to different tasks where they have to make a decision on lexical accuracy after being presented with a prime lexical item that might or might not have a semantic relation with the stimulus word. Favreu and Segalowitz found that the group of L2 learners with higher levels of fluency were able to make more accurate decisions, as they showed that neither the prime word nor the time that they were given to assess the stimulus word had an effect on

the accuracy of their decisions. Thus, Favreu and Segalowitz conclude that L2 automaticity is not related with speed, but rather with an efficient processing of the second language that is unstoppable for every aspect of the L2 that it is required.

In this line of thought, Segalowitz and Segalowitz' (1993) argue that "one can study automaticity by distinguishing processing speed from processing stability" (Segalowitz, 2010, p.85). In their view, automaticity in processing means processing stability, which means that the underlying mechanisms carrying out L2 tasks have been and are being re-structured and re-organized not only for a faster, but also for a more efficient processing. In order to test that automaticity equals processing stability, the authors propose a shift in how research on L2 fluency and automatic processing is conducted.

Segalowitz and Segalowitz (1993) rationalized that automaticity is usually measured basing on the L2 learner's means of reaction times (RT) and standard deviations (SD). They observed, though, that there is always a linear correspondence between means of RTs and SDs among the participant's responses; so, if the L2 speech production task required a faster response, the means in RTs and SDs would increase and decrease depending on the speed required for the task. Instead, Segalowitz and Segalowitz (1993) proposed that research on L2 fluency and automaticity should be taken beyond and concentrate on the coefficient of variation (CV), which is the resulting ratio between the means of reaction times and the means of standard deviations, and its correspondence with the L2 learner's RTs alone. Using coefficient variations and reaction times, according to Segalowitz and Segalowitz, is a better formula to assess changes in underlying mechanisms that have to do with re-organizing and re-structuring in L2 processing.

The results in their study show a significant difference between RTs and CVs among more fluent learners when performing in L2 tasks; the authors report that there is no significant difference between CVs and RTs among these learners when the task is in their L1. Segalowitz and Segalowitz (1993) conclude then, that automaticity is not only related to fast processing, but also to efficiency and optimal operation; they add that automaticity, as a stability processing is observable when the learner performs in their L2 given that the underlying mechanisms of their L1 have been stabilized.

The interpretations of automaticity in the research of Favreu and Segalowitz (1983) and Segalowitz and Segalowitz (1993) here discussed are of great importance to fathom the depth in connection between L2 fluency and automaticity. Additionally, and equally

important to the present study, their empirical findings illustrate that L2 automatic processing is more than cognitive speed; L2 automaticity represents that learners have managed to re-structure and re-organize their underlying grammatical mechanisms to make them more adequate and efficient for L2 usage.

All in all, the present segment has discussed that second language oral fluency is an observable aspect of language that is intimately connected to automaticity (Anderson, 1983, 1995; DeKeyser, 1997, 2001; Kormos, 2006; Segalowitz, 2010; inter alia). In this sense, there seems to be a latent agreement in the emergentist literature, as well as in other cognitive theory, that L2 fluency serves as a vehicle to expand the knowledge on what automaticity entails in L2 acquisition. It was also argued that the definition of automaticity should not be simplified to cognitive proceduralization of L2 knowledge, or as a glass ceiling in L2 acquisition as some theorists conceptualize it (see DeKeyser, 1997, 2001). For this reason, the treatment of automaticity in Favreu and Segalowitz (1983) and Segalowitz and Segalowitz (1993) (inter alia) discussed above is highly significant for the present research as it presents L2 automaticity as an aspect that implies a restructuring of cognitive mechanisms, which make an effective functioning in the L2 possible.

Nonetheless, as it was shown in these studies, the levels in L2 fluency vary even among groups of learners with similar acquisition trajectories (age of L2 acquisition, age of L2 onset, learning environments, etc.). Therefore, it can be concluded that L2 fluency, and by default automaticity, are levels of acquisition that are not reached by all second language late learners. That is why, the current study aims at exploring if the variation in L2 fluency is related to individual differences in working memory capacity (WMC) (see 2.2.1 for a detailed description of working memory capacity). Granted, this section has reaffirmed the notion that L2 fluency is a complex aspect of L2 acquisition that requires L2 automaticity; this in itself represents not only a speed up in cognitive processing, but also a major reorganization and resetting of underlying mechanisms that lead to efficient L2 processing. For this reason, finding a correlation between L2 fluency and WMC among L2Aers might suggest that for automaticity to be developed, L2Aers require a higher capacity in working memory while acquiring their second language.

2.3.4 Research on second language oral fluency

As stated in the previous segment, one of the research objectives of the present work is to explore the relation between L2 oral fluency and working memory capacity on L2Aers;

this is due to the variation in the attainment of this language aspect among this group of learners. Moreover, as it was concluded in 2.3.3 above, the reason for exploring L2 oral fluency with respect to WMC among L2Aers is that L2 oral is a product of automaticity; which might only be reached through high capacity in working memory among these learners. This research objective, however, requires analysing these two variables from a quantitative research point of view.

Consequently, as it was argued in 2.3.2, L2 oral fluency can be systematically studied by focusing on its utterance fluency component. This multimodal type of L2 oral fluency can be systematically approached since, as previously explained, it is composed of measurable fluency features such as speech rate, length of pauses, number of repairs, etc. (see Table 2-1 for a description of these features). This segment is intended to provide an account of some research in second language utterance fluency. Hence, the present segment is intended to discuss some relevant empirical studies on L2 utterance fluency in regard to the research objective concerning this aspect in the current work.

In this respect, a pioneering work of L2 fluency analysis is the study by Towell et al. (1996). In their study, twelve English speaking late learners of French were considered. The aim of their study was to spot variables of fluency that included: speech rate, phonation/time ratio, articulation rate, and mean length of runs. To extract these measures, participants were asked to retell a story from a short, mute motion picture. Participants were recorded at the start of a-year-abroad study program in a French-speaking country. A year later, the learners were recorded again retelling the same story in French, their L2.

The results in the Towell et al. (1996) study show that performance on articulation rate, mean length of runs, and speech rate increased significantly from the first time of data collection to the second. Important to notice is that results do not yield any significant difference in terms of phonation/time ratio from one sample take to the other. Another finding is that mean pause length did not change from their first data collection to their second. The authors conclude that speech rate increases were determined by the increases in the mean length of runs, and not by a decrease in pauses.

Granted, Towell et al. (1996) reported that fluency was related to procedural knowledge (based on Anderson's, 1983 ACT Theory), as well as they comment that their findings are useful to know what cognitive mechanisms are involved in L2 oral fluency. These conclusions are relevant to the present research given that they illustrate that measures of

L2 utterance are closely related to proceduralization, which is a component of automaticity. These results also highlight the lack of assertion in DeKeyser's (2001) definition of automaticity (see 2.3.3 for a direct quotation on DeKeyser's definition of automaticity).

Evidently, the increase in speech and articulation rates after a year of exposure to the L2 among the learners in the Towell et al. (1996) study indicates that procedural knowledge, or automaticity is not an ending result when learners are developing their L2 speech production skill; so, DeKeyser's claims about automaticity not being affected by practice, and/or by working memory (WM) need to be reformulated. The results in the present study complement this study as they show that measures related to speech and articulation rate significantly correlate with higher capacity in WM among intermediate L2Aers.

Other studies have offered a wider panorama in terms of the assumptions that can be drawn from L2 fluency measures. Owed to the availability of more advanced computer-based programs, it has become easier to analyse and measure utterance fluency features. An example of this is the study performed by Cucchiriani et al. (2002) who used automated software techniques to make a comparison between L2 read and spontaneous speech among beginner and intermediate L2 learners. The authors found that speech rate and phonation/time ratio were significant correlates of fluency among L2 beginner learners, while mean length of runs was significant only among intermediate learners. Although, further tasks might be required to have a better interpretation on the results of this study, it can be observed that L2 fluency has features that are not necessarily task or skill dependent. In this sense, there is a possibility that late learners apply similar strategies to cope with activities that require L2 fluency; though, as learners increase their L2 proficiency, their strategies become more specialized.

The latter adds to the Segalowitz and Segalowitz' (1993) perspective on processing stability in that higher levels of fluency or proficiency in L2Aers are linked to a re-structuring of cognitive mechanisms that may be deemed as more effective for L2 use (see 2.3.3 for a further discussion on processing stability). Thus, these results add to the notion of L2 automaticity as processing stability, which is a conceptualization of automaticity adopted in the present work. Overall, the study by Cucchiriani and colleagues contributes to the quantitative study of L2 fluency. In addition, it shows that automated programs that measure utterance fluency allow for the study of L2 fluency in larger groups of learners, which leads to more reliable knowledge of L2 acquisition (Segalowitz, 2010, p. 34).

Another important study is the one by de Jong et al. (2009b) who also used an automated software to analyse samples of speech to calculate L2 fluency measures. Of particular interest was the inclusion of multiple elicitation tasks to obtain speech data in both the L1 and the L2 of English and Turkish speaking learners of Dutch. Over the samples obtained, the authors calculated multiple fluency measures that included syllables and words per second (including and excluding filled pauses), and length of silent pauses.

The first finding in the De Jong et al. (2009b) study is the significant correlations between L1 and L2 fluency performances among their participants; specifically, on the measures of syllables per second and phonation time. Also, they found correlations for length of pauses, and filled pauses; no significant correlations were found for silent pauses per word between the L2 and the L1. The second relevant finding is the results in terms of "differences scores", which are related to differences in proficiency between the L1 and the L2. To calculate said difference scores, de Jong et al. (2009b) calculated effect sizes by converting the Cohen's *d* to R^2. Over this, they found significant effect sizes for words per second and for percentage of silent pauses per word. The final important result observed in their study was that repair fluency measures did not significantly correlate between the L1 and the L2; however, they did find effect sizes in the percentage of words corrected when difference scores were estimated.

Overall, the findings in De Jong et al. (2009b) as interpreted by Segalowitz (2010) is that "the oral variables best reflecting L2 fluency, using the L1 as a baseline -at least for the English-Dutch L1-L2 language pair- are effect sizes for L1-L2 differences on three measures - percent of silent pauses per word (but not length of silent pauses); words per second speech rate, especially excluding filled pauses; and percentage of corrections or self-repairs per word" (p. 36). The results in the de Jong et al. (2009b) are significant as they reflect what aspects of utterance fluency are shared in the L1 and the L2; in this case, measures of breakdown fluency (cf. Skehan, 2003) such as a length of pauses and filled pauses can be transferable characteristics from the L1 to the L2, and thus, are not attributable to the L2 acquisition process. Nonetheless, it also serves to show that there are certain features of fluency that are indeed a part of the L2 process of acquisition, such as words per second and the percentage of silent pauses per word (measures that are more related to speech fluency).

De Jong et al. (2009) argue that similarities reflect the existence of individual differences on fluency. This means that fluency is an aspect common to the two languages and not just specific to the L2. More recently, Tavakoli and Wright (2016) report that "recent

research findings highlight L2 fluency as a reliable predictor of L2 proficiency […], but also as a characteristic that retains some traits of L1 speech production […]" (p. 1). The latter adds to the notion of automaticity as processing stability; as Segalowitz and Segalowitz (1993) sustain, L2 fluency improves as learners develop more automaticity. However, such development in automaticity is only observable in the learners' L2 given that certain aspects underlying L2 fluency are just a reflection of the L2Aers' fluency in their native language. Hence, the development of automaticity might rely on the learner's ability to change and/or improve those characteristics of fluency that are apparently more related to speech rate or speed fluency as in the De Jong et al. (2009b) study since these are measures that are linked to the learners' L2 proficiency.

As I recognize that using the L2Aer's L1 as a baseline to understand what aspects of their L1 fluency are reflected on their L2, I did not consider the extraction of fluency measures of the participants' L1 for the current research as in the De Jong et al. (2009b) study. Mainly, the reason for not taking into account the L2Aer's L1 as a baseline responded to the limitation that I faced in terms of time, human and economic resources as this type of analysis requires extensive hours or transcription and the calculation of temporal measures in a special software such as PRAAT (cf. De Jong et al., 2009a); due to the constraints faced when undergoing empirical research for doctoral purposes in a distance program, it was difficult to meet the conditions to be able to extract the participants' L1 speech samples and, most importantly, to analyse such data in terms of temporal measures.

However, in order to achieve the purpose of the present study, I considered the inclusion of L2 speech samples of L2 adult learners at two distinct levels of proficiency as well as the speech samples of native speakers (whose samples can serve as a baseline to compare the L2Aer's fluency). The latter consideration is more in line with a recent study by De Jong (2016), who investigated the distribution of silent and filled pauses in L1 and L2 speech. For this, De Jong (2016) collected the speech samples of 25 Turkish and 29 English speaking learners of Dutch; and of 18 native speakers of Dutch who served as a baseline for comparison to the L2 speech samples. The participants were asked to respond to eight different speech generation tasks, which varied in terms of formality, difficulty, and discourse mode as reported by de Jong (2016). In order to analyse the speech samples, the CLAN software was used (cf. MacWhinney, 2000); the pauses were added manually to the transcriptions of the speech samples.

Important for the data analysis was that the transcriptions were also classified into analysis of speech units (ASU) (cf. Foster et al., 2000); ASUs are utterances that comprehend an independent or subordinate clause, or a subclausal unit. De Jong (2016) used these ASUs to indicate where the major boundaries in speech were located. The measure of silent pauses was then categorized as occurring between or within an ASU. Moreover, the transcriptions were also analysed in terms of filled pauses, which included non-lexical forms such as "uhm", "uh", "er", etc. In addition, as part of the second major goal of the study, the transcriptions were analysed in terms of pauses and word frequency; thus, silent and filled pauses were also divided into those that occurred before "low frequency" or "high frequency" words.

The results in the de Jong (2016) study show that L1 and L2 speakers pause in a similar manner at ASU boundaries; however, L2 speakers tend to pause more within ASUs compared to the native speakers of Dutch. De Jong (2016) notices that "as learners were more proficient, they produced fewer silent pauses within ASU's" (p. 54). The second relevant finding was that both L1 and L2 speakers were more propense to pause before low-frequency nouns; and, in the case of the L2 speakers, their proficiency level did not have an effect on this occurrence. Overall, as reported by De Jong (2016), L2 speakers' samples were characterized by containing more high-frequency nouns than those of the L1 speakers, who used more low-frequency nouns.

De Jong (2016) interprets these results as an indication that both L1 and L2 speakers pause at ASU boundaries to plan their conceptual message (semantic and discursive information); however, L2 speakers need to pause within ASU boundaries as well given that they need to work on formulating the linguistic message (morphosyntactic information). The latter coincides with the differences in the underlying processing stages between L1 and L2 speakers based on the Levelt (1989, 1999) model and its adaptation to the L2 speech by De Bot (1992). Furthermore, L1 and L2 speakers tend to pause more before low-frequency nouns given that they might require some time to retrieve these kinds of lexical items; as interpreted by De Jong (2016).

As De Jong (2016) emphasizes, the implications of the latter findings are that the main difference between L1 and L2 fluency occurs at the linguistic planning level; and it is at this level where L2 proficiency has an effect: as L2 proficiency increases, the less pausing there will be within ASU boundaries. First, these results are related to this research work in that the level of proficiency does seem to be related to pausing; as I will demonstrate and discuss later, the L2Aers in my study did seem to differ significantly in pausing time with respect to

the advanced L2 and L1 speakers. However, as my findings demonstrate, IDs in the L2Aer's WMC serve to explain the variability in both pausing and speed fluency measures (only) at the intermediate level.

Second, the consideration of using L1 speakers' speech to compare L2 speech regarding fluency (as in the de Jong, 2016) is helpful to understand what aspects of L2 fluency are characteristic of the L2 acquisition process. Nonetheless, as in the present study, they are also useful to analyse the effect of general cognitive mechanisms, such as WMC, on L2 oral fluency in comparison to L1 oral fluency. As I mentioned before, the present research was developed with considerable time and aid constraints; still, its findings can complement the results in studies such as the one by de Jong (2016) in the sense that not only is it important to understand L2 fluency in terms of proficiency levels, but also in terms of the variability observed in this aspect among L2Aers.

To conclude, these studies are samples of the shift that has been made in terms of L2 fluency from a quantitative research perspective under a cognitive approach. Particularly, the focus on utterance fluency offers a more systematic way to approach this complex aspect in L2 acquisition. Therefore, these studies are central for the present research work as they underscore the relevance of conducting research based on quantifiable features of fluency as these lead to a more reliable analysis and knowledge on the underlying mechanisms involved in L2 acquisition, and the variation in attainment among second language late learners.

2.4 Resumptive pronouns

The previous section, 2.3, was intended to describe L2 oral fluency since one of the two specific objectives of the current work is to explore the effects of working memory capacity (WMC) on L2 oral fluency. As it was established in 1.2, this is an aspect in L2 output that is differently accomplished by L2Aers. Nonetheless, the second research objective here pursued is to explore the effects of WMC on L2 grammatical aspects; given that certain features of the grammar are attained only by some L2Aers. Given that resumptive (R) pronouns are 1) an uninterpretable feature, which differs in terms of grammatical strategy from Spanish (L1) to English (L2) (cf. Leal-Mendez & Slabakova, 2012), and 2) possible "saving devices" that can alleviate the parsing load imposed by the antecedent and its referent in long-distance sentences (cf. Ariel, 1999; Hawkins, 2004; Alexopoulou, 2009; Alexopoulou &

Keller, 2007, 2013), the second research objective is to focus on assessing if the acceptability of this L2 feature is related to WMC.

2.4.1 Description and definition of resumptive (R) pronouns

This subsection will start with a general recount on how resumptive pronouns have been approached in the literature. To begin with, it is important to mention that the definition of resumptive (R) pronouns has changed over the last three decades. For example, much of the research on resumptive pronouns in the 90s focused on the distinction between gap constructions and resumptive constructions (Rouveret, 2011). For the most part, resumptive constructions were treated as a property of movement.

Further research on resumptives indicated that the properties of these type of pronouns varied from one language to the other (ibid). Owed to this, some literature in linguistics has focused on studying the set of languages where the syntactic properties of R pronouns are akin to the relation between a trace and its wh-antecedent. Some other researchers have focused on studying the set of languages in which R pronouns have a pronominal behaviour (Rouveret, 2011). Nevertheless, Rouveret (2011) argues that "the label resumptive pronoun usually refers to the overt pronominal elements found in some languages in the variable position of unbounded Ā-dependency constructions - the latter includes relative clauses, constituent questions, comparative clauses, dislocation and focus constructions-" (p.2).

Therefore, one distinction that can be made about resumptive pronouns is that they can either act as a "saving device, avoiding the violation of a general principle", or they can act as a "productive strategy to form Ā-dependencies" (Rouveret, 2011, p.6). Consequently, languages can be classified in terms of the behaviour that R pronouns have in their grammatical system. In other words, there can be a category for languages that use resumptives productively (used where a gap is legitimate), and a category for languages that use them as a saving device (used to avoid violations on grammatical principles) (Rouveret, 2011). As an example, Rouveret reports that Hebrew and standard Arabic enter the classification of languages that use resumptives as productive, whereas English enters the category that uses them as a saving device.

In this line of thought, since the present study focuses on the acceptability of resumptive (R) pronouns in English, it is useful to illustrate how R pronouns act in English. In

this regard, Creswell (2002) explains that English speakers place an R pronoun "in place of the gap or trace which would be an island-violation if a wh-word had been extracted from that position" (p.102). The function of R pronouns as a remedying or saving device in island violations is exemplified in sentences 1a, 1b, 1c, and 1d (Creswell, 2002, p. 102) below, in which they occur on a relative, an adverbial, and a subject clause respectively (R pronouns are marked in bold letters).

(1) a. There are always guests who I am curious about what **they** are going to say. (Prince, 1990)

b. That asshole X, who I loathe and despise the ground **he** walks on, pointed out that… (Prince, 1990)

c. Apparently, there are such things as bees in the area which if you're stung by **them**, you die. (Prince, 1990)

d. You have the top 20% that are just doing incredible service, and then you have the group in the middle that a high percentage of **those** are giving you as a good day's work… (retrieved from http://www.ssa.gov/history/WEIKEL.html)

Moreover, Leal-Mendez and Slabakova (2014) mention that given the property of R pronouns as saving devices of subjacency violations, English does not permit them (R pronouns) in structures where there is not a violation of subjacency. The authors illustrate the latter in examples 2a and 2b below (Leal-Mendez & Slabakova, 2012, p.3) (* indicates ungrammaticality).

(2) a. *I saw the boy that Mary loves him

b. I saw the boy that Mary loves__

In the same vein, Creswell reports that R pronouns "appear in relative clauses in English in non-island violation contexts" (p. 103). The author exemplifies the latter in sample sentences 3a, 3b, and, 3c below, extracted from Prince (1990) (Creswell, 2002, p.103) (R pronouns are marked in bold).

(3) a. My son, God bless him, he married this girl which I like **her**.

b. If there's any message that she can forward **it** to us, then…

c. I have a friend who **she** does all the platters.

Furthermore, Leal-Mendez and Slabakova (2014) add that R constructions are those "in which the moved constituent has landed too far away from the position where it started"

(p. 3). This can be observed in the examples provided by Ross (1967), as shown in sentences 4a and 4b below.

(4) a. I just saw the girl who Long John's claim that **she** was a Venusian made all the headlines.

b. The only kind of car which I can never seem to get **its** carburettor adjusted right is them Stanley Steamers.

Nonetheless, it is important to describe how R pronouns act in Spanish since this is the native language of the participants in this study. According to Suárez Fernández (2013), resumptives in Spanish can appear in non-restrictive restrictive clauses, and can be applied to any syntactical function; but they are not obligatory in any of them (p. 64).

This is reiterated in Leal-Mendez and Slabakova (2014) who add that although R pronouns are "legitimately used to save island violations" (p. 543), they might or might not appear in the sentence given that Spanish is a null-subject type of language. For example, in the sentences 5a and 5b from Senn (2004, p. 131) below, the R pronoun can be removed without altering the grammaticality of the sentences (the R pronoun is highlighted in bold) ("_" indicates where the R pronoun has been omitted).

(5) a. Quiso que le arreglara a Ana el vestido$_i$ que no me acuerdo dónde **lo** compré.
Quiso que le arreglara a Ana el vestido que no me acuerdo dónde **_** compré.
Wanted-1sg that DAT.fix-3sg to Ana the dress that not NOM.remember-1sg where **[it]** bought-1sg
'[Someone] wanted me to fix the dress for Ana that I don't remember where I bought.'

b. Fuimos a pasear por una carretera que **la** recorríamos cuando éramos chicos.
Fuimos a pasear por una carretera$_i$ **que _ recorríamos** cuando éramos chicos.
Went-3pl to travel by a highway that **[it].**DAT travelled-3pl when were-3pl kids.
'We took a trip over a highway that we travelled by when we were kids.'

As in English, R pronouns in Spanish function as a saving device (Rouveret, 2011) and do not admit subjacency violations within the sentence. Additionally, as in English, resumptives are more commonly found in colloquial spoken varieties of Spanish (Leal-Mendez & Slabakova, 2014; Senn, 2004; Suárez Fernández, 2013; Suñer, 1998). However, Leal-Mendez and Slabakova (2014) explain that "Spanish is a null subject language that has clitic-like subject agreement and object clitics" (p. 8). In English, on the other hand, when an R pronoun appears in a sentence, it does as an overt pronoun given that English is a non-null subject language; in Spanish, R pronouns can be assimilated on the verb and/or appear as clitics.

Examples of the usage of R pronouns as an alleviating strategy in English can be observed in 4a and 4b extracted from Ross (1967). On the other hand, sentences like 6a and 6b taken from Cerron-Palomino (2009) can serve as examples to demonstrate that the resumption strategy in Spanish is more syntactic given the use of subject and object clitics. The latter can be seen in the case of 6a, with "qué libro" (which book), which agrees in gender and number with "lo", the subject clitic; in 6b, "un tipo" (a guy) agrees in gender and number with "él" the object clitic. The clitics working as resumptives have been emphasized in bold letters.

(6) a. Qué libro$_i$ te preguntas quién **lo** escribió?
Which book you ask-2sg who it wrote
'Which book do you wonder who wrote **it**?'

b. Conozco a un tipo$_i$ que **él** me aconseja a mí.
Know-1sg ACC a guy that he me.DAT advises to me
'I know a guy that he advises me.'

Overall, R pronouns are elements that may vary in grammatical strategy from one language to the other. In some languages, R pronouns have a productive function; whereas in some others, like English and Spanish, they are said to function as "saving devices". Nonetheless, even when it appears that R pronouns are used similarly in English and in Spanish, the resumption strategy differs among these two languages at the grammatical level.

On the one hand, R pronouns may appear overtly in sentences in English in A-dependencies where a gap is legitimate and in which there is not subjacency violation. On the other hand, in Spanish, R pronouns might not overtly appear in the sentence given that Spanish is a null-subject language, and thus, can be assimilated onto the verb or appear in the form of clitics. These differences are relevant to the present work since the purpose is to explore if L2Aers' variability in grammatical behaviours can be explained by their individual differences in WMC.

2.4.2 R pronouns from an SLA generative perspective

The purpose of this section is to expand on the complexity that R pronouns represent in the context of L2 grammars; mainly for adult Spanish Speaking learners of English under generative approaches to SLA (from here on referred to as "genSLA"). With this in mind, I deem it important to first have a recapitulation of the theory surrounding L2 grammars and late learners in genSLA since, as I have discussed before, generative research to SLA has demonstrated that L2Aers are capable of comprehending morphosyntactic features that are not necessarily transferable from their L1 to their L2, and that are not accessible through the

information provided by the L2 input. Nonetheless, under the genSLA perspective, there are views that seem to oppose the notion that late learners are able to fully access all of the L2 grammatical features; and thus, sustain that L2Aers' grammars are characterized for being incomplete compared to the grammars of native speakers.

In order to set the bases for the differing views in genSLA, it is necessary to touch on certain aspects of the generative grammar theory. In the most recent linguistic framework, the Minimalist Program (Chomsky, 1995), features of the language are classified in two categories: interpretable and uninterpretable. On the hand, interpretable syntactic features "are those which, while relevant to the syntactic computation, are also used by the semantic component in determining the meaning of syntactic expressions" (Hawkins & Hattori, 2006, p. 270).

According to Hawkins and Hattori (2006), third person, past tense and questions are examples of interpretable features in the English language because they are useful to convey both the grammatical and semantic sense of a clause or sentence; these types of features can also be grasped and/or explained by the information provided on the input or the surface structure. On the other hand, uninterpretable features "are those that are purely grammatical and only relevant to the morphosyntax" (Leal-Mendez & Slabakova, 2012, p. 2). According to Leal-Mendez and Slabakova (2012), case and agreement are examples of uninterpretable features.

In terms of grammatical features, there are, some authors that claim that L2 late learners can have knowledge of properties of the L2 grammar, being interpretable and/or uninterpretable; but, that this subset of learners will occasionally show problems at making associations with the correct morphological or phonological form (e.g. Lardiere, 1998; Prévost & White, 2000; Goad & White, 2004). For example, Lydia White (2007) claims that "if it can be shown that L2ers acquire abstract and subtle properties that are underdetermined by the L2 input, this suggests that interlanguage competence must be subject to the same constraints as native competence" (p.36). This is demonstrated in White and Juffs (1998) who also propose that if L2 learners signal the processing of linguistic features that only operate under the constraints of the L2 (and are not available in their L1), they have access to UG.

To do so, Juffs and Rodriguez (2014) report that findings from recent research put forward the possibility that highly fluent L2 learners, who have reached this level via naturalistic linguistic exposure, show knowledge of abstract grammatical properties of the

L2, which require a deep structural processing ability (p. 137). An example of this research is Juffs and Harrington's (1995) study, which examined the processing of grammatical and ungrammatical wh-movement in 25 Chinese-speaking learners of English. The authors find that these L2 learners show an accurate judgment on wh-movement and knowledge of constraints. Thus, the acquisition of L2 complex features (abstract grammatical properties) is at ad rem in research with a generative approach.

Nevertheless, other authors who have approached L2 acquisition from a generative perspective argue that L2 late learners' grammars will be characterized for being incomplete, compared to those of the native speakers; particularly at the morpho-syntactic level (e.g. Tsimpli & Roussou, 1991; Hawkins & Chan, 1997). Based on the Interpretability Hypothesis (cf. Tsimpli & Mastrapavlou, 2007), Tsimpli and Dimitrakopoulou (2007) claim that uninterpretable features that are not part of the repertoire of the L2Aer's L1 will not be a part of their L2 grammar. Under the Interpretability Hypothesis, L2 late learners' grammatical knowledge will be constrained due to maturational processes that are based on the assumptions made in the critical period hypothesis for language acquisition (cf. Johnson & Newport, 1989; Smith & Tsimpli, 1995; Meisel, 1997). More specifically, "[the Interpretability Hypothesis] maintains that uninterpretable features are subject to critical period constraints and, as such, they are inaccessible to L2 learners" (Tsimpli & Dimitrakopoulou, 2007, p. 224).

Explained in minimalist terms, L2Aers can understand features that are interpretable at Lexical Form (LF) since these grammatical properties "are represented both in the language system and in the LF- interface, implying that they have a dual status in the mental lexicon: a linguistic and a conceptual one" (Tsimpli & Dimitrakopoulou, 2007, p. 224). In other words, L2Aers do have access to grammatical features that can be comprehended through the information provided by the L2 input. For example, Spanish speaking learners of English will be able to comprehend subject-verb agreement given that the number and person morphemes are available in LF, and also because they have semantic and grammatical value to convey the meaning of the sentence (plus, this parameter is available in their L1 Spanish). On the other hand, these same speakers will have problems with sentences with a gap condition in English since the necessary information to interpret this feature is only available through the grammar and not in LF.

Of particular interest for the current investigation is, precisely, the study by Tsimpli and Dimitrakopoulou (2007) (from now on referred to as T&D (2007)) in which the Interpretability Hypothesis for L2 language learners is tested in the acceptability of

resumptive pronouns among 48 Greek speaking learners of English and 26 native speakers of English. T&D (2007) argue that resumptive pronouns will be incorrectly accepted by L1 Greek speakers owed to the following: 1) Greek has a 3rd person subject agreement and object clitics (which are a bundle of uninterpretable features), and this will cause them to incorrectly accept subject and object pronouns in gap position of wh-interrogatives in English; and, 2) Greek has an optionality for resumptive object clitics and an obligatory subject-verb agreement, thus Greek learners will be more accepting of R pronouns in subject than object position. In other words, T&D (2007) based their study on the premise that the learners will use the same resumptive strategy of their L1 (Greek) on their L2 (English) given that this parameter will resist resetting given the uninterpretability conditions of this property between the L1 and the L2.

In order to test their predictions, T&D (2007) asked participants to indicate the degree of acceptability of 51 sentences (30 test items and 21 fillers) containing resumption in subject and object position, and sentences with a gap in resumptive position. The study also tested animacy and discourse-linking (d-linking) effects with the premise that these properties might aid Greek learners to correctly accept the sentence conditions given their LF-interpretability status. The results corroborate T&D's (2007) predictions based on the Interpretability Hypothesis as they find that L2 learners at lower levels of proficiency were more accepting of R pronouns (regardless of site of extraction), whereas most of the more proficient learners rejected R pronouns in object position.

However, the L2 learners significantly differed from the L1 English speakers in that the former were more accepting of R pronouns in ungrammatical sentences, as well as of ungrammatical sentences with a gap condition. These results led T&D (2007) to conclude that "uninterpretable formal features, such as (subject, object) agreement, cause learnability problems even at advanced levels of acquisition" (p. 237). For the authors, L2Aers "impose" the resumptive strategy in their L2 English grammar as they are following what T&D (2007) call a "process of morphological misanalysis of these L2 items" (p.237).

Nonetheless, Leal-Mendez and Slabakova (2012) do not advocate for the Interpretability Hypothesis, and instead argue that late learners are able to correctly accept formal uninterpretable features of the L2 in spite of the lack of availability of the same or similar features in their L1. For Leal-Mendez and Slabakova (2012), the T&D (2007) study needs to be replicated because assumptions made by the Interpretability Hypothesis regarding L2 acquisition need to be carefully considered in light of the multiple findings that

prove that L2Aers are able to successfully reset their parameters and acquire even formal non-interpretable features of the L2 (cf. Bolotin, 1996; Li, 1998; Martohardjono, 1993; interalia).Hence, Leal-Mendez & Slabakova (2012) propose that "learner performance on resumptives has to be evaluated in its own right, and it has to be established whether learners display a contrast between acceptable and unacceptable sentences" (p. 16).

In order to study the acceptability of R pronouns among L2Aers, Leal-Mendez and Slabakova (2012) considered the grammatical judgment of 29 Spanish speaking learners of English on 10 items containing resumptive pronouns in both subject and object position, and 4 fillers. The learners were tested in terms of proficiency and classified accordingly. Also, Leal-Mendez and Slabakova (2012) divided the learners into those who were accepting of the R condition in their L1 (Spanish), and those who were not. Crucial for the present investigation are the assumptions that the authors make about the grammatical description of R pronouns in English and Spanish. Leal-Mendez and Slabakova (2012) find that in Spanish, as in Greek (the L1 used in the T&D, 2007 study), R pronouns are found in "clitic-like subject agreement and object clitics" (p. 8); and, said subject and object clitics can be assumed to reflect "uninterpretable features on the functional categories T and little *v*" (p.8). The learnability task for L2 Spanish-speaking learners of English is the same as for the Greek speaking learners of English in the T&D (2007) study: "to realize that English is a non-null subject language and lacks the resumptive strategy" (Leal-Mendez & Slabakova, 2012, p. 4).

The results in the Leal-Mendez and Slabakova (2012) study indicate first, that there is no significant difference in performance between the L2 learners and the L1 speakers; second, both groups of learners, those who accept R pronouns and those who do not, show a grammatical contrast with regards to ungrammatical and ungrammatical sentences containing resumptives. Lastly, and most importantly, those learners who accept the R condition in their L1 were significantly more likely to accept sentences with resumption in English. Moreover, Leal-Mendez and Slabakova (2012) found that those learners with an advanced L2 proficiency and who do not accept R pronouns in Spanish were better at distinguishing the grammaticality/ungrammaticality conditions on the stimulus sentences. The authors interpret their findings as negative evidence for the Interpretability Hypothesis and consequently as contradictory to the results in T&D (2007). Therefore, Leal-Mendez and Slabakova (2012) conclude that L2 learners can successfully learn properties of the L2 grammar that involve formal non-interpretable features; what is more, the authors add that "the linguistic analysis of the phenomenon of resumptive pronouns in Spanish, Greek and

English is related to the appropriateness of looking for a contrast within interlanguage grammars (and not so much between learners and natives)" (p. 16).

The findings in both the T&D (2007) and the Leal-Mendez and Slabakova (2012) are relevant for the current research in that they explain, from different perspectives of genSLA, what aspects of the L2 grammar pose a challenge for L2Aers. What is more, based on linguistic principles of the cognitive architecture of grammars (which are mainly guided by the Minimalist Program), they offer an explanation as to why L2 learners' grammars are built in the way they are. As I discussed, formal grammatical properties such as resumptives have been used to show that L2Aers' grammar will be constrained by the availability or lack thereof of the uninterpretable features of their L1; nonetheless, studies focusing on the same property seem to support the notion that L2Aers are able to accurately comprehend features that are not available in their L1 and that are non-interpretable in the L2 as in the Leal-Mendez and Slabakova (2012) study.

Although in the latter, there seems to be an explanation for the variability in grammatical comprehension among L2Aers (namely that the accurate distinction of L2 uninterpretable features depends on the individual L1 grammars of late learners), it is still not clear why L2 grammars vary among L2Aers. If L2Aers' grammars are formed based on the individual's L1 grammatical preferences, then the present study on L2 grammaticality should yield similar results as the Leal-Mendez and Slabakova (2012) study. However, as Leal-Mendez and Slabakova (2012) point out, "if resumptive strategies are a processing phenomenon, as some analyses would have it, then the differences between Greek and Spanish on one side, and English on the other, are a matter of degree and the learners have nothing to learn apart from aligning their processing strategies to the L2 ones" (p. 16).

That is why, in my view, the treatment of R pronouns should go beyond the scope of its uninterpretability classification in generative terms as this might not be enough to understand variability in how L2Aers parse these types of linguistic features when learning a second language. In other words, for as much as genSLA theory can explain what morphosyntactic properties L2Aers are able to attain at the competence level, said explanation requires a further examination of these grammatical properties at a cognitive level to comprehend: 1) what factors are demonstrably (theoretically and empirically) involved in the parsing of R pronouns, and 2) how these factors affect the parsing of these features amid L2Aers.

Therefore, an approach to R pronouns beyond its classification as uninterpretable features will be offered in the following section to establish their potential interaction to WMC. The latter will ultimately serve to set the ground for the proposal of the current research in terms of how differences in WMC affect the acceptability of R pronouns; thus, offering a possible explanation as to what causes variability at the grammatical level among L2Aers.

2.4.3 R Pronouns and their implications for the L2 grammatical parser

Nonetheless, as I mentioned before the analysis of R pronouns in the context of L2 attainment also needs to be considered with respect to what it represents in terms of processing for both L1 speakers of English and L2 speakers of Spanish. With this in mind, Alexopoulou (2009) points out that "intrusive resumption has been viewed as a strategy to overcome processing complexity involved in long-distance Ā-bar dependencies" (p. 488). The author adds that the view of resumptives, and the embedded clauses to which they belong need to be considered "at the interface between grammar and processing rather than solely due to grammatical principles". Therefore, not only are R pronouns abstract grammatical features, but also "devices" that might help with the difficulty of processing embedded clauses (McCloskey, 2006).

Other authors have made similar observations with regards to what R pronouns represent for processing. For example, Ariel (1999) and Hawkins (2004) state that R pronouns are utilized to overcome complexity owed to their "accessibility" (cf. Ariel, 1990). In their view, an R pronoun being the head of a relative clause facilitates the realization of an antecedent, or makes it more accessible, than the lack of an R pronoun (or a gap) in the same position. However, in a study by Alexopoulou and Keller (2002), sentences with a gap condition were significantly more accepted than sentences with

resumption.

In their study, Alexopoulou and Keller (2002) set out to "investigate the nature of the interaction between resumption and islands, establish the "saving" effect of resumptives and investigate these interactions from a crosslinguistic perspective" (p. 2). In order to do so, Alexopoulou and Keller (2002) conducted two experiments; one involving 36 native speakers of English, and one with 59 native speakers of Greek. In both experiments, participants were asked to determine the acceptability of the following sentence conditions: a) single and

double embedding, b) islands with that-, relative clauses, and whether- complement clauses, and c) resumption (or gap in the place of the R pronoun).

The predictions in Alexopoulou and Keller (2002) regarding resumption were that: 1) in no island conditions, gaps should be equally acceptable at any level of embeddings, and better than resumption at any level of embedding as well; 2) sentences with resumption are expected to save island violations and in turn they should be more accepted than gaps; and, 3) in the strong island condition, resumption cannot save the island violation, and thus both gaps and resumptives should be unacceptable (the authors also predicted that gaps should be less acceptable in single and double embeddings as this is a strong island violation).

The experiments in Alexopoulou and Keller (2002) yielded the following results. In the English resumptives experiment, the authors found that gaps were significantly more accepted than the sentences with a resumptive condition. Similarly, in single or double embedding conditions, gaps were more accepted than resumptives; for gaps, however, acceptability decayed in sentences with double island embeddings. Furthermore, Alexopoulou and Keller (2002) found that there is not a significant difference between the acceptability of gaps and resumptives, even in sentences with a strong island condition. The results in the experiment that analysed the acceptability of resumptives in Greek show that Greek speakers seem more accepting of resumption. Although as in English, embedding (single and double) in Greek did not seem to improve acceptability of either gap or resumptive conditions; Greek speakers were more accepting of resumptive conditions in weak islands as well as in non-embedded sentence conditions.

The authors conclude that their studies demonstrate that resumptives do not have a "saving" effect in either Greek or English as they were equally unaccepted as gaps even when interacting with different types of island and embedding conditions. Nonetheless, they emphasize that "an important crosslinguistic difference is that in English, resumptives are generally significantly worse than gaps (in all levels of embedding)" (p. 14). Alexopoulou and Keller (2002) suggest that these results call for a "processing explanation" (p. 14) to further understand the role of resumptives.

Provided the results observed in the latter study, Alexopoulou and Keller (2007) (from now on referred to as A&K, 2007) replicated their 2002 study on resumptives to further explore the effects of resumption cross-linguistically. To do so, the authors employed a methodology that allowed them to obtain quantifiable distinctions between type of clause

embeddings (-that, -wh, -whether islands), and the number of embeddings in English, Greek, and German. The study was comprised of 4 experiments: experiment 1 involved 55 native speakers of English, experiment 2 included 59 Greek speakers, and 37 German speakers were considered for experiment 3. The first 3 experiments were intended to test the sentence conditions in each language using the Magnitude Estimation (ME) technique (cf. Bard et al., 1996; Schütze, 1996; Cowart, 1997); which according to A&K (2007) "was developed to determine to what extent subjects can reliably indicate proportional judgments corresponding to degrees of magnitude in perceived physical stimuli, […]" (p. 115). The fourth experiment was intended to measure the effects of triply embedded sentence conditions in Greek using the stimuli for experiment 2 in their study.

Overall, the results in A&K (2007) demonstrate that embeddings have an effect on the acceptability of sentences with a gap condition; the more embeddings, the less acceptability there is for gapped clauses. Moreover, the authors found that resumption interacts with embedding; as embedding increases, the resumptive condition tends to be significantly more accepted. However, the most relevant crosslinguistic finding observed in the A&K (2007) study for the present investigation was that, in English, resumption was never more acceptable than the gap condition, which is in line with the authors' results in their Alexopoulou and Keller (2002) study. A&K (2007) noticed a crosslinguistic difference: while gaps and resumptive conditions were equally accepted in German and Greek, gaps were always significantly more acceptable than resumptives in English under any of the conditions tested in the study.

The finding regarding the unacceptability of resumptives in English in the A&K (2007) study is crucially relevant for the present research owed to the processing explanation that the authors provide for this occurrence. For A&K (2007), the preference for gap conditions among native speakers of English can be explained with the Syntactic Prediction Locality Theory (SPLT) by Gibson (1998). Based on the linguistic integration cost principle in Gibson's theory, R pronouns are more costly because they represent a new discourse referent to process as well as they create more distance "between the head being integrated and the head it attaches to" (Alexopoulou & Keller, 2007, p. 140).

More specifically, the SPLT (Gibson, 1998) states that the following principles are involved in the processing of long-distance sentences: 1) a syntactic prediction memory cost, and 2) a linguistic integration cost. Under the first principle, there are two specifications: a) the prediction of the predicate linked to the main verb does not imply memory costs; and, b)

the more discourse referents that are added after the prediction of the main verb's predicate has been made, the more memory costs there are for the parser. The latter, according to A&K (2007) implies that "local resolutions are always less costly than nonlocal ones" (p. 139). The second principle, the linguistic integration cost, consists of two factors. The first factor has to do with the cost that the type of element that is integrated to the sentence implies for the parser; for example, "a new discourse referent is more costly than an old established referent such as a pronominal" (Alexopoulou & Keller, 2007, p. 140). The second factor is that the integration of elements to a sentence is sensitive to the distance caused between the referent that is integrated and the head that said referent is attached to.

A&K (2007) use the syntactic prediction memory cost and the linguistic integration cost in Gibson's (1998) SPLT to explain their findings in terms of resumption; their explanation is crucially useful to understand the processing implications of resumption in English for the research objectives of the current investigation. According to A&K (2007), R pronouns are less acceptable than gaps in English because the English grammatical parser predicts that a gap will occur in the Ā-bound long-distance dependency (in which R pronouns are likely to occur); the latter means that the gap does not represent memory costs for the English parser.

Resumptives in English, on the other hand, add more referents to be parsed to the sentence because they act as new discourse referents (cf. A&K, 2007). Based on the linguistic integration cost in Gibson' (1998) SPLT, R pronouns, acting as new discourse referents, require for the parser to "go back" and search for the head antecedent to which the resumptive is attached. Hence, the R pronoun represents a higher processing cost for the English grammatical parser, which has already predicted that a gap might take place later in the long-distance sentence structure, as interpreted by A&K (2007).

The findings in A&K (2007) are further attested in their 2013 study, in which the authors explore the effects of resumption and (d)iscourse-linking in the acceptability of that-clauses and whether-islands in Greek and English. Departing from the notion that resumption did not seem to have a "saving" effect in long-distance sentences in their previous study, Alexopoulou and Keller (2013) considered 29 native speakers of Greek and 25 native speakers of English to test their acceptability of the sentence conditions mentioned above using the ME technique (cf. Bard et al., 1996; Schütze, 1996; Cowart, 1997). The authors based their study and predictions on Anagnostopoulous's (1994) referential hierarchy, in which it is proposed that the acceptability of clitic left dislocation (CLLD) and wh-phrases increases if the sentence conditions are given in the following order: "overt

partitive wh-phrases (which of your books) < which-phrases (which book) < what-phrases (what book) < bare wh-phrases (who, what)" (Alexopoulou & Keller, 2013, p. 315). Alexopoulou and Keller (2013) based on the referential hierarchy to expand the effects of resumptives given that studies as the one by Frazier and Clinton(2002) "provide evidence from judgment experiments showing that whether-islands with resumptive pronouns […] receive higher acceptability scores when the wh-phrase is d-linked […] than when it is not […]" (p. 315).

Alexopoulou and Keller (2013) found that 1) d-linking did seem to improve the acceptability of whether-islands in both English and Greek; however, it did seem to have an increasing acceptability effect when it interacted with resumption in Greek; 2) resumption did not increase the acceptability of whether-islands (but at least in Greek, embedded resumption did seem to reverse the unacceptability effect of these types of islands); and 3) d-linking does not seem to reduce the unacceptability of whether-islands containing resumption (at best, d-linking seemed to make whether-islands with resumption as acceptable as whether-islands with gap conditions). The authors find a confirmation for the referential hierarchy (Anagnostopoulous, 1994), but observed that "what + X" phrases contrast with "what (only)" phrases, and that "who-" phrases are more accepted than any other wh- phrases.

Nevertheless, the most important result and interpretation in this study is that d-linking has a stronger "saving" or rescuing effect than resumption. Alexopoulou and Keller (2013) explain that the "amelioration of integration costs, which is linked to d-linking/animacy, is more effective than (partial) cancelling of locality/distance-based costs, which is linked to resumption" (p. 340). The latter echoes the processing explanation for resumptives provided in A&K (2007) based on Gibson's (1989) SLPT. In other words, d-linking aids the parser with processing costs as it lessens the difficulty of carrying the referent until the end of the structure (thus, helping with the memory costs of the long-distance structure); while resumptives add a syntactic load (more in English than in Greek) in that they make the parser to search where the antecedent of such referent is located in the sentence (hence, creating more linguistic integration costs).

Overall, A&K's (2002, 2007, 2013) findings differ from Ariel (1999) and Hawkins (2004) who have implied that R pronouns have a saving effect (at least in Hebrew) and are, therefore, conventionally more accepted than structures with gaps in the position of an R pronoun. Although A&K (2007, 2013) agree that R pronouns are more than saving grammatical devices

to island violations, they also state that resumption is not a condition that necessarily lessens the difficulty of processing structures that have a long-distance condition and/or that create a filler-gap dependency as explained by the (memory and linguistic integration) cost principles in the Syntactic Prediction Locality Theory by Gibson (1998); at least, not in English.

Nevertheless, for late L2 learners, R pronouns might aid with the processing difficulty of long-distance structures. As stated by Leal-Mendez and Slabakova (2012), the grammatical strategy for resumption in Spanish is the same as in Greek. Spanish, as Greek, has a clitic left dislocation (CLLD), which interacts with resumption; making resumption slightly more acceptable and/or expected for the parser to occur. In English, on the other hand, there are no CLLD type of structures; this causes for the parser to create a filler-gap dependency strategy, in which the gap is predicted to occur from the moment that the head antecedent to which the gap is attached appears in the sentence. Therefore, R pronouns might help Spanish-speaking L2Aers who 1) might have a lower WMC ,and 2) are in the process of developing (or have not developed) the grammatical strategy of predicting gap conditions in English; and thus, these learners might need more elements to cope with the processing load imposed by long-distance structures.

With this in mind, the latter serves as rationale for one of the major objectives of the present investigation: exploring the relation between the IDs in WMC and the acceptability of R pronouns amid L2Aers. As explained in the previous subsection, from a genSLA point of view, the learnability task for Spanish-speakers, when it comes to R pronouns, is to comprehend that English is a non-null subject language; and hence, lacks the CLLD grammatical property. The latter implies for the L2Aers to abandon the Spanish resumptive strategy to correctly distinguish long-distance structures in English with or without resumption. Said learnability task poses a grammatical challenge for the L2Aer as it is characterized by linguistic properties (CLLD, null-subject v. non null-subject, gap realization and optionality, wh-movement) that are classified as uninterpretable features (cf. Chomsky, 1995).

As I discussed, some researchers claim that the resetting of the resumption parameter will not be possible for L2Aers given that L2 uninterpretable features are not accessible for these types of learners given maturational constraints (cf. The Interpretability Hypothesis by Tsimpli & Mastrapavlou, 2007). Opposite to this view, Leal-Mendez and Slabakova (2012) found evidence that demonstrates that L2Aers can accurately distinguish sentences with R pronouns in English; in spite of the difference in strategy for resumption in their L1 (Spanish).

According to the authors the availability of these types of properties in late learners' L2 grammars depends on their individual L1 grammatical preferences.

Nevertheless, given the evidence provided by Alexopoulou and Keller (2002) and their processing take on resumption (Alexopoulou & Keller, 2007, 2013) (based on Gibson's 1998 SPLT), I propose that the examination of resumption should be made in terms of the effects of general cognitive mechanisms such as WM, and on the L2Aer's individual differences in WMC. Mostly, because the findings and interpretations provided from a genSLA perspective are not sufficient to explain the variability found in L2Aers' grammars.

Chapter 3 Methodology

3.1 Introduction

As discussed in chapter 2, the need to explore the effects of working memory capacity (WMC) on the development of L2 oral fluency and acceptability of resumptive (R) pronouns among late Spanish speaking learners of English is to determine the role of this cognitive mechanism in the variability of attainment observed in late second language learners. With this in mind, the purpose of this chapter is to present the methodological design employed to study the role that WMC plays in the attainment of these two complex aspects of acquisition amid L2Aers.

First, in order to establish if there is a correlation between WMC and these two aforementioned aspects of L2 attainment, I have focused on quantitatively studying L2 oral fluency as an aspect that derives from the development of L2 automaticity, and on R pronoun acceptability since it is a property that 1) varies at the morphosyntactic interface in Spanish (L1) and English (L2) (cf. Leal-Mendez & Slabakova, 2012), and 2) might alleviate the parsing challenge imposed by Ā-bound long-distance dependencies (cf. Alexopoulou & Keller, 2007, 2013).

3.2 Rationale and research objectives and questions of the present study

The rationale behind the present research is that late second language learners (L2Aers) show various levels of attainment (Schmidt, 2011) and this is a well-known phenomenon in the research and literature of second language acquisition (SLA). Given this, much psycholinguistic research in SLA has concentrated on studying individual differences (IDs) among L2Aers in order to determine if said IDs, mainly in cognition, determine the degree of L2 attainment on these particular group of learners. Among these IDs, SLA studies with a cognitive approach have been concerned with the effects of working memory capacity (WMC) on the learners' L2 developmental process.

The reason that working memory (WM) is considered an influential ID on late learner's second language attainment is that it is an underlying mechanism responsible for temporarily storing information, while it underpins higher-order skills (Baddeley, 1983, 2000, 2003, 2007; Baddeley & Hitch, 1974). Based on the latter, major research studies have proven that WMC

is connected with skills that require deep linguistic knowledge such as reading (cf. Daneman & Carpenter, 1980), learning new vocabulary (e.g. Baddeley et al., 1988; Vallar & Baddeley, 1988), and comprehension of grammatically complex sentences (cf. Just & Carpenter, 1992). Generally speaking, this empirical research has provided substantial evidence to believe that WM is a necessary mechanism to perform linguistic tasks, and that not all individuals process information through WM in the same manner (cf. Cowan, 2005; Miyake & Freedman, 1998; interalia); said dissimilarities in storage and processing have been acknowledged as individual differences of capacity in working memory.

Therefore, since late learners might have less neuronal plasticity to rely on (cf. DeKeyser, 2010; Lenneberg, 1967; Penfield & Roberts, 1959; inter alia) or transfer from their L1 to their L2 during the language learning process (e.g. MacWhinney, 1997; Schmid, 2011; among others), there is a high possibility that they depend on their WMC to cope with the attainment of an L2. However, as it has been mentioned, since such capacity in working memory varies from individual to individual, the level of attainment among L2Aers might respond to said dissimilarities in WMC.

The previous remark is not strange to the cognitive field in SLA. As reviewed in Chapter 2, specifically in 2.2.4, many theorists and researchers have made the claim that working memory is necessary to acquire a second language; nevertheless, most of this research work needs clarification, more evidence, and a more unified conceptualization on how WM is addressed and what its role is in SLA among L2Aers. In lieu of this, the present work seeks to shed light on the role and influence that differences in WMC has on the level of attainment that late learners reach in their L2.

In order to do this, I take into consideration two central aspects of SLA emergentist and generative research and theory. As discussed in 2.1 in the previous chapter, it can be found that research evidence from both of these frameworks establishes that late learners are able to reach high levels of L2 attainment (in emergentism, e.g. Ellis, 1998; Ellis & Schmidt, 1998; MacWhinney, 2001; inter alia reviewed in 2.1.1) (in generative SLA, e.g. Rothman & Slabakova, 2017; White, 2003, 2007; White & Juffs, 1998; among others reviewed in 2.1.2).

On the one hand, proponents of emergentist research attest that L2Aers attain high levels of proficiency by developing what has been referred to as L2 automaticity (see 2.1.1.4 for a discussion on this concept in emergentist research). On the other hand, generative theorists have demonstrated that late L2 learners' grammars can be as complete and

accurate as that of L1 speakers (cf. Prévost & White, 2000; Goad & White, 2004; Rothman & Slabakova, 2018; interalia). Nonetheless, it can be observed that the development of L2 automaticity (cf. Ellis & Sinclair, 1996; Hulstijn, 1990; McLaughlin & Heredia, 1996) as well as the L2 grammatical competence vary among L2Aers.

Considering the latter, this investigation aims at focusing on the variability in L2 attainment observed amid L2Aers by examining aspects of L2 production (L2 oral fluency) and of the L2 grammatical competence (R pronouns) to explore up to what extent these aspects are related to IDs in WMC. For the purpose of testing the influence of WMC on the development of automaticity, I target L2 oral fluency as an aspect that derives from the development of automatic processing; to explore WMC effects on the L2 grammar, I have chosen to study the acceptability of resumptive (R) pronouns.

With this in mind, this study aims at answering the following general question: What is the role that working memory capacity plays in the development of oral fluency (as an automatic process) and the acceptability of object resumptive pronouns (as complex grammatical features) among late L2 learners? In order to respond to this question, the following specific questions will be addressed in the present research:

1) **Does working memory capacity have an effect on the oral fluency of late L2 learners?**
2) **Does working memory capacity have an effect on the acceptability of object resumptives of late L2 learners?**

In order to answer these questions, a quantitative research methodology has been adopted, and two studies were conducted. The following sections should serve to describe the tasks that were designed to obtain the data that could lead to respond to the research questions of the present research.

3.3 Participants

Originally, a total of 95 participants were recruited; however, only those participants that completed all the tasks comprising the research studies were selected and were distributed in three groups. After carefully verifying that participants have completed all of the tasks, a total of 72 individuals were considered for this research. One group of participants is composed of 27 advanced learners of English; another group is composed of 22 intermediate learners of English. The participants in both of these groups are native

speakers of Spanish and were tested with an Oral Proficiency Interview (OPI) by the American Council on the Teaching of Foreign Languages (ACTFL) to assert their level of proficiency in English and also to be classified on one of the aforementioned groups. The third group is composed of 24 native speakers of English.

The reason to test the proficiency level of the native speakers of Spanish is to establish how individual differences (IDs) in WMC operates at different levels on adult L2 learners. As it has been previously discussed, studies on WMC have been limited in terms of not including groups of participants with different levels of proficiency (e.g. Mizera, 2006; Dussias & Piñar, 2010). The importance of considering level of proficiency when measuring the effects of WMC on L2 oral fluency and acceptability of sentences with or without resumption is to have more evidence as to when in the L2 learning process, L2Aers rely more on WMC. With this in mind, the Spanish-speaking learners of English in this research were asked to take the OPI applied by the ACTFL.

The OPI test was applied in the following manner. First, a total of 59 individual OPI tests were purchased through the Language Testing International (LTI) website (https://www.languagetesting.com/), which is the only authorized company to apply language proficiency tests validated by the ACTFL; each license was priced at USD$75.00 (United States dollars). Some of the money granted as part of my scholarship was used to pay for the cost of the OPI tests for speaking proficiency in the English language. The participants were not charged to take the test; instead, they were given the results of the OPI and the certificate that comes with it, validated by the ACTFL, as a compensation for their participation in this investigation.

In order to take the test, participants were invited to go into a computer lab were each one of them had access to an individual desktop computer with internet access and headphones; each computer equipment was tested prior to the application of the OPI tests to make sure that it met the required hardware and software criteria specified by the LTI providers to avoid possible interruptions and/or technical difficulties. The participants took the OPIs in two groups in the computer labs. A group of students enrolled in the English as a Foreign Language program at the Universidad Autónoma de Ciudad Juárez (UACJ) took the test first; the next group was comprised of university students who were about to graduate or have graduated from the B.A. in English Language Teaching at the UACJ. Both groups were personally supervised by me and I was assisted by a computer lab technician who could provide me with guidance and support in case there was a technical issue as the participants

took the OPI tests simultaneously. There were no reported incidences during the application of the tests.

The OPI consists of four stages: a) the first stage is a "background survey", in which test takers are asked about general topics and about some personal preferences regarding work, school, hobbies, etc.; b) the second stage is a "self-assessment", in which the test takers indicate the level of proficiency that they feel "most accurately describes their language ability"; c) the third stage is the oral interview, in which an "avatar figure" asks questions to the test taker (the interview is preceded by a couple of trial sessions to assess and guarantee the quality of the audio, of the voice recording, and of the clarity of the questions and instructions of the OPI test); finally, d) the test taker receives confirmation of the recording of their responses, which are saved online and secured by the LTI company.

Once the test takers completed the OPI tests, their responses were evaluated and assessed by two external professional examiners assigned by the ACTFL. The examiners evaluate and grade the OPIs based on the ACTFL guidelines for language proficiency. According to their website the trained raters "listen to the sample and select the best match between the sample and the assessment criteria of the rating scale" (https://www.languagetesting.com/pub/media/wysiwyg/manuals/actfl-fam-manual-opic.pdf).

A total of 59 speakers of Spanish (potential participants) took the OPI test. The participants were notified that should their OPI test results meet the proficiency criteria, they would be invited to continue with the rest of the tasks in the research. The results of the OPIs were e-mailed to me (including the official certificates validating the results by ACTFL). The results are described in the next table.

Table 3-1 Level assigned by the ACTFL in OPIs classified by number of Spanish speaking participants

Number of participants	Proficiency level assigned by ACTFL on OPIs
Superior	12
Advanced high	18
Advanced Mid	3
Advanced Low	1

Intermediate High	7
Intermediate Mid	15
Intermediate Low	3

To form the group of advanced L2 learners, the participants who obtained either a "superior" or an "advanced high" were selected; resulting in a group of 30 participants. The group of intermediate L2 learners was comprised of those participants that were assigned any of the intermediate levels (intermediate high, mid, and low); thus, forming a group of 25 participants. This criterion on selection was taken to make sure that those participants in the advanced group were not closed in English proficiency to those in the intermediate group. Therefore, those participants who were assigned an "advanced mid" or "advanced low" level were not selected to continue in the investigation (*N* = 4). Consequently, the two groups of L2Aers were separated by at least two sublevels of English proficiency in accordance to the ACTFL guidelines; the latter serves to secure that the L2 linguistic knowledge and performance between the groups does not overlap, approximate, or assimilate.

Most of the participants in the group of advanced English learners were either about to finish a B.A. program in English Language Teaching or already had a degree in this or a related area at time of study. In their majority, the participants in the advanced group were teaching EFL courses to adult learners at college level. The participants in the group of intermediate English learners was mostly comprised by students who were taking an intermediate course of EFL at the Language Centre of the Universidad Autónoma de Ciudad Juárez (Autonomous University of Juarez) in Mexico.

In the case of the EFL students, one was assigned a "superior" level of proficiency; this participant was then reassigned to the advanced group. Also, among the EFL students, one individual had to be removed from the research project as they obtained an "advanced mid" level. In the group of ELT graduates and EFL teachers, two obtained an "advanced mid" level, and one more obtained a "high low" level of proficiency. These three individuals were then asked not to continue in the research.

Table 3-2 Participants' Ages

Group	**Number of participants**	**Minimum Age of Participants**	**Maximum Age of Participants**	**Mean Age of Participants**
Intermediate L2Aers	22	20	35	23.27
Advanced L2Aers	27	20	31	24.00
Native Speakers of English	24	20	35	29.21

The third group of participants is comprised of 24 native speakers of English whose ages range from 20 to 35 as indicated in Table 3-2; 35 native speakers of English had agreed to participate, but only 24 completed all of the tasks involved in the two studies here developed. Participants in both the Advanced and Intermediate groups of L2 learners reported to have started their acquisition process after puberty. The latter was an important consideration given that an important number of theorists and researchers in the SLA field agree that there is an evident variability in the levels of L2 attainment among individuals who start their acquisition process after puberty and/or during adulthood (cf. Schmid, 2011; Johnson & Newport, 1989; DeKeyser, 2010; Penfield & Roberts, 1959; MacWhinney, 1997; interalia).

Lastly, the participants in the group of native speakers of English were recruited by invitation via e-mail, and social networks. All the participants had either a college degree or were about to obtain a college degree at the time of this study. This was an important requirement to recruit the participants as it was necessary for them to have a solid education background to cope with the level of complexity of the sentences included in the grammaticality judgment task for Study 2.

Also, age was an important requirement as the purpose of the present work is to shed light on the reliance of WMC in the L2 developmental amid adult learners. Table 3-2 shows the number of participants, the range of ages (by indicating the minimum and maximum age), and the mean age in each group. As it can be seen on this table, after having 25 participants in the intermediate group and 30 in the advanced, only 22 participants were considered as a part of the intermediate group, and 27 for the advanced group at the time the data was stored, organized and analysed.

Participants were recruited with a written invitation to this study, and their participation was voluntary. In the case of the participants in the advanced and intermediate groups, their participation was rewarded, as aforementioned, by providing them an official certificate of their results in the OPI test provided by the ACTFL. Participants in the group of native speakers did not receive any kind of remuneration. Furthermore, participants were given a format containing all the necessary information concerning their participation on this study. Only participants who signed a consent form to be a part of this study were considered.

Participation was kept anonymous (participants were asked to use nicknames for these tests), and their data was stored in data bases in a single personal computer secured with a password. Participants were reminded that they could withdraw their participation from this study at any time. The formats, type of recruitment and tasks employed in this study were submitted for approval to the University of Southampton's Ethics and Research Governance Online (ERGO) committee on July 22, 2016. This research project was approved by ERGO on October 6, 2016 under the ID No. 23509.

3.4 Study 1

3.4.1 Tasks and materials for Study 2

3.4.1.1 Working memory capacity: materials and tasks

In order to measure working memory capacity, two span tasks were applied. These two tests consisted of a reading span task (RST), and a listening span task (LST) (Daneman & Carpenter, 1980). The reason to choose these specific span tasks is because they measure the functions of WM in the Baddeley (2007) model, including the control executive (cf. Wen, 2016). In addition, the software allows to obtain more reliable measures of WMC since the task is executed automatically rather than manually (cf. Juffs, 2011; Juffs & Harrington, 2014).

Furthermore, the design of these span tasks is based on Daneman and Carpenter's (1980) design to measure WMC. However, the span tasks used in this experiment are adaptations to automated versions of the RST and LST (Conway et al., 2005; Redick et al., 2012). The software to apply these tests was obtained by purchasing two Inquisit Lab licenses (one for each task) at the Millisecond Software, LLC online company. This software makes it possible to administer the tests on a computer (PC or Mac, desktop or laptop), or online. The participants in the two experimental groups of advanced and intermediate Spanish-speaking English learners took these tests on a computer lab. The participants in the control group of native speakers of English took these tests online.

The RST consists of 15 trials. These fifteen trials are composed of three repetitions of five sets; the order of sets is randomly determined. Afterwards, participants are asked to recall sequences of letters (from a set of 3 to 7 letters); each letter is preceded by either a sentence that makes sense or one that does not. Letter recall is done by picking out letters from a provided letter matrix. Prior to the 15 trials that comprise the actual test, participants have a practice session in which they recall sequences of letters in sets of 2 or 3 in an

ascending order (there are 4 trials); after the practice session, participants engage in a practice of semantic evaluations (which is made up of 15 trials). Finally, this test offers a combined practice of recalling sequence of letters (sets of only 2 letters) and semantic evaluation of sentences. The user manual script provided by the Milliseconds Company, together with the web address were this test can be taken are provided in 0.

The LST also consists of 15 trials as the RST. However, for the LST, participants listen to sequences of letters (that range from 3 to 7), which as in the RST, need to be recalled at the end. Each letter in the sequence is preceded by an auditory semantic categorization test, which consists of logical and non-logical sentences. Letter recall is tested by asking participants to select letters from a provided letter matrix. Participants have a trial session that follows the same logistics of the RST, with an auditory stimuli variation. The user script for this test provided by the Millisecond company, together with the web address were this test can be taken are detailed in Appendix A.

Especial attention should be paid to the fact that the tasks selected to measure WMC are in English, the L2 of the participants in the groups of L2Aers. However, the reason for selecting these specific automated span tasks is because of their high validity, their efficiency and overall proven reliability. As described in subsection 2.2.3, the validity of span tasks such as the RST and LST chosen for the present research is that they use stimuli that triggers the usage of storage and central executive functions of WM (cf. Conway et al., 2005); in the case of these tasks, participants have to store the letters presented to them for a later recall, while they make use of higher-order skills when deciding if the sentences presented to them are true or false.

Furthermore, the tasks can be easily and simultaneously applied to large groups of participants given that they can be individually taken on computer devices with internet access; this allows for the stimulus to appear on the screen in a systematic and timed manner, while the participants' responses are recorded, automatically stored, analysed and calculated (cf. Redick et al., 2012; Unsworth et al., 2005). The latter reduces the risk of human error, which was a factor that affected the efficiency and reliability of non-automated versions of span tasks (cf. Juffs, 2011; Juffs & Rodriguez, 2014) based on the Daneman and Carpenter (1980) model to measure WMC. Moreover, the LST and RST are shortened versions (Oswald et al., 2014), which allows for the participants to take less time to engage on the actual WM span task and have time for rehearsal trials.

The characteristics just mentioned were crucial when seeking for what WM span tasks to select. Unfortunately, none of the tasks available to measure WMC in Spanish had these characteristics. The majority of the tasks in Spanish, the L2Aers' native language, had to be applied manually, which might have affected the reliability and efficiency in the application of the tasks. However, in order to sort this problem, I revised the sentences that were going to be used for the true or false judgments to check for vocabulary or structural choices that could potentially represent a challenge for the Spanish speaking participants.

I did not find any vocabulary and/or structural items that could cause a problem for the participants; the sentences were adequate for the level of proficiency of the L2 adult participants in both groups. In any case, I was present when both groups of participants were taking the test in the computer lab and offered to translate any word, part of the sentence or complete sentence in either the RST and LST which they might not know or with which they were not familiar. Participants were also told in written form, in Spanish, that they could stop taking any of these WMC tests if they found that the vocabulary choices or the whole task were too complex; none of the Spanish speaking participants requested to be assisted and/or withdrew their participation when taking the LST and RST.

In addition, other studies that have explored the effects of WMC in L2 acquisition with adult learners have also opted for using only WMC span tasks in English and have not reported any shortcomings in their investigation because of this methodological choice (e.g. Dussias & Piñar, 2010; Miyake & Friedman, 1998). Other option that was considered to overcome the shortcoming of the language on the span tasks was to apply an operation span task (cf. Unsworth et al., 2005); this span task could not be selected for the present study given that it was very costly and out of my reach due to my economical limitations when conducting this investigation. Nonetheless, one of the main goals that I have for the near future is to look for technological support to develop WMC span tasks in Spanish that are as reliable, valid, and efficient as the ones employed in this research work.

3.4.1.2 **Oral fluency: materials and tasks**

In order to obtain measures of oral fluency, a speech generation task (SGT) (Daneman, 1991; Segalowitz, 2010) was designed. This task consisted of an elicitation of speech task, in which a couple of drawings were presented to participants. One drawing showed a family sharing a meal in a kitchen. The other drawing showed a park with adults and children engaged in multiple activities. For this task, participants were instructed to observe the two drawings and choose one.

These drawings can be observed in Appendix E and were retrieved from the website http://www.apic.es/imágenes. When participants were ready and had decided on one of the two drawings, they were asked to describe their chosen drawing for approximately thirty seconds. The reason for asking participants to describe these drawings within 30 seconds was to have individual audio samples that could be more efficiently measured in the PRAAT program used for the present study by each of the assistants that volunteered to help me with this endeavour. Given that there are 72 participants and I was restricted on both time and assistance for the measurement of each sample, I decided to have 30-second samples that do provide a significant window to the participant's oral fluency.

Moreover, participants could observe the drawing while they were describing it to avoid possible pauses on the sample owed to the participant's attempts to recall details of the drawing from memory. Participants were recorded while describing the pictures. The participants in both of the experimental groups went into a silent room in order to perform the task. Written and oral instructions on the task were provided to the participants. It is important to mention that the instructions for this task were provided both in Spanish and in English to the participants in the advanced and the intermediate groups of learners to minimize confusion for the participant and to make sure that the participant felt comfortable performing on this speech generation task.

In the case of the participants in the control group, instructions and files for this speech generation task were e-mailed to them individually in the form of two PDF document attachments. In addition, since there were going to be more steps for the participants in this group to undergo, they were presented with only one drawing to describe. The set of instructions for this task, as well as the drawings can be seen in Appendix I. Additionally, the file that was sent in an e-mail to these participants is included in said appendix. This appendix also includes the instructions that were given to the participants in the advanced and intermediate groups in I.1 and I.2.

While the task was being conducted, the participants were being recorded and timed as they were describing the chosen drawing. However, participants were not limited and were not asked to stop immediately after reaching the 30-second mark on the task; it was important for the purposes of this study to have participants perform in the most natural way possible. It is important to point out that ideally, the speech samples should have been longer as the formulae used to extract temporal measures of fluency are based on a time

frame of 1 minute. Nonetheless, the latter was not a difficulty as the formulae was adjusted to the 30 second time limit. Also, since the task used for the SGT was not demanding (since participants were not asked to engage in a conversation and/or description of abstract topics), the time allotted allowed for participants to provide sufficient details and extensive descriptions on the picture of their choice. Furthermore, the 30-second time frame has been used in similar speaking tasks on internationally validated English proficiency tests such as the Aptis Speaking test as reported by Tavakoli et al., (2020, p. 175).

The 30-second mark could imply some limitations if the present study involved locating where in their speech the L2 learners paused more (e.g. Tavakoli, 2011; de Jong, 2016; Kahng, 2014; Skehan & Foster, 2008), for example, as specified in Tavakoli et al. (2020). It could also represent a limitation if the purpose of the study included discourse analysis of the L2 speech samples as in Dewaele (1996) since the 30-second time limit could pose a challenge for the participant when organizing and structuring the description involved in the task. Given that the present study aims at analysing utterance fluency and its relation to the L2Aers' WMC from a quantitative perspective, the time allotted to complete the SGT did not represent any challenges for the analysis of the data; and, there were no reported samples that had to be significantly cut off and/or adjusted to fit the 30-second frame for data analysis. Moreover, the extraction of the temporal measures of utterance fluency was done using PRAAT, which permits to analyse the speech samples at a syllable level. Thus, no problems were encountered with lexical and/or phrasal completion on the samples taken into account for this study.

3.5 Study 2

3.5.1 Tasks and materials for Study 2

3.5.1.1 Resumptive (R) pronouns: tasks and materials

In order to measure the acceptability of resumptive (R) pronouns a Grammaticality Judgment Tasks (GJT) (Mackey & Gass, 2012; White, 2007) was designed. The type of sentences used for this task are 1) object resumptive pronoun, 2) a set of the latter sentences with the object resumptive pronoun removed from the clause, 3) grammatical relative pronoun clauses (used as fillers), and 4) ungrammatical relative pronoun clauses (used as fillers as well). Given this, the sentences used for the GJT are comprised of six sentences containing object resumptive pronoun type, and a set of six sentences, which are the same as the latter (object resumptive types) but have the resumptive pronoun feature removed

from the sentence. Also, these tasks contain a set of twelve grammatical sentences with a relative pronoun and a set of twelve ungrammatical sentences with relative pronoun, which are used as fillers. These fillers are also used as measures of grammatical accuracy for the L2 learner groups. Hence, the task consists of thirty sentences.

Two versions of this task were created; one GJT in English, and one in Spanish. The latter will allow to compare the degree of acceptance of resumption from the L1 to the L2 among Spanish-speaking participants. The purpose of evaluating the degree of acceptance of the resumptive and gap in resumptive position conditions in English and Spanish is to verify if the acceptability of these conditions in Spanish (L1) is transferred as a grammatical strategy to the L2 (following what the findings in the Leal-Mendez and Slabakova (2012) study indicate). Therefore, a Pearson co-variance test was conducted to verify if there is a significant correlation between the degree of acceptability to the sentence conditions in the GJT in Spanish and English among the two groups of L2Aers; the results of the Pearson correlational analysis are reported on Appendix J.

The sentences used for both of the GJT tests were extracted from corpora data bases (O'Keeffe et al., 2007). The sentences used for the English GJT were retrieved from the Corpus of Contemporary American English (COCA) (Davies, 2008). The sentences used for the Spanish GJT were retrieved from the Corpus de Referencia del Español Actual (CREA). To make sure that participants did not relate the R sentences with their non-resumptive pronoun counterpart sentence, a separation of ten sentences was considered when arranging the order of sentences in the test tasks.

Additionally, the selected sentences from the two corpora were modified in the following manner. All sentences were modified to have a similar number of syllables so that syllable length was not an issue that could influence the participants' acceptability rate. Moreover, the sentences that contained an object resumption condition were manipulated to create sentences with a gap in resumptive position; finally, in order to create the ungrammatical fillers, the sentences that were extracted from the corpora data bases were altered by removing an obligatory syntactic element from the relative clause (the relative pronoun). An account of the sentences utilized for the design of this test can be seen in Appendix B for the GJT in English, and in Appendix C for Spanish (sentences are arranged by type).

The purpose of the GJT was to determine degree of acceptance among participants (Mackey & Gass, 2012; White, 2007). Thus, for each sentence, participants were asked to indicate how natural the sentence was using a 4-point scale (Leal-Mendez & Slabakova, 2012). On this scale, "4" indicated that the sentence was "natural", and "1" indicated that the sentence was "not natural". The tasks were applied to participants using Google Forms. The latter allowed for participants to take this test online at their own pace. A sample of the GJT as presented to participants and the link to access it can be seen in Appendix G. In order to test the degree of acceptability of the sentences included in the GJT, a pilot was conducted with 32 native speakers of English and 27 native speakers of Spanish. The results of the statistical analyses of the data obtained from the pilot of the GJT are detailed in Appendix D.

The data extracted from both of the WM span tasks (RST and LST) has been used to determine the participants' working memory capacity. In addition, the data extracted from the GJT has been organized by sentence type to determine degree of acceptance of each sentence grammatical condition. Thus, in order to answer question 1, this experiment consists of correlating the data extracted from the WMC tasks and the GJT. Also, the data obtained from the WM span tasks and the GJT tests will be analysed per aspect (an individual analysis of working memory capacity tasks, and an individual analysis of grammaticality tasks). Each of these aspects is reported per group. Hence, the tasks will be further analysed on a per-group basis, so that a detailed account can be obtained on the aspects that comprise this study.

Chapter 4 Results of Study 1: Measuring the Effects of Working Memory Capacity on Oral Fluency

The present chapter is intended to provide the results obtained from the statistical analyses performed to the data of Study 1. These analyses were executed to obtain the necessary quantitative evidence that can lead to respond the question that motivates the first study of the current research:

Does working memory capacity have an effect on the oral fluency of late L2 learners?

In order to reach such statistical proof, the data from the tasks measuring WMC was organised and classified to extract specific quantities per group of participants that allowed to determine the capacity in working memory, and to spot any significant differences and/or similarities among the groups of participants. The same consideration was made for the data extracted from the task that measured oral fluency.

After analysing the data, the present chapter presents the results of the statistical model that was undertaken to explore up to what extent the WMC scores of the three groups of participants were related to their measures of oral fluency. In this manner, the significance or lack thereof found must serve as quantitative evidence to respond to the question of this study.

4.1 Data used for Study 1

4.1.1 Data used to measure working memory capacity

The scores gathered from the RST and LST are automatically calculated by the software employed to apply these tasks. Thus, the total scores are reported with the value “RSPAN” for the RST, and with the value “LSPAN” for the LST. To calculate these values the software uses an absolute span scoring method. The latter is the sum of all perfectly recalled sets. This measure was used given that it represents, in a precise manner, the number of times that participants were able to successfully use both the storage and processing functions of working memory. The extraction of these absolute span values did not pose a challenge as they are automatically calculated by the software of the automated span tasks that were applied. This was the score that was considered to report results on RST and LST. The highest score that participants could obtain on these tasks was 75.

The scores obtained by each participant on both of these tasks were saved and organized in an excel data base. Scores were organized by groups, and by task within the data base of each group. In order to obtain a single measure for WMC, the sum of the total scores that participants obtained in the RST and the LST were stored and used as a total WMC value. Also, to classify participants in groups of higher and lower WMC, the total value under the total of WMC was considered. Thus, averages for the WMC total value were estimated per group.

The reason to conflate the results obtained from the RST and the LST into a large sum instead of estimating the mean of both and using it as a value is because this could potentially reduce the number of times that the participants were able to accurately recall the sets in the order that they were presented and be able to make true or false judgments (on the sentences presented afterwards). A total sum of the results obtained on both tests represents in a more exact manner the extent to which the participant is able to employ the storage and central executive functions of WM.

The means between the scores obtained in the two tasks were not considered as a total measure of WMC given that they are central tendencies; and, as such, they have the disadvantage of not representing a meaningful value (cf. Manikandan, 2011). In the current research, using the sum of the total span values from both WMC span tasks is crucial to have a precise, rather than an approximate, number of what the participants' WMC is. Creating means for statistical analysis might be more useful and practical for studies in which there is a larger number of participants (Manikandan, 2011). Given that the number of participants in this research is not considerably large, a sum of the results from both WM span tasks has been opted as a measure of WMC to conduct the statistical analyses of the two studies in this investigation.

Hence, four measures were considered to analyse WMC: "rspan" value for RST, "lspan" value for the LST, the sum of "rspan" and "lspan" score values for WMC total span value. The "rspan" and the "lspan" scores represent the number of times a participant was able to recall the letters presented as well as to judge if the sentence that appears afterwards is true or false accurately. The data base with these four measures was created to be analysed with SPSS statistical software. These measures were saved as numerical, continuous variables. A One-way ANOVA analysis was run to obtain means and standard deviations, as well as to

explore differences between groups on each of the variable measures for working memory capacity.

4.1.2 Data used to measure oral fluency

In order to obtain measures for oral fluency, the elicited speech data obtained in the SGT was converted into audio files per participant, and then saved into folders created for each group. Each audio file was analysed using an automated software (De Jong et al., 2009) called PRAAT, which allows to measure audible speech in terms of time elapsed for phonemes, syllables, words, pauses, etc. Thus, the audios were measured to obtain three aspects of fluency: speed fluency, breakdown fluency, and repair fluency (cf. Skehan, 2003).

The latter are crucial to obtain a quantitative reliable measure of utterance fluency (De Jong et al., 2009; Segalowitz, 2010; also see 2.3.2 for a full discussion on utterance fluency and its role in SLA). To obtain measures for these three main aspects of oral fluency, it was necessary to establish the total time (total time of duration of the audio, 30 secs), phonation time (the percentage of time spent speaking as a percentage proportion of the time taken to produce the speech sample), number of syllables (total number of syllables produced in the sample), number of silent pauses (number of pauses above 0.2 seconds), total number of repairs and/or repetitions (number of restarts, fillers such as uhm, err, etc.), inter alia (Kormos, 2006; see Table 2-1 for a description on utterance fluency measures). The calculation of these measures allows to estimate participants' performance on oral fluency. Thus, the following measures were obtained per participants' audio:

a) **speed fluency**

Speech rate (syllables divided per total time)

Articulation rate (syllables divided by phonation time)

Mean syllable duration (phonation time divided by number of syllables)

Mean length of utterance (in syllables) (syllables divided by silent pauses +1)

Mean length of utterance (in seconds) (phonation time divided by silent pauses +1)

a) **breakdown fluency**

Number of pauses per minute (total time)

Number of pauses per minute (speaking time)

Mean pause duration (total length of silent pauses)

c**) repair fluency**

Total number of Repairs/Repetitions

Total length of Repairs/Repetitions

The values for each of these measures were saved in an excel data base, organized by participant. A data base was created for each group. This data base was used to analyse, and report results on measures for oral fluency using SPSS software. The measures that comprised the three aspects of oral fluency were used as numerical variables for SPSS analysis. A One-way ANOVA analysis was run to obtain means and standard deviations, as well as to explore differences between groups on each of the variable measures for oral fluency. Simple and multiple regression analyses were used to measure the effects of WMC on oral fluency; for the latter, it was necessary to take the measures that comprise each of the three dimensions of utterance fluency mentioned above. In the case, of the multiple regression analysis, the resulting measures of speed fluency obtained per participant in each group were conflated and considered as positive measures of oral fluency; the results of the measures for breakdown and repair fluency gathered for each participant were conflated as well and regarded as negative measures of oral fluency for the purposes of the present study and to have a more precise result of the effects of WMC on L2Aers' oral fluency.

4.2 Results of Study 1

4.2.1 Working memory capacity

As it was mentioned, this study comprehends the application of two tasks to measure working memory capacity. Hence, this section presents the results obtained in each of these WMC tasks, organized by group, and in a statistically descriptive manner. With this in mind, I present the results obtained in the Reading Span Task (RST) first, followed by the results gathered from the Listening Span Task (LST). I also present the resulting scores of the sum of these two tasks at the end, since said sum is used as the measure of Working Memory Capacity (WMC) in order to answer the research question of the present study.

As previously explained, all participants in the three groups took both the RST and the LST on a computer. Also, as previously mentioned, in these tasks, participants are granted a value of 1 every time that they manage to solve if the sentence that was presented to them is correct or incorrect, and if they can remember the input "letters" that they were exposed to (visually or auditorily). The maximum score that can be obtained in each of these tasks is 75; which means that there are 75 possible sets in which the participant can assess the semantic likelihood of a sentence, while remembering the order of letters that are presented afterwards.

The reason to also test the group of native speakers for WMC is to use their scores and results as a baseline to compare the results observed in the groups of intermediate and advanced learners; particularly, to explore if WMC has an effect on oral fluency and acceptability of object resumptive pronouns amid L1 speakers. Nevertheless, the findings of the ANOVA analysis for significant differences among the two groups of L2Aers and native speakers are reported to provide a more detailed description of the results obtained by the participants in the three groups that were considered for this investigation. The latter might serve to complement the findings observed in similar studies; especially those that relate WMC to language proficiency (cf. Temple, 2000; Prebianca, 2009; interalia), or to those that distinguish between verbal WM and visuo-spatial WM (cf. Miyake & Shah, 1996).

The main findings in this statistical analysis are that when comparing the scores between the intermediate and advanced learners, there is not an observed significant difference in terms of WMC (on the total measure of WMC that combines the scores obtained on both the RST and LST span tasks). Both groups of learners significantly differ from the group of native speakers on the total measure of WMC; the group of native speakers obtained considerably higher scores than both the intermediate and advanced groups in the global measure of WMC. The latter can be attested in Table 4-4.

However, when paying attention, the mean scores obtained by each group on the RST and LST tasks separately, it can be observed that the intermediate group obtained higher scores on both tasks. The differences in scores on these tasks (between the groups of intermediate and advanced learners) did not prove to be statistically significant; the only significant difference between the groups was found between the group of native speakers (who consistently showed higher scores on the WM tasks) and the two groups of L2Aers. What follows is a detailed description of the statistical findings on the RST, LST and total measures of WMC.

First, it can be observed that for the RST, participants in the group of intermediate English learners the mean score is *38.36 (N = 22, M = 38.36, SD = 26.72, CI = 27.37, 48.64)*. The mean score reached among the participants in the intermediate group of English learners is higher than that observed in the group of advanced English learners which is *23.19 (N = 24, M = 23.19, SD = 17.23, CI = 16.96, 29.70).* However, the native speakers of English reached the highest average score for this task, obtaining a *44.38 (N = 24, M = 44.38, SD = 18.01, CI = 37.17, 51.37).*

These numbers were estimated by running a descriptive statistical analysis on SPSS, which included a bootstrapping analysis of the data to report Confidence Intervals (CI) on the average scores of participants per group to have a better understanding of both effect sizes of the resulting numbers and the validity of the mean scores observed (Cumming, 2012; Kline, 2004; Larson-Hall, 2016). These results are summarized on **Table 4-1** below.

Table 4-1 Descriptive Statistics for the RST per Group of Participants

Group of Participants	Number of participants in group	Mean score	Standard Deviation	Bootstrap 95% Confidence Intervals	
				Lower	Upper
Intermediate English Learners	22	38.36	26.72	27.37	48.64
Advanced English Learners	27	23.19	17.23	16.96	29.70
Native Speakers of English	24	44.38	18.01	37.17	51.37

Furthermore, for the LST, participants in the intermediate group averaged a score of *19.55* (*N = 22, M = 19.55, SD = 19.67, CI = 12.09, 27.68*). On the other hand, participants in the advanced group obtained a mean score of *14.37* (*N = 27, M = 14.37, SD = 8.78, CI = 11.22, 17.70*), which is lower than the average score reached in the group of intermediate English learners. Once again, though, the participants in the group of native speakers have the highest mean score, which is *40.25* (*N = 24, M = 40.25, SD = 21.88, CI = 32.04, 48.37*). As with the data from the RST, the means, standard deviations, and confidence interval numbers for the LST were calculated by applying a descriptive statistical analysis on SPSS. These results are summarized on **Table 4-2** below.

Table 4-2 Descriptive Statistic Results for the LST per Group of Participants

Group of Participants	Mean score	Standard Deviation	Bootstrap 95% Confidence Intervals	
			Lower	Upper
Intermediate English Learners	19.55	19.67	12.09	27.68
Advanced English Learners	14.37	8.78	11.22	17.70
Native Speakers of English	40.25	21.88	32.04	48.37

*The descriptive values observed in the intermediate and advanced group have been highlighted in grey.

As it can be observed, the group of native speakers obtained scores that are seemingly at mid-level (*M = 44.38* for the RST and *M = 40.25* for the LST) with regards to the highest score (which is 75) that can be obtained in the WM span tasks. The latter can indicate that 1) the participants in the native speaker group have WM capacities that are more evenly distributed, and 2) the level of difficulty of the span tasks is considerably high regardless of the participants' L1. Hence, the span tasks meet the requirement of posing an incremental demand for the participants in terms of both storage and processing functions of WM (cf. Conway et al. 2005; Miyake, 2001).

Finally, in order to have a more integrative and global measure of working memory capacity (WMC), I conflated the scores of the RST and the LST by making a sum which is used in this study as a total WMC score. As explained before, the reason to use the sum of the RST and LST results instead of calculating the means between these two sums is to have a value that represents WMC in a precise and meaningful manner. Means, on the other hand, might have the disadvantage of representing a more general value to conduct statistical analysis that might be more applicable for studies in which the number of participants is considerably large (Manikandan, 2011).

The results of integrating the scores of the participants in these two tasks has also been measured using a descriptive statistical analysis in SPSS. Therefore, the results of this analysis show that the participants in the group of intermediate English learners have an average score of *57.91 (N = 22, M = 57.91, SD = 35.79, CI = 43.64, 71.82)* for the total WMC measure. As was predicted by the tendencies in the mean scores of the RST and LST, the participants in the groups of advanced English learners group had the lowest average score on this measure *(N = 27, M = 37.56, SD = 22.60, CI = 29.41, 46.32);* and as expected, the highest average score can be observed among the participants in the group of native speakers *(N = 24, M = 84.63, SD = 36.51, CI = 70.25, 98.87)*. These scores are summarized in Table 4-3 below.

Table 4-3 Descriptive Statistic Results for total WMC measures per Group of Participants

Group of Participants	**Number of participants in group**	**Mean score**	**Standard Deviation**	**Bootstrap 95% Confidence Intervals**	
				Lower	Upper
Intermediate English Learners	22	*57.91*	*35.79*	*43.64*	*71.82*

Advanced English Learners	27	*37.56*	*22.60*	*29.41*	*46.32*
Native Speakers of English	24	*84.63*	*36.51*	*70.25*	*98.87*

As it can be observed in the Confidence Interval (CI) values in Table 4-3, the advanced group scored significantly lower than the intermediate group. The latter can be explained in light of the findings observed in Finardi and Weissenheimer (2009), Prebianca (2009) and Prebianca et al. (2014) who report that L2Aers with lower proficiency levels tend to score higher on WMC span tasks. This will be further discussed in chapter 6 (Discussion of the results) in section 6.1.1. The scores obtained by the intermediate and the advanced groups in the LST and RST span tasks as well as on the WMC total value have been highlighted in grey in the tables above.

The mean scores obtained in these two tasks and on the WMC total measure have been analysed using a One-way ANOVA test in SPSS to determine if the differences in the average of mean scores is statistically significant among these groups of participants. This analysis shows that there is a significant difference between the mean scores of the three groups on the RST *[F (2,70) = 7.096, p < .002],* on the LST *[F (2,70) = 15.32, p < .0001],* and on the scores of WMC total measure *[F (2,70) = 13.930, p < .0001].*

By running a Post-Hoc Games-Howell analysis it can be observed that for the RST there is a significant difference between the mean scores of the advanced English learners and the native speakers of English of *-21.19* units *(CI = -33.17, -9.21, p < .0001);* there is no significant mean difference between the advanced and the intermediate groups of learners. The reason to use the Games-Howell analysis was to identify with more precision where the significant difference was among the groups. Moreover, the Games-Howell post hoc test is more useful in this specific statistical analysis than other post hoc tests, such as Tukey or Bonferroni, given that it does not assume that there is normality in mean distribution; and thus, it is more suitable to work with means that are not necessarily equally distributed (cf. Ruxton & Beauchamp, 2008), as it is the case of the data obtained in these tasks. In addition, there is no significant difference between the native speakers and the intermediate learners in the RST.

However, for the LST, there is a significant difference of *-20.705 (CI = -35.57, -5.84, p < .0001)* between the intermediate and the group of native speakers*;* similarly, there is a

significant mean difference of *-25.88 (CI = -37.57, -5.84, p < .0001)* among the advanced and the native speakers groups of participants. While there is a significant mean difference between the native speakers and both of the groups of learners, there is not an observed significant difference in mean scores between the intermediate and advanced learners.

Lastly, as it has been shown in the tendencies of the results of mean scores, the Post Hoc Games-Howell analysis shows that there is a significant difference in mean scores between the native speakers group and both groups of learners; the mean difference between the group of native speakers and the intermediate group is statistically different by *26.71* units *(CI = .84,52.59, p < .042),* and by *47.06* units with regards to the group of advanced learners *(CI = 26.01,68.13, p < .0001).* There is not a significant difference in mean values between the two groups of learners. A summary of the statistically difference in mean scores among the three groups of participants is provided on **Table 4-4** below.

Table 4-4 Post Hoc Games-Howell Significant Mean Differences among Groups of Participants on WMC tasks.

Group of participants	Task of WMC	Mean Difference	Significance	95% Confidence Interval	
				Lower Bound	Upper Bound
Advanced and Native Speakers	**RST**	-21.19	.0001	-33.17	-9.21
Intermediate and Native Speakers	**LST**	-20.70	.0001	-35.57	-5.84
Advanced and Native Speakers		-25.88	.0001	-37.67	-14.09
Native Speakers and Intermediate	**WMC Total Measure**	26.71	.042	.84	52.59
Native Speakers and Advanced		47.06	.0001	26.01	68.13

*Significant differences among the groups have been highlighted in grey.

In this manner, the results on the analysis of the data for the two tasks to measure WMC for the first study of the present research have been introduced from a statistically descriptive point of view. The numbers here presented are helpful to know the capacity in working memory of the English learners involved in this study, and to compare them with the

scores of the participants in the native speaker of English group. Of particular relevance was to observe the statistical differences in mean scores between the groups of learners in the different tasks and measures for WMC, as well as how they differ from the mean scores in the control group. These results will be analyzed in 6.1 later in this section.

4.2.2 Second language oral fluency

To continue with the development of Study 1, it is necessary to describe the results obtained in the descriptive analysis of the data obtained in the Speech Generation Task (SGT) used to measure the oral fluency of the participants in this research. As mentioned before, in 4.1.2, this task allowed us to obtain quantifiable measures in terms of time for speed fluency, breakdown fluency, and repair fluency which comprise the participant's utterance fluency. In this line of thought, the following is a descriptive statistical analysis report on the results observed in the SGT based on the estimations calculated in SPSS. Said results are presented by aspect of utterance of fluency and its distinct measures, organized by group and also containing the results of the 95% bootstrapping analysis for confidence intervals performed using SPSS. In the last part of this section, the results of the One-way ANOVA analysis for the SGT data is described to expose significance in difference of mean scores between the groups of participants.

First, the data from the SGT on speed fluency is composed of the following measures: speech rate, articulation rate, mean syllable duration, and mean length of utterance (seconds), and mean length of utterance (syllables); the measures were calculated per 30 seconds, which was the time allotted for the speech generation task. The descriptive statistical results for each of these measures is the following. For the group of intermediate learners of English, the mean score for speech rate is *M = 2.38* seconds (s) *(N = 22, SD = .52, CI = 2.16, 2.61);* for articulation rate, the mean observed is *M = 3.60s (SD = .73, CI = 3.31, 3.90).* Also, in the intermediate group, it can be observed that the average score for mean syllable duration is *M = .28 (SD = .06, CI = .26, .31),* a *M = 1.05s (SD = .53, CI = .85, 1.2)* for mean length of utterance in seconds, and a *M* = 3.65s *(SD = 1.71, CI = 2.98, 4.41)* for mean length of utterance in syllables.

In the group of advanced English learners, the mean scores obtained for speed fluency are higher than those in the intermediate group. This can be seen in the average score for speech rate, which is *M = 2.76s (N = 27, SD = .52, CI = 2.57, 2.95),* and on the *M = 3.96s (SD = .68, CI = 3.70, 4.23)* for articulation rate. Moreover, for the mean syllable duration measure

the average is *M = .25s (SD = .04, CI = .24, .27), M = 1.41s* for mean length of utterance in seconds *(SD = .59, CI = 1.18, 1.63),* and an *M = 5.39s (SD = 2.02, CI = 4.58,6.17)* for mean length utterance in syllables. Nevertheless, the mean scores in the group of the native speakers of English are higher in certain measures of speed fluency such as speed rate *(N = 24, M = 3.23, SD = .57, CI = 3.01, 3.46),* articulation rate *(M = 4.25, SD = .64, CI = 4.01, 4.51),* and mean length of utterance in syllables *(M = 6.86, SD = 4.23, CI = 5.43, 8.69);* whereas, certain measures such as mean syllable duration *(M = .24, SD = .03, CI = .22, .25),* and mean length of utterance in seconds *(M = 1.60, SD = .98, CI = 1.28, 2.04)* are fairly similar to those observed among the participants in the advanced group. The descriptive statistic results of speed fluency are summarized in **Table 4-5** below.

Table 4-5 Descriptive statistical results on Speed Fluency measures

Group of participants	Measure of Speed fluency	Mean	Standard Deviation	95% Confidence Intervals	
				Upper bound	Lower Bound
Intermediate English learners	**Speech Rate**	2.38	.52	2.16	2.61
	Articulation Rate	3.60	.73	3.31	3.90
	Mean syllable duration	.28	.06	.26	.31
	Mean length of utterance (seconds)	1.05	.53	.85	1.2
	Mean length utterance (syllables)	3.65	1.71	2.98	4.41
Advanced English learners	**Speech Rate**	2.76	.52	2.57	2.95
	Articulation Rate	3.96	.68	3.7	4.2
	Mean syllable duration	.25	.04	.24	.27
	Mean length of utterance (seconds)	1.41	.59	1.18	1.63
	Mean length utterance (syllables)	5.39	2.02	4.58	6.17
Native Speakers of English	**Speech Rate**	3.23	.57	3.01	3.46
	Articulation Rate	4.25	.64	4.01	4.5
	Mean syllable duration	.24	.03	.22	.25

	Mean length of utterance (seconds)	1.60	.98	1.28	2.04
	Mean length utterance (syllables)	6.86	4.23	5.43	8.69

In addition, the data subtracted from the Speech Generation Task was also used to calculate measures for breakdown fluency. The resulting means, standard deviations, and confidence intervals of these measures per group are described as follows. In the groups of intermediate English learners, it can be observed a *M = .72s* for number of pauses per 30 seconds "total time" *(N = 22, SD = .30, CI = .61, .86*), a *M = 1.14s (SD = .55, CI = .92, 1.38)* for the measure number of pauses per 30 seconds "speaking time/phonation time", and *M = .49s (SD = .19, CI = .41, .57)* for mean pause duration. In the group of advanced English learners, these means are lower with respect to the groups of intermediate English learners. This can be observed in the *M* = .58s *(N = 27, SD = .40, CI = .45, .75)* obtained in number of pauses per 30 seconds, in the *M = .86s (SD = .67, CI = .65, 1.14)* for number of pauses per 30 seconds "speaking time/phonation time", and in the *M = .62s* for mean pause duration *(SD = .29, CI = .51, .75)*.

Nonetheless, as it can be predicted by the nature of the type of measures and with the statistical tendencies observed in the measures for speed fluency, the participants in the groups of native speakers of English show lower means for number of pauses per 30 seconds "speaking time/phonation time" *(N = 24, M = .77, SD = .39, CI = .62, .92)*, and for mean pause duration *(M = .55, SD = .42, CI = .41, .74)*. However, the observed average for the number of pauses per 30 seconds "total time" *(M = 58, SD = .28, CI = .47, .68)* is the same as that obtained in the groups of advanced English learners. These results are summarized in the table below.

Table 4-6 Descriptive statistical results of Breakdown Fluency measures

Group of participants	**Measure of Breakdown fluency**	**Mean**	**Standard Deviation**	**95% Confidence Intervals**	
				Upper Bound	Lower Bound
Intermediate English learners	**No. of pauses per 30 seconds (total time)**	.72	.30	.61	.86
	No. of pauses per 30 seconds (speaking time/phonation time)	1.14	.55	.92	1.38

	Mean pause duration	.49	.19	.41	.57
Advanced English learners	**No. of pauses per 30 seconds (total time)**	.58	.40	.45	.75
	No. of pauses per 30 seconds (speaking time/phonation time)	.86	.67	.65	1.14
	Mean pause duration	.62	.29	.51	.75
Native Speakers of English	**No. of pauses per 30 seconds (total time)**	.58	.28	.47	.68
	No. of pauses per 30 seconds (speaking time/phonation time)	.77	.39	.62	.92
	Mean pause duration	.55	.42	.41	.74

Furthermore, the data of the SGT was also analysed for measuring the participant's repair fluency, which is the last dimension of fluency that is approached in this study to quantitively describe the participant's utterance fluency. Regarding this, two measures were taken from the SGT data to estimate repair fluency: total number of repairs and/or repetitions, and the total length of repairs and/or repetitions. Thus, the statistical descriptive analysis demonstrates that the average mean in the intermediate is the largest *(N = 22, M = 4.36, SD = 3.60, CI = 3.00, 6.05)*, compared the mean in the group of advanced learners *(N = 27, M = 2.30, SD = 2.78, CI = 1.29, 3.43)*, and to group of native speakers *(N = 24, M = .88, SD = 1.07, CI = .48, 1.33)* which is the shortest. Consequently, for total length of repairs/repetition the longest mean can be observed in the intermediate group *(M = 1.71, SD = 1.21, CI = 1.17, 2.24)*, with respect to the *M = .81s* in the group of advanced learners *(SD = .98, CI = .42, 1.19),* and the *M = .30s (SD = .36, CI = .15, .46)* in the group of native speakers, which is the shortest in terms of time. These results are provided in a summarized version the following table.

Table 4-7 Descriptive statistical results of Repair Fluency measures

Group of participants	**Measure of Repair Fluency**	**Mean**	**Standard Deviation**	**95% Confidence Intervals**	
				Upper Bound	Lower Bound
	Total number of repairs/repetitions	4.36	3.60	3.00	6.05

Intermediate English Learners	**Total length of repairs/repetitions**	1.71	1.21	1.17	2.24
Advanced English Learners	**Total number of repairs/repetitions**	2.30	2.78	1.29	3.43
	Total length of repairs/repetitions	.81	.98	.42	1.19
Native Speakers of English	**Total number of repairs/repetitions**	.88	1.07	.48	1.33
	Total length of repairs/repetitions	.30	.36	.15	.46

Nevertheless, part of this statistical report on oral fluency consists of confirming if the differences and similarities presented above are significant, not just at a descriptive level, but an inferential level. In this sense, I used a One-way ANOVA analysis to confirm if the groups of participants are statistically alike or distinct among each other. Moreover, I applied a Post-Hoc Games-Howell analysis using SPSS software to know what groups significantly differed, and on what measure of utterance fluency this difference resided.

Hence, it can be observed that there is a significant difference in all measures that comprise speed fluency between all three groups of participants. These variances are reported in Table 4-8, where it can be observed that there is a major difference among groups on mean syllable duration *[F (2,700 = 6.051, p < .004],* speech rate *[F (2,70) = 14.176, p < .0001],* mean length of utterance in syllables *[F (2,70) = 7.109, p < .002],* and articulation rate *[F (2,70)= 5.15, p < .008]*. In not as a statistically significant manner, the groups differ on mean length of utterance in seconds *[F (2,70) = 3.32, p < .042].*

Table 4-8 One-way ANOVA analysis of measures of Speed Fluency among the three groups of participants.

Measure of Speed Fluency	Degrees of Freedom	F-value	Significance
Mean syllable duration	2,70	6.051	.004
Speech Rate	2,70	14.176	.0001
Mean length of utterance (syllables)	2,70	7.109	.002
Articulation Rate	2,70	5.15	.008
Mean length of utterance (seconds)	2,70	3.32	.042

Taking a closer look at the statistical reported dissimilarities among groups, a post hoc Games-Howell analysis in SPSS shows that for speech rate there is a difference of *-.37s (CI = -.74, -.01, p < .040)* between the intermediate and the advanced groups of English learners, and of -.84s *(CI = -1.24, -.45, p < .0001)* between the group of intermediate English learners

and the group of native speakers of English. Also, in this measure, there is a significant difference of *-.46s (CI = -.84, -.09, p < .011)* in the advanced group with respect to the native speaker group. In a similar pattern, the intermediate group distances itself by *-1.73s (CI = -3.03, -.44, p < .006)* from the advanced group, and by *-3.20s (CI = -5.51, -.89, p < .005)* from the native speaker group in mean length of utterance in syllables. However, this post hoc analysis shows that the advanced group does not significantly vary from the group of native speakers on this particular measure.

Furthermore, the post hoc Games-Howell shows that the significant difference in mean syllable duration is between the intermediate and the native speaker groups by *.049* seconds *(CI = -.08, -.01, p < .010).* Interestingly, the differentiation between the three groups in terms of mean length of utterance in seconds is not significant after conducting the Post Hoc Games-Howell analysis. Finally, for articulation rate, the variance between groups is only significant between the intermediate group by *-.65s (CI = -1.14, -.15, p < .008)* with respect to the group of native speakers. These results can be observed in Table 4-9 below.

Table 4-9 Post hoc Games-Howell analysis in measures of Speed Fluency

Group of participants	Measure of Speed Fluency	Mean Difference	Significance	95% Confidence Interval	
				Lower Bound	Upper Bound
Intermediate and Native Speakers	**Mean syllable duration**	-.04	.010	-.088	-.01
Intermediate and Native Speakers	**Articulation Rate**	-.65	.008	-1.14	-.15
Intermediate and Advanced	**Mean Length of utterance (syllables)**	-1.73	.006	-3.03	-.44
Intermediate and Native Speakers		-3.20	.005	-5.51	-.89
Intermediate and Advanced	**Speech Rate**	-.37	.040	-.74	-.01
Intermediate and Native Speakers		-.84	.0001	-1.24	-.45
Advanced and Native Speakers		-.46	.011	-.84	-.09

On the other hand, there is not an observed significant difference between these three groups on the measures that comprise breakdown fluency (number of pauses per minute in total time, number of pauses per minute on speaking time/phonation time and mean pause

duration). Nevertheless, when analysing these results through the post hoc Games-Howell statistical filter, it is found that there is a significant difference of *.37s (CI = .01, .72, p < .037)* between the intermediate group and the group of native speakers for the measure of number of pauses per minute (speaking time/phonation time).

The latter is not the case, however, for the measures of repair fluency. There is a significant dissimilarity among groups on total number of repairs and/or repetitions *[F (2,70) = 9.82, p < .0001]* and on total length of repairs and/or repetitions *[F (2,70) = 13.61, p < .0001].* These significances are better perceived in the post hoc Games-Howell analysis, in which it can be observed that there is a mean difference of *3.48* number of repairs/repetitions between the intermediate and the native speakers groups *(CI = 2.13,5.27, p < .001)*, and a mean difference of *1.42 (CI = .30, 2.63, p < .049)* between the advanced and the native speakers groups in this same measure.

In addition, all three groups significantly differ in terms of total length of repairs/repetitions among each other; the group of intermediate learners varies with a mean difference of *.90s (CI = .24,1.55, p < .021)* with respect to the advanced group, and with a mean difference of *1.40s (CI = .92, 1.9, p < .0001)* in relation to the group of native speakers. Lastly, for this measure, the advanced group significantly differs from the group of native speakers by *.50s (CI = .14,.92, p < .048).* A summarized report of these results can be seen in Table 4-10 below.

Table 4-10 Post hoc Games-Howell analysis for measures of Repair Fluency.

Group of participants	Measure of Repair Fluency	Mean Difference	Significance	95% Bootstrap Confidence Intervals	
				Lower bound	Upper Bound
Intermediate and Native Speakers	Total no. of repairs/repetitions	3.48	.001	2.13	5.2
Advanced and Native Speakers		1.42	.049	.30	2.63
Intermediate and Advanced	Total length of repairs/repetitions	.90	.021	.24	1.55
Intermediate and Native Speakers		1.40	.0001	.92	1.97
Advanced and Native Speakers		.50	.048	.14	.92

With the latter, I conclude the section on the results obtained in the data drawn from the speech generation task to measure oral fluency on the groups of participants involved in the present research study. With this mind, there is a base to understand the tendencies,

distinctiveness and similarities between the groups of participants regarding oral fluency. These results are approached further in this section to analyse the role that they play in the current research and the objectives here pursued.

4.2.3 Working memory capacity effects on oral fluency

As specified earlier, the main purpose of Study 1 is to find the extent to which working memory capacity (WMC) measures correlate with oral proficiency measures on adult second language learners (L2Aers). Thus, so far, I have provided the results of the tasks that comprise this study individually. Nevertheless, in this section, I reveal the results of the statistical test that served to accomplish the aim of the current study. In order to obtain these results, I ran a multilinear regression data analysis using SPSS software.

Hence, as explained in 4.1.1, the data considered as the participant's WMC measure was the sum of the scores that they (the participants) obtained in the RST and LST; this measure will be referred to as WMC total to report the statistical results of the current study. Furthermore, the data considered to analyse oral fluency were the measures that constitute the participants' speed, breakdown, and repair fluency; these measures are described in 2.3.2 and illustrated in Table 2-1 An overview of Measures of Fluency (Kormos, 2006, p. 163).

In order to be able to apply the multilinear regression model, the measures encompassing speed fluency (speech rate, articulation rate, mean syllable duration, mean length of utterance (seconds), and mean length of utterance (syllables)) were conflated and regarded as a variable that represented positive measures of oral fluency. Furthermore, the measures of breakdown (no. of pauses per minute in total time, no. of pauses per minute (speaking time/phonation time), mean pause duration) and repair (total number of pauses, total length of repairs/repetitions) fluency were conflated as well and referred to as a variable that indicates negative measures of oral fluency.

The conflation of the temporal measures into two composite variables of utterance fluency is based on the observations made by Tavakoli (2016) who claims that "recent research findings suggest that some [temporal] measures may overlap with others" (p. 138). This overlapping indicates that the temporal measures that comprehend the major dimensions of utterance fluency (breakdown, repair and speed fluency) (Skehan, 2003; Tavakoli & Skehan, 2005) correlate with one another; and therefore, can be put together as reliable quantities that describe one of these three aspects of fluency. Moreover, Witton-

Davies (2010) makes the observation that the combination of complex measures of fluency can account for a composite measure of utterance fluency. For example, Witton-Davies (2010) highlights that "speech rate" and "mean length of run" have been used in combination in studies such as the one by Towell et al. (1996) and have proven to be "good indicators" to measure utterance fluency (p. 120).

Furthermore, Tavakoli (2020) emphasizes that in spite of the wide range of studies that utilize temporal measures to analyse fluency, "there has been little agreement about which measures can best characterize speakers' fluency , or whether certain measures are more relevant to describing fluency in different L2 tasks and contexts or at different stages of L2 development" (p. 172). With this in mind, I deemed it necessary to calculate various temporal measures, so that I could put together a more integral quantitative description of fluency. However, as mentioned before, these measures can be combined and serve as a reliable composite measure of fluency. Also, integrating these measures into a composite measure of oral fluency makes it possible to study L2 oral fluency and WMC with advanced inferential statistical models; such as the multilinear regression analysis used in this research. Therefore, what follows is the report of the significant regression equations that resulted of the analysis of the WMC total variable and the variables of oral fluency as they were observed per group of participants.

In the case of the intermediate group there was an observed significant regression *(F (2,18) = 5.098, p < .003)* with an R^2 *of .48* for the effects of WMC on positive and negative measures of oral fluency. This means that for the L2Aers in the intermediate group their positive measures of oral fluency will increase by *1%* for every *2.045%* obtained in WMC measures, while their negative measures of oral fluency will decrease by *1%* for every *-.717%* less in WMC measures. The *DW value* of this analysis is found to be within the permitted limits *(DW = 1.446)* which indicates that there is an absence of autocorrelation between the variables. Also, the Variance Factor and the Tolerance Values are within the permitted ranges *(VIF = 1.081)*; hence, demonstrating that there is not multicollinearity between the variable measures of oral fluency. Therefore, effects of WMC can explain the oral fluency performance of L2Aers in the intermediate group in a *69%* of the cases as indicated by this statistical analysis.

Nonetheless, there were no significant regressions found in any of the groups of advanced learners or the native speakers. In the case of the advanced group, the regression indicated that WMC does not explain the results in neither positive nor negative measures

of oral fluency among these participants *[F (2,23) = .611, p = .551]*, where the *DW (= .513)* are below the permitted limits, thus, indicating autocorrelation between the variables. Similarly, in the group of native speakers the regression showed that WMC cannot explain the results obtained by participants of this group in terms of positive or negative measures of oral fluency *[F (2,21) = 2.019, p = .158],* for which the *DW (= .505)* indicates autocorrelation among the variables. Hence, WMC effects cannot explain the oral fluency performance of the participants in the native speakers of English group.

Chapter 5 Results of Study 2

5.1 Introduction

The purpose of this chapter is to report the results obtained from the statistical analysis performed on the data extracted from the Grammatical Judgment Task (GJT), intended to test the acceptability of object R pronouns. Secondly, this chapter describes the results obtained from measuring the effects of Working Memory Capacity (WMC) on the degree of acceptability of object R pronouns used in the GJT.

5.1.1 Data used to measure the acceptability of R pronouns

The data obtained from GJT required to make a series of classifications. First, the ratings for each sentence were gathered and arranged per participant and per group in an excel data base. These ratings were coded into type of sentence as follows: object resumptive pronoun type (+RO), gap in object resumptive pronoun position (-RO), grammatical relative pronoun (GRL) (used as fillers), and ungrammatical relative pronoun (UGRL) (used as fillers). Afterwards, a data base was created for each sentence type. In each of these data bases the "1" (not natural) to "4" (natural) answers were conflated to two definite values as follows: "1" and "2" were conflated to "1" and taken as a rejection for the sentence type; "3" and "4" were conflated to "4" and taken as an acceptance of the sentence type. The conflated number answers were calculated and converted to two measures: a measure for degree of acceptance of the sentence type, and a measure for percentage of acceptance of the sentence type.

To calculate the measure for degree of acceptance the following considerations and calculations were made:

A) If participants had given an answer of "4" on four (or more) of the six sentences of the "+RO" or "-RO" sentence types, then they were classified with the value "accept"; this same classification was made for the "GRL" and "UGRL" filler sentence types, but there had to be a "4" answer on seven (or more) of the twelve sentences within these types.

B) If participants rated the majority (4+ out of 6, and 7+ out 12) of any of the type of sentences with a "1", they were classified as "reject" on the data base.

C) If participants did not have a majority of either "4" or "1" ratings per sentence type, they were classified as "no tendency".

To calculate the measure of percentage of acceptance, all "4" ratings were divided into the total of sentences in each type and reported in percentages per participant in each group. This calculation was made in order to have a more precise indication on the degree of acceptance that each participant had on each sentence type.

Given these two measures, a general data base was created containing the total values calculated for each measure in each sentence type data base. A general data base was created per group. This general data base was used to analyse sentence types per group using SPSS statistical software. The degree of acceptance measure was saved as a categorical variable, and the percentage of acceptance measure as a numerical, continuous variable. How these data were analysed is provided below. Furthermore, a pilot of the GJT was conducted to test the degree of acceptability among native speakers of English and Spanish. The results of the analyses of the data obtained from the pilot are included in Appendix D.

5.2 Results of Study 2

The first purpose of the present section is to show the results of the statistical analyses that were conducted to measure the participants' acceptability of object R pronouns. The second goal is to report on the results retrieved by the multiple regression analysis run on SPSS to measure the degree of correlation between the data of WMC measures and the data related to degree of acceptance of R pronouns among the participants in the three groups of this study.

5.2.1 Results of the statistical analysis on GJT measuring the acceptability of R pronouns

This section is organized as follows. First, I present a statistical description of the data measuring the acceptability of the +RO, -RO types of sentences, which involves a report on the means and standard deviations of the +RO and -RO sentences in the GJT organized by sentence type and by group of participants. Moreover, the results of a chi-square analysis of the data resulting from the English GJT is provided to determine if there are significant differences in the degree of acceptance of the sentence conditions +RO and -RO in English among the two groups of L2Aers and the group of native speakers of English. The results of the chi-square analysis will serve to determine how L2Aers react to these grammatical conditions with respect to English speakers; and, in such manner, said results will help to examine if the tendencies for acceptance or rejection regarding the sentence conditions can provide further information about the grammatical strategies that are being applied as a

group and to observe if there is any variability in the degree of acceptance among the two groups of L2Aers and the native speakers of English.

Moreover, a Pearson co-variance analysis was conducted on the numerical data obtained from the application of the GJT in English and Spanish to the intermediate and advanced groups of L2Aers. The aim of this analysis was to observe if there is any significant degree of acceptability for the object resumptive (+RO) and gap in object resumptive (-RO) conditions in English among the L2Aers in their L1 (Spanish) and their L2 (English). The results there found will shed light on the degree of L1 transfer of the grammatical strategies of these conditions to their L2. The results obtained in the Pearson co-variance analysis performed on SPSS are reported on Appendix J.

Finally, a multiple regression statistical analysis between the total measure of WMC and the percentage of acceptability of the +RO and -RO sentence conditions was conducted. The purpose of this analysis is to answer the second research question of the present investigation: does working memory capacity correlate with the acceptability of the object resumptive pronoun among L2Aers? The results of the multiple regression analysis will allow to understand if the acceptability of the +RO and -RO conditions can be explained by the individual differences in WMC among these learners. The latter will serve to verify if WMC has a significant mediating effect on the L2 grammatical behaviours of L2Aers; and thus, lead to understand why there is variability among them at the grammatical level.

It is important to mention that in order to test the degree of acceptability of the sentences included in the GJT, a pilot (using both the English and Spanish versions of the GJT) was conducted with separate groups of participants (meaning that these groups of participants were not involved with the rest of the tasks involved for study 1 and/or 2), who reported to be native speakers of English (*N = 32*) and Spanish (*N = 27).* The pilot has the purpose of verifying the degree of acceptability of the sentences among groups of native speakers; which will serve as bases of comparison for the grammatical behaviours observed among the groups of participants involved in the investigation with regards to the resumption (+RO) and gap in resumption position (-RO) conditions. The results of statistical analysis of the data extracted from the pilot is reported in Appendix D.

5.2.1.1 **Descriptive analysis of the sentences in the GJT**

To begin with this report, it is necessary to remember that the type of sentences included in this task, which have been previously detailed in 5.1.1, are:

Table 5-1 List of referents used to indicate the sentence types included in the GJT

Referent used for the sentence type	Description of the referent used to indicate sentence type
+RO	Sentences with an object resumptive pronoun.
-RO	Sentences with the object pronoun removed.
GRL	Sentences containing grammatical relative clauses functioning as fillers.
UGRL	Sentences containing ungrammatical relative clauses functioning as fillers.

Considering the latter, the present subsection is intended to provide the descriptive statistics of each of the sentences with a +RO and a -RO condition. The descriptive statistics of the filler sentences are provided in Appendix D. It is important to remember that the participants indicated the degree of acceptability using a 4-point Likert scale, in which "1" indicated "not natural" and "4" "natural". The values were conflated in the following manner: values of 1 and 2 were taken as "rejection" of the sentence; and the values 3 and 4 were taken as acceptance of the sentence condition. To this end, the means and standard deviations of the sentences were estimated using excel formulas. The latter were performed on the data bases that contain the conflated numbers that indicate degree of acceptance on each sentence. This calculation was made using the columns that resulted from organizing the numerical responses of the three groups of participants by sentence type, and by group.

Also, the responses of each participant to each sentence was accessible given that each sentence was categorised using the number that it was given in the GJT. This statistical analysis was only performed to the sentences belonging to the GJT in English since the purposes are to 1) have a reference of the reactions to each sentence by comparing the reactions of the groups of learners with regards to the native speakers as a control group, and 2) to assess the validity of each sentence as a token to measure grammatical knowledge of complex features in English.

To begin with, Table 5-2 shows the results obtained for each of the sentences comprising the +RO type. In the table, the clauses containing the R pronoun are highlighted in bold within the column containing the sentences.

Table 5-2 Descriptive statistics of the +RO sentences

Sentences with a Resumptive Pronoun	Descriptive Statistics	Advanced	Intermediate	Native Speakers
	Mean	**2.66**	**2.09**	**1.75**

1. Despite the attention given in recent days to allegations by an Okinawan woman **who an American airman raped her**, molestation of women and sexual harassment received little coverage	**Standard Deviation**	1.49	1.44	1.29
2. He talked a lot from the speech he gave the other day **that if you've not looked at it** is a good summary of the work the United States has been doing with respect to the Asian regional economy.	**Mean**	**2.44**	**3.04**	**1.87**
	Standard Deviation	1.49	1.39	1.36
3. The issue is not if gays are in the military. It is if they can be there without lying about it since it is a strict code of conduct **that if they violate it** would lead to dismissal from the service.	**Mean**	**3.33**	**3.18**	**2.5**
	Standard Deviation	1.24	1.33	1.5
4. I know the cast felt the same and our producers as it's the best and one of the greatest living playwrights, Jon Robin Baitz **who he wrote this piece** with such eloquence and compassion for all the different characters.	**Mean**	**2.66**	**2.90**	**1.37**
	Standard Deviation	1.24	1.44	.99
5. We were expecting this foreign man, Maroof Farooq, to show up after having a sexually explicit chat with a decoy **who he told him** that this young girl was a 12-year-old female.	**Mean**	**2.55**	**2.90**	**1.5**
	Standard Deviation	1.49	1.44	1.18
6. Medavoy was involved in some great films such as "One Flew Over the Cuckoo's Nest" during his tenure at United Artists; "Amadeus" and "Silence of the Lambs" at Orion Pictures, **that he co-founded it**.	**Mean**	**3.44**	**3.59**	**1.5**
	Standard Deviation	1.16	1.02	1.11

*The means have been emphasized in bold.

Also, the means and standard deviations of the sentences under the -RO category are reported in the table below. In Table 5-3, the asterisk (*) indicates where the R pronoun has been removed in the sentence.

Table 5-3 Descriptive statistics of the -RO sentences

Sentences with R pronoun removed in object position	**Descriptive Statistics**	**Advanced**	**Intermediate**	**Native Speakers**
7. Despite the attention given in recent days to allegations by an Okinawan woman **who an American airman raped ***, molestation of women and sexual harassment received little coverage.	**Mean**	2.66	2.5	3.37
	Standard Deviation	1.49	1.5	1.21
	Mean	2.44	2.77	2.0

8. He talked a lot from the speech he gave the other day **that if you've not looked at *** is a good summary of the work the United States has been doing with respect to the Asian regional economy.	**Standard Deviation**	1.49	1.47	1.41
9. The issue is not if gays are in the military. It is if they can be there without lying about it since it is a strict code of conduct **that if they violate *** would lead to dismissal from the service.	**Mean**	3.66	3.45	2.62
	Standard Deviation	.94	1.15	1.49
10. I know the cast felt the same and our producers as it's the best and one of the greatest living playwrights, Jon Robin Baitz **who he wrote *** with such eloquence and compassion for all the different characters.	**Mean**	2.44	2.5	1.37
	Standard Deviation	1.49	1.5	.99
11. We were expecting this foreign man, Maroof Farooq, to show up after having a sexually explicit chat with a decoy **who he told *** that this young girl was a 12-year-old female.	**Mean**	3.11	2.63	2.25
	Standard Deviation	1.36	1.49	1.47
12. Medavoy was involved in some great films such as "One Flew Over the Cuckoo's Nest" during his tenure at United Artists; "Amadeus" and "Silence of the Lambs" at Orion Pictures, **that he co-founded ***.	**Mean**	3.66	3.72	2.62
	Standard Deviation	.94	.86	1.49

*The means have been emphasized in bold.

Finally, as part of the analysis of the sentence conditions +RO and -RO, I conducted a descriptive analysis of the means obtained per group of participants. The results of the analysis are reported on the tables below. It can be observed that the overall acceptability for the +RO condition is very similar among the two groups of L2Aers, which is closer to "3"; indicating that there is acceptability for this condition among the L2Aers. However, the acceptability of the +RO condition among the native speakers of English is close to "2", which indicates rejection for this condition.

The acceptability of the -RO among the L2Aer groups is again very similar and also close to "3", which indicates acceptability of the condition; however, as it can be seen in the individual analysis of the sentences, there is high variability in the degree of acceptance of these sentences. Lastly, the native speakers of English show a low acceptability of the -RO sentence condition, which is close to "2". These descriptive results demonstrate what was observed by Alexopoulou and Keller (2002, 2007, 2013); English speakers reject both resumption and gap conditions but seem to prefer gaps over resumptives.

Table 5-4 **Descriptive Statistics of the GJT Means in the Group of Intermediate L2Aers**

Sentence Condition	No. of participants	Minimum	Maximum	Mean	Std. Deviation
+RO	22	1.0	4.0	**2.955**	**.8004**
-RO	22	1.5	4.0	**2.932**	**.7761**

*The means and standard deviations have been emphasized in bold

Table 5-5 **Descriptive Statistics of the GJT Means in the Group of Advanced L2Aers**

Sentence Condition	No. of participants	Minimum	Maximum	Mean	Std. Deviation
+RO	27	1.0	4.0	**2.852**	**.7444**
-RO	27	1.0	4.0	**3.000**	**.7206**

Table 5-6 **Descriptive Statistics of the GJT Means in the Group of Native Speakers of English**

Sentence Condition	No. of participants	Minimum	Maximum	Mean	Std. Deviation
+RO	24	1.0	3.5	**1.750**	**.7223**
-RO	24	1.0	4.0	**2.375**	**.7837**

In conclusion, the analysis of the degree of acceptability of the sentences and the descriptive statistics should serve to understand the variability on the degree of acceptability of the resumptive (+RO) and gap in resumptive position (-RO) conditions among L2Aers and native speakers. These results are also useful to understand if said degree of variability depends on the speakers' parsing strategies (which might be linked to their IDs in WMC); the results of the statistical tests in the following subsections should serve to comprehend the latter with more specificity.

5.2.1.2 Chi-square analysis on the acceptability of the sentence types

In order to find if there was a significant difference in the responses that the three groups of participants indicated to the type of sentences, a chi-square analysis on SPSS was applied using the categorical values assigned to the data of this task. The results of the chi-square analysis of the filler sentences are provided in Appendix F. The following is a report of the results pertaining to this analysis organized by type of sentence.

With this in mind, there was a significant difference between groups in terms of acceptance of the sentence type +RO *(X2(4) = 29.26, p < 0.05)*. The results in this analysis indicate that while *68.2%* of the participants in the intermediate group, and *59.3%* of the participants in the advanced group of learners accept this type of sentences, *83.3%* of the participants in the group of native speakers rejected them. These results can be observed the next table. The latter is comparable to the results obtained in the chi-square analysis conducted on the data of the pilot of the GJT, which indicate that *75.0%* of the English

speakers (*N = 32*) also reject the +RO sentence condition; significantly more than the group of Spanish speakers *(X2(2) = 24.22, p < 0.0001)* who took the GJT in Spanish. Only *11.1%* of the Spanish speakers *(N = 27)* in the pilot GJT rejected the +RO condition. The specifications of the chi-square analysis of the pilot are provided in appendix D.3.

Table 5-7 **Results in percentages per group of participants for Sentences with an Object Resumptive pronoun (+RO).**

Categorical response to +RO		**Intermediate Group**	**Advanced Group**	**Native Speakers Group**
Reject	Percentage	**18.2%**	**18.5%**	**83.3%**
	Count	4	5	20
No Tendency	Percentage	13.6%	22.2%	4.2%
	Count	3	6	1
Accept	Percentage	**68.2%**	**59.3%**	**12.5%**
	Count	15	16	3
	Total no. of participants per group	22	27	24

Also, there was a significant difference among the three groups of participants in how they reacted to -RO sentence type (*X2(4) = 13.194, p < 0.05).* These results indicate that the participants in both of the groups of learners, *68.2%* in the intermediate and *74.1%* in the advanced, accept this type of sentence. However, only *29.2%* of the participants in the group of native speakers accepted the grammatical condition of this type of sentences. These results are reported in Table 5-8.

Table 5-8 **Results in percentages per group of participants for Sentences with the Object Resumptive pronoun removed (-RO).**

Categorical response to -RO		**Intermediate Group**	**Advanced Group**	**Native Speakers Group**
Reject	Percentage	22.7%	18.5%	37.5%
	Count	5	5	9
No Tendency	Percentage	9.1%	7.4%	33.3%
	Count	2	2	8
Accept	Percentage	**68.2%**	**74.1%**	**29.2%**
	Count	15	20	7
	Total no. of participants per group	22	27	24

The degree of acceptability of the -RO condition among native speakers of English is also visible in the results observed in the chi-squared analysis conducted on the data

extracted from the GTJ pilot study; the results indicate that only *25.0%* of the English speakers accepted the -RO condition. There is also a significant difference between the degree of acceptance of the -RO in English and Spanish (*X2(2) = 23.70, p < 0.0001)*; just a *3.7%* of the Spanish speakers rejected the -RO condition in Spanish. Moreover, these results further confirm the results observed in Alexopoulous and Keller (2002, 2007, 2013) as it was mentioned in the previous section.

5.2.2 Working memory capacity and acceptability of R pronouns

As it has been stated, the main purpose of Study 2 is to explore if working memory capacity (WMC) has any effects on the degree of acceptability of R pronouns as this grammatical property differs from Spanish (L1) to English (L2) (cf. Leal-Mendez & Slabakova, 2012); but, more importantly, because R pronouns might aid L2Aer speakers of Spanish in the parsing (cf. Alexopoulous & Keller, 2007, 2013) of long-distance English sentences where these type of pronouns are likely to occur (cf. Leal-Mendez & Slabakova, 2012). Thus, this subsection is dedicated to report on the results of the multiple regression analysis used in this study to measure the effects of WMC on R pronoun acquisition among L2Aers.

In order to use this statistical model, it was necessary to transform the data of the WMC and GJT variables to correct some correlational problems as well as to make the data stable. Also, it was necessary to make the data of the GJT less disperse, since the percentage values were taken as the data for this statistical analysis. With this in mind, the following is a description of the results observed in this multiple regression analysis organized per group of participants.

First, for the group of intermediate learners, a significant regression equation was found *(F (6,15) = 3.504, p < .023)* with an R^2 of *.58*, which indicates that WMC can explain the degree of acceptability of the types of sentences +RO and -RO involved in the GJT in English. The coefficients in the multiple regression analysis indicate that one sentence condition in particular is significant to this statistical model: the -RO (sentences with a gap in object resumptive position). The results indicate that for every *1.753* higher scores obtained in the WMC span tasks, the intermediate L2Aers will be more accepting of the -RO condition (by *1%* more); in other words, if WMC increases, so does the acceptability of the -RO condition among the group of intermediate L2Aers.

Moreover, given the Durbin Watson values shown in this analysis *(DW = 1.218)* as well as the covariance values and the tolerance factor, it can be concluded that WMC has a significant effect on the acceptability of the sentences that have the R pronoun removed in object position (-RO) among the intermediate L2Aers. The reason to use Durbin Watson values is that they allow to observe the Variance Inflection Factor (VIF), which indicates if there is an effect of multicollinearity between constant or independent variables (in the case of this study the sentence conditions are the constant predictors and/or independent variables in the multiple regression analysis model) (Frost, 2017). Multicollinearity between constant variables can obscure the significance values in a multiple regression model as it indicates that the dependent variable (in this case WMC measures) does not necessarily explain the changes in the independent variables (Frost, 2017). The overall results of the multiple regression analysis between the type of sentences of the GJT and WMC performed with the data extracted from the Intermediate group is summarized on Table 5-9 below.

Table 5-9 **Results of the Multiple Regression Analysis between WMC and the type of sentences in the GJT for the Intermediate Group**

Type of Sentence		Non – standardized coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Standard Error	Beta			Tolerance	VIF
	+RO	-.592	.354	-.441	-1.673	.115	.400	2.502
	-RO	1.753	.433	1.015	4.052	.001	.443	2.259

*The significant regression equations found in this analysis are highlighted in grey.

Furthermore, for the advanced group, there was not a significant regression found *(F (6,20) = 2.180, p < .089)* with an R^2 *of .39* between WMC data and the data for the type of sentences comprising the GJT to test the object R pronoun (+RO) acceptability. However, by observing the coefficients of this analysis, the +RO sentence condition seems to be significantly influenced by WMC among the advanced L2Aers. This might mean that for every *.766%* drop in WMC scores, the advanced L2Aers will accept +RO sentences by a *1%*. Nonetheless, it is important to mention that according to the statistical values of collinearity observed in this model, such as the Durbin Watson values *(DW = .934),* WMC cannot significantly explain the acceptability of the sentence conditions -RO and +RO included in the GJT among the advanced L2Aers. Thus, even when the model for multiple regression found significant relations between sentence conditions such as -RO and GRL and WMC in the advanced group, these results cannot be trusted as the VIF values indicate that there is a multicollinearity effect between the variables that represent the sentence conditions (the VIF values should approximate 1 or less to indicate that there is no multicollinearity effect)

(Frost, 2017). A relation of the results found on the multiple regression analysis on the type of sentences of the GJT and WMC for the advanced group of learners can be seen in Table 5-10 below. The tables with the overall results obtained in the multiple regression analysis for the advanced group can be found in Appendix K under K.2.1.

Table 5-10 **Results of the Multiple Regression Analysis between WMC and the type of sentences in the GJT for the Advanced Group**

Type of Sentence		Non-standardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Standard Error	Beta			Tolerance	VIF
	+RO	.617	.330	.608	1.866	.077	.285	3.506
	-RO	-.766	.369	-.737	-2.078	.051	.241	4.157

Finally, there was a non-significant regression equation observed *(F (6,17) = 2.221, p < .091)* with an R^2 of *.43* between WMC and the sentence conditions in the GJT among the native speakers of English. Although, the Durbin Watson values indicate that there is not an autocorrelation effect between the variables involved in this analysis, the *probability value* is greater than *.05*; the latter confirms that there is not a significant effect between WMC and the acceptability of the sentence conditions +RO and -RO. The complete results of the multiple regression analysis performed for the group of native speakers is provided in K.2.2.

Chapter 6 Discussion

This research has been driven by the need to understand the effects of working memory capacity on the acceptability of L2 grammatical properties on late learners. The reason for this is to understand if the variability in the attainment of L2Aers can be explained by their individual differences (IDs) in working memory capacity (cf. Cowan, 2005). Research in second language acquisition has focused on working memory (e.g. Dussias & Piñar, 2010; Fortkamp, 1999; Juffs & Harrington, 2011; Ellis & Sinclair, 1996; inter alia) as a factor that might have a major influence in the attainment of the L2 given that WM is a cognitive mechanism that stores information temporarily while it utilizes higher-order skills from long-term memory (Baddeley, 2000, 2007, 2017).

Considering the latter, the current work has been dedicated to exploring if IDs in WMC can explain aspects of the L2 in which L2Aers show variable degrees of attainment. The aspects chosen in the current research are 1) oral fluency and 2) object resumptive (R) pronouns. The first aspect, oral fluency, has been selected owed to the many cognitive functions that are required for its development (cf. Skehan, 2003; Segalowitz, 2010; inter alia). Similarly, the acceptability of R pronouns is undertaken because it is an uninterpretable feature (cf. Chomsky, 1995; Hawkins & Hattori, 2006; Tsimpli & Dimitrakopoulou, 2007) whose grammatical strategy differs from Spanish (L1) to English (L2) (cf. Leal-Mendez & Slabakova, 2012); but most importantly, owed to its grammatical parsing repercussions: R pronouns might alleviate the processing load imposed by the distance between the head antecedent and its referent in Ā-bound long-distance dependencies (cf. Alexopoulou & Keller, 2007, 2013).

Hence, these two aspects show variability at the production and grammatical levels of L2 development and attainment among L2Aers. As it has been discussed, the present investigation takes into account that theorists and researchers under emergentism have proposed that L2 learners are able to reach L2 automaticity through the frequency of encounter with L2 linguistic items, and rehearsal (cf. Ellis, 1998; Ellis & Schmidt, 1998; MacWhinney, 1986; Matesa & Anderson, 2000; McLaughlin & Heredia, 1996). Furthermore, I have discussed that some proponents of genSLA have found that late learners are able to accurately distinguish L2 grammatical features that are purely morphosyntactic and not available in their L1 (cf. Rothman & Slabakova, 2017; Juffs, 1998, 2005; White, 2003, 2007; White & Juffs, 1998). Within these two perspectives, it needs to be further addressed that, as I have pointed out, not all L2Aers reach a level of attainment characterised by L2

automaticity or by a complete L2 grammatical repertoire; this, in spite of the L2Aers' frequent encounter with L2 items, rehearsal, and exposure to the L2.

That is why, WMC has been considered an influential factor on the degree of attainment among this particular group of learners. Nonetheless, I seek to approach WMC from an integral perspective, as I have measured it using automated computer software (Conway et al., 2005; Oswald et al., 2014; Redick et al., 2012; Unsworth et al., 2005) that allows to test all of its components, including central executive functions (Daneman & Carpenter, 1980; Wen, 2016). Moreover, the present research involves groups of learners with different levels of proficiency: intermediate and advanced; a characteristic that has been missing in studies of similar nature (e.g. Fortkamp, 1999; Dussias & Piñar, 2010; Mizera, 2006; inter alia). With the latter in mind, this chapter aims at interpreting the results from Study 1 and Study 2 of the current research

6.1 Discussion of results in Study 1

6.1.1 Working memory capacity

The current study involved two groups of Spanish-speaking learners of English; one group was comprised of intermediate learners *(N = 22),* and the other of advanced learners *(N = 27).* In addition, there was a group of native speakers of English *(N = 24).* Thus, the first part of the study involved asking the participants in these three groups to take two working memory span tasks: a listening span task and a reading span task (Conway et al., 2005; Oswald et al., 2014; Redick et al., 2012; Unsworth et al., 2005). Both of these span tasks measure executive working memory, and thus are known as complex WM tasks (Juffs, 2011; Wen, 2016).

In other words, these tasks were designed based on the Daneman and Carpenter (1980) WM verbal span task, which measure the functions of WM envisioned in the Baddeley and Hitch (1979) model. Therefore, the results for WM in this study indicate the participants' capacity to temporarily store information while making use of higher-order skills retrieved from long-term memory (Baddeley, 2000, 2003a, 2007, 2017).

With this in mind, the first statistical base regarding WMC in Study 1 is the means obtained per group in both the RST and LST. As seen in chapter 4, the means for the total WMC measure per group are the following: intermediate *M = 57.91*, advanced *M = 37.56*,

native speakers *M = 84.63*. Statistically, there were significant differences between the three groups of participants *[F (2,70) = 13.930, p < .0001].*

In this regard, the post hoc Games-Howell analysis shows that the significant differences are observed particularly between the native speakers and the intermediate learners *(CI = .84, 52.59, p < .042),* and between the native speakers and the advanced learners *(CI = 26.01, 68.13, p < .0001)* in the Total Working Memory value. There were no observed significant differences among the two groups of learners in this same value. The only particular result here found was that there is no significant difference amid the native speakers and the intermediate group in the Reading Span Task *(CI = -22.58, 10.56 p < .652);* the advanced group and the native speakers did differ significantly in this task *(CI = -33.17, -9.21, p < .0001).*

The results obtained regarding WMC should be discussed in light of previous studies that have investigated the correlation between WM and the proficiency of L2ers. For example, in the study by Prebianca et al. (2014), it is noticeable that learners with a lower proficiency scored significantly higher in three automated span tasks that measure WMC. The research included three groups of learners of Portuguese who were categorized into low, intermediate, and advanced proficiency. The chi-squares used in this study show that there were significant differences between the three groups. However, the post hoc Games-Howell analysis in this study does not show any significant differences between the intermediate and the advanced groups of learners.

Moreover, L2 proficiency has also been a factor that significantly correlates with WMC in the research work by Finardi and Weissenheimer (2009), and Prebianca (2009). Altogether, these research studies explain that the cognitive mechanisms involved in working memory are more “present” and active for L2 adult learners who are still undergoing the process of acquisition. What this might represent for L2 acquisition is that learners with a lower linguistic knowledge of the L2 are relying on working memory more than advanced learners to perform in L2 tasks. In both of these studies, the suggestion about WM is that it is a cognitive mechanism that L2Aers can activate when the L2 learning process is in a more demanding stage.

In the work of Prebianca et al. (2014), Finardi and Weissenheimer (2008), and Prebianca (2009), WM is not necessarily an individual difference, but an obligatory mechanism that L2Aers need in beginning stages. Stages in which, according to these

authors, the L2 process tasks involve attention focus, storing and retaining input, inhibiting interference from their L1, while operating with complex knowledge from long-term memory. Thus, the only resource that can assist the L2Aer with such cognitive load would be working memory given its functions. Nevertheless, as it has pointed out by the statistical findings in the present study, working memory is not correlated by level of proficiency. Although the means show that the intermediate learners obtained higher scores in the WM span tasks, the statistical analysis demonstrates that this is not significant.

Therefore, the results observed in this study should point to the approach of WM as an individual difference rather than as a general learning mechanism that is available to L2Aers in beginning stages of acquisition; and/or reduce WM as a mechanism that L2Aers need to "activate" in order to undergo the L2 acquisition process. The focus, as in this research, is to specify the effects that the individual capacity in working memory of L2Aers have in the acquisition of complex L2 aspects; and establish, by means of testing groups of participants with different proficiency levels, at what stage in the L2 learning process, the L2Aer's WMC determines the attainment of said complex L2 features.

6.1.2 Oral fluency

In order to analyse the participants' oral fluency, they were asked to describe one of two drawings presented to them in at least 30 seconds. One of the drawings consists of a family in a kitchen sharing a meal, while the other portrays a park with people engaging in multiple activities; the drawings used for this Speech Generation Task can be seen in Appendix E. As described in detail in 4.1.2, the oral description of each participant was recorded and analysed using PRAAT (De Jong et al., 2009).

This software allows to measure the participant's utterance fluency (Segalowitz, 2010) by estimating: 1) speed fluency, 2) breakdown fluency, and 3) repair/repetition fluency (Skehan, 2003; Tavakoli & Skehan, 2005). Each of these three dimensions is comprised by measures of time in seconds (except for one measure under repair/repetition fluency), which are estimated following formulae proposed by De Jong et al., (2009). The formulae and description of the measures of the three dimensions of oral fluency are provided in Table 2-1 (Kormos, 2006). All in all, the measures that comprise the three dimensions of utterance fluency make it possible to have quantifiable data of oral fluency, as well as to study it more precisely from a cognitive perspective (Segalowitz, 2010).

With the latter in mind, the statistical analysis performed on the three groups are summarized as follows. With regards to speed fluency, the One-way ANOVA analysis indicates that there are significant differences between the three groups of participants in all of the five measures that encompass this dimension: mean syllable duration *[F (2,70) = 6.051, p < .004],* speech rate *[F (2,70) = 14.176, p < .0001],* mean length of utterance in syllables [*F* (2,70) = 7.109, *p* < .002)], articulation rate *[F (2,70)= 5.15, p < .008]*, and mean length of utterance in seconds *[F (2,70) = 3.32, p < .042]*.

The post hoc Games-Howell analysis under the One-way ANOVA allowed to specify where these significance differences could be found; meaning, among what groups of participants. The findings show that the significant differences in these five measures are mainly between the intermediate and both the native speakers and advanced groups. The latter is evidenced in the statistical results presented in Table 4-8. In this table it is clear that the only significant difference in terms of speed fluency between the advanced group *(M = 2.38 SD = .52)* and the group of native speakers *(M = 2.76 SD = .52)* is in terms of speech rate *(p < .011),* the differences between the intermediate and the advanced in native speakers is significant in all of the measures under this dimension.

These findings demonstrate that although the group of native speakers' speed fluency is the highest among the three groups, it is not significantly different than that of the advanced learners of this study. However, the speed fluency of the intermediate learners is significantly the lowest amid the three groups; it is even significantly lower than that of the advanced learners. The means and standards deviations for the measures of speed fluency can be observed in Table 4-5.

As Bosker et al. (2012) point out, based on Skehan (2003, 2009) and Tavakoli & Skehan (2005), speed fluency can be "[…] characterized as the rate and density of speech delivery" (p. 160). Thus, the results mentioned above show that the advanced learners in this study have a quantifiably denser L2 speech delivery, which has no significant differences as to the delivery and rate shown by the native speakers. The intermediate learners' speed fluency, on the other hand, is significantly different in rate and density of speech delivery. These assertions could be further interpreted in light of what Segalowitz (2016) and De Jong (2018) comment in terms of measures, or what they refer to as temporary or objective measures comprising utterance fluency. In their research, they have found that these measures usually correlate with ratings of perceived fluency. This means that measures of speed fluency, such as speech rate, tend to correlate with the ratings granted to L2 speakers' oral fluency in

empirical studies (e.g. Cucchiarini et al., 2002; Derwing et al., 2004; Kormos & Dénes, 2004, Rossiter, 2009) as reported by De Jong (2018). More explicitly, De Jong (2018) adds that "highly fluent speakers (as attested by judges) were found to speak faster, […], than those who were judged to be highly nonfluent" (p. 241).

In terms of breakdown fluency, there are no observed significant differences between the groups of participants in any of its three measures (no. of pauses per minute total time, no. of pauses (speaking time/phonation time), and mean pause duration). This is confirmed in the One-way ANOVA analysis. However, after performing the post hoc Games-Howell analysis, there is a significant difference between the group of native speakers and the intermediate learners in the measure of no. of pauses per minute (speaking time/phonation time) *($p < .037$).* The significance is not major, but it does reveal that the intermediate speakers would take *.37* seconds more to pause in between syllables than the native speakers. Nevertheless, there is no significant differences in this measure between the groups of learners.

According to Bosker et al. (2012), breakdown fluency "concerns the extent to which a continuous signal is interrupted" (p. 160). Therefore, the statistical analysis of the objective measures of this aspect of utterance fluency evidences that the pausing time that intermediate and advanced learners is overall similar; and, in the case of the advanced learners, their pauses are not significantly longer than those of the native speakers of English.

Nevertheless, the significant dissimilarity in the temporary measure of no. of pauses per minute (speaking time/phonation time) here found differs with the results observed in a similar study conducted by Temple (2000). In her research, Temple included 30 adult learners of French as foreign language, who she reports are at intermediate and advanced levels, and 20 adult French native speakers. The participants were interviewed and recorded; the recordings were used to extract temporal measures of what the author calls disfluency. These temporal measures are similar to the ones utilized in this research to estimate the breakdown fluency aspect; namely, Temple calculated silence ratio, pause rate, and repair rate. Her results show that there is a significant correlation between the learners and the native speakers in all of these measures (Temple, 2000, p. 293).

It is important to highlight that Temple did not separate the learners by L2 proficiency. If all of the learners considered in her study were more in an advanced than in an intermediate stage of acquisition, then her findings are in line with the ones observed in the

current research. There is no way, however, to verify if this is the case. In any instance, the fact that the intermediate learners' utterance fluency is characterized by pauses that are significantly longer than that of the native speakers could signal that the processes underlying their L2 speech production is not yet carried out in long-term and/or procedural memory (Hulstijn, 1990). Therefore, the significant longer pauses in the group of intermediate learners might indicate that their L2 oral fluency is not undertaken by automatic processing (see 2.3.3 for a detailed discussion on the relation between L2 oral fluency and automaticity).

In reference to the evidence of longer pauses in measures of breakdown fluency, Lahman et al. (2017) and De Jong (2018) suggest that when these pauses are significant is owed to a lack of efficient connection between the conceptualizer, the formulator and the articulator; these are the agreed cognitive processing stages of speech in the research literature related to L2 fluency (De Jong, 2018, p. 245). The latter stages underlying speech can be seen in the model by Levelt (1989, p. 9) in Figure 2-5. Nonetheless, the current study did not involve a task that could gather data regarding the real-time speech behavior of the participants; nor did it include evaluators that could provide further information with regards to the type of pauses made by the intermediate learners.

Finally, the results regarding oral fluency also involve repair fluency. The One-way analysis performed on the objective measures that comprise this dimension of fluency shows that there are significant differences in total no. of repairs/repetitions *[F (2,70) = 9.82, p < .0001]* and on total length of repairs/ repetitions *[F (2,70) = 13.61, p < .0001]* amid the three groups of participants. To be more precise, there is a significant difference between the native speakers and both the advanced *(p < .049)* and the intermediate *(p < .0001)* groups of learners in no. of repairs/repetitions as observed in the post hoc Games-Howell analysis. Moreover, the intermediate group significantly differs from both the advanced *(p < .021)* and the group of native speakers *(p < .0001),* as well as the advanced from the native speakers *(p < .048)* in total length of repairs/repetitions. The latter is confirmed in the post hoc Games-Howell analysis for this measure, which can be observed in Table 4-10.

The statistical analyses performed on the measures of repair fluency demonstrate that the intermediate group of learners makes considerably more repairs in their speech, as well as they tend to have more instances of repeated speech items than the group of native speakers. Similarly, the advanced learners' speech production is also significantly characterised by a higher number of repairs and repetitions when compared to that of the

native speakers. Furthermore, the time spent by both groups of learners at making these repairs and repetitions is significantly higher than the native speakers *(M = .30 SD = .36).* Again, it is observable that the intermediate group has the highest mean *(M = 1.71 SD = 1.21)* in the time spent to these instances of speech; it is even significantly higher than the advanced learner's *(M = .81 SD = .98)* length of time for repairs and repetitions.

In spite of not finding significant differences between the advanced learners and the native speakers in terms of breakdown fluency, the significant dissimilarities in terms of repair fluency are a finding worth discussing. For example, in his research, Mizera (2006) finds that there are no significant differences in breakdown and repair fluency between an advanced group of learners and a group of native speakers. In light of this, the author comments that the proximity in these two aspects fluency amid advanced learners and native speakers has to do with the general ability of individuals to engage in what Ejzenberg (2000) calls "the juggling act" when producing speech.

This juggling act is described as the characteristic that more fluent native speakers possess, and it consists of being more "creative" when pausing or making repairs; thus, giving the impression of being more fluent. Nevertheless, the case of the group of advanced learners in this study with regards to repair fluency could not be interpreted as a general ability of the individual to seem more fluent given the significant differences found among this group and the native speakers. However, the latter is a discussion that requires further research consisting of an analysis of data from a perceived fluency perspective as I mentioned before. Data from a perceived fluency task could provide an insight on how the length of pauses, the repairs, and the repetitions in the L2 speech of the learners in this study affect their fluidity (Segalowitz, 2016), and the extent to which they may or may not cause a perceived disfluency (De Jong, 2018; Segalowitz, 2016; Temple, 2000).

All in all, the discussion on the oral fluency findings in the current study contribute to understand adult L2 learners' fluency a) at different levels of proficiency, b) in comparison to native speakers, and c) from a quantitative cognitive point of view. This should be important given that Lahmann et al. (2017) points out that "a few studies have compared fluency measures of non-native speakers with those of native speakers mostly in terms of the duration of utterances spoken in a L2" (p. 231). According to the authors, the need to expand L2 fluency quantitative research to the comparison of late learners and native speakers would shed light to the distinctive features underlying this complex aspect of acquisition.

Furthermore, studies that are restricted to the listener's evaluation need to be complemented with tasks that involve objective measures of oral fluency. As De Jong (2018) highlights, "fewer studies have related objective measures of fluency to overall (oral) proficiency or have tracked gains within speakers over time" (p. 240). Therefore, it is crucial to set an empirical preceding for L2 oral fluency that allows to comprehend its dynamism, its complexity, and to identify the factors that play a role in its acquisition amid late learners. The latter should serve as an introduction to the following subsection since it approaches the results obtained from analyzing the temporal measures of oral fluency here discussed with regards to the capacity in working memory of the three groups of participants.

6.1.3 Question 1: Does working memory capacity have an effect on the oral fluency of late L2 learners?

The previous section discussed the results observed in the analysis of the oral fluency of the groups of intermediate and advanced learners, and the native speakers involved in this study. In this discussion, it was shown that the intermediate learners' oral fluency is significantly different from the groups of advanced learners in native speakers in terms of speed fluency, a temporary measure of breakdown fluency, and in repair fluency. The most significant differences between the advanced learners and the group of native speakers is only in terms of breakdown fluency. The statistical evidence in these results points to the conclusion that the group of intermediate learners' oral fluency is characterized by shorter time lengths of continuous speech, longer pauses in between syllables, and more repairs and repetitions in their utterances.

The previous discussion is relevant for this section provided that the purpose is to determine the extent up to which working memory capacity can explain the observed results in the oral fluency of L2Aers; as well as to have be able to compare them with the effects that it has on native speakers' fluency. In order to understand the effects of WMC in oral fluency, I conducted a multiple regression analysis in SPSS. To be able to perform this statistical method, the data for the temporal measures of oral fluency had to be conflated so that they could fit into the requirements of this analysis. Thus, the temporal measures of breakdown and repair fluency were assimilated into one measure of negative quantities of oral fluency; whereas the measures comprising speed fluency were reconfigured into a single measure that indicated a positive component of oral fluency.

Such conflation is in line with what recent empirical studies have found in terms of the multiple temporal measures under the three aspects of oral fluency. For example, Segalowitz et al. (2017) finds that only three temporal measures have a significant correlation with the factors taken into account by listener's when grading fluency. This is also reported in Bosker et al. (2012) in which it is concluded that speed, breakdown, and repair fluency are aspects that significantly predict the listener's perceived fluency (Segalowitz, 2010). However, they suggest that certain measures under the three dimensions of fluency are "interrelated" (Bosker et al., 2012), and highly correlated; thus, they make reference to the same characteristic of fluency.

With this in mind, the measures in this study were conflated to avoid multicollinearity effects (cf. Kormos & Dénes, 2004). Nevertheless, it was important to weigh all the temporal measures that were considered for Study 1 to have a more accurate referent of the participants' L2 utterance fluency. As De Jong (2018) comments "[all temporal measures] together [are] significantly better at predicting perceived fluency than any other parsimonious model" (p. 240). The data utilized for WMC is the sum of the scores from the RST and LST tasks, which has been referred to as "total WMC measure". After obviating the considerations made for the oral fluency data, what follows is a summary of the results observed in the multiple regression analysis.

The results show that the only significant correlations between working memory capacity and second language oral fluency are found in the group of intermediate learners *($F(2,18) = 5.098, p < .003$).* This regression also indicates that working memory capacity can significantly explain an increase in the measures that indicate higher levels of fluency, such as speech rate, articulation rate, syllable duration, etcetera. In addition, the results indicate that, in the case of the intermediate learners, a higher WMC score means a reduction of time in measures of breakdown fluency (mean pause duration, no. of pauses per minute) and in measures of repair fluency (total number of pauses, total length of repairs/repetitions). This result is also validated by the *R^2 of .48* obtained in the regression equation between WMC and L2 oral fluency that indicates a large size effect for the intermediate group of learners (Larsen-Hall, 2015). Thus, there is enough statistical data that substantiates that as working memory capacity increases, so does oral fluency for L2Aers in an intermediate stage of acquisition.

These results need to be interpreted first by paying attention to the stages that underlie the cognitive processes of oral fluency, second by considering emergentist

theoretical principles, and lastly by relating to the development of L2 automaticity. For the first point, in the previous subsection, 6.1.2, I made a reference to the Levelt (1989) model and commented that researchers have agreed that the conceptualizer, the formulator and the articulator are the stages that cognitively precede speech production (cf. De Jong, 2018; Lahnman et al., 2017).

Central to the research work of Temple (2000), is to emphasize that two of the aforementioned stages, namely the Formulator and the Articulator, are carried out in long-term memory, and thus are automatized. For Temple, then, what causes fluent speech is the automatic nature of the processes underlying speech production, which are "beyond the realm of working memory" (p. 292). This assertion can account for the results observed in oral fluency, in which the group of intermediate L2Aers reached significantly lower temporal measures of speed fluency, and higher measures for breakdown and repair fluency; thus, serving as evidence that this particular group of learners is not undergoing L2 fluency processes automatedly, but in working memory. In addition, the fact that the L2 oral fluency of the advanced group is not significantly distinct to that of the native speakers signals that they have managed to automatize L2 speech production processing.

The results here observed are in line with the proposal that WMC is used by L2Aers at the Conceptualizer and Formulator levels (cf. Temple, 2000) based on the framework by De Bot (1992) and his adaptation of the L2 speech production process by Levelt (1989,1999). It is also in line with the notion proposed by Fortkamp and Bergsleithner (2007) who claim that those L2 learners with a higher WM span are able to produce more "error-free" speech than those with a lower WMC. As I discussed in 2.3.1, the involvement of WMC in the underlying speech production process of L2 speech would be necessary to cope with the demands found at the Conceptualizer; particularly at the microplanning level, since it is where the information to meet the communicative goals is selected (which might require to choose either the L1 or the L2 as suggested by De Bot, 1992). Furthermore, WMC might also be needed at the Formulator stage, where the lemma retrieval, grammatical and phonetic encoding take place; in this stage, the L2 learner might not only be dealing with the limitations of an incomplete vocabulary and grammatical acquisition, but also with the interference of their L1 lexical and grammatical information (De Bot, 1992; Kormos, 2006; Segalowitz, 2010; Temple, 2000; Fortkamp & Bergsleithner, 2007).

Overall, as mentioned in Skehan (2009; 2014), WM is active at the Conceptualizer stage, in which the influence of dynamic relationships proposed by Segalowitz (2010) (e.g.

frequency of exposure, social context, etc.) is present as well. At this point both L1 and L2 speakers engage in WM as they need it to integrate linguistic and non-linguistic information that allows them to make decisions regarding the communicative task at hand; which needs both storage and executive functions of WM. However, as indicated before, the L2Aer might encounter more processing challenges at the microplanning level of the Conceptualizer than the native or L2 advanced speaker. The challenge in processing and the need to use WM extends to the Formulator stage for the intermediate L2 learner; given that they need to maintain information under attentional control, while they select the appropriate L2 lexical items as well as the required morphosyntactic and phonetic operations.

In the visual representations of both De Bot's (1992) Blueprint for L2 speech production and Segalowitz' (2010) framework of the dynamic relationships that influence fluency, I added an indication of where WM is used for L1/native speakers or L2 advanced learners as well as for L2 intermediate learners (see Figure 2-3 and Figure 2-4 for a visual representation). There, I signaled that considering the storage and higher-order processing functions of WM (cf. Baddeley & Hitch, 1974; Baddeley, 2000; inter alia), the utilization of this mechanism extends up to the Formulator stage; nevertheless, for L1/native speakers and/or advanced L2 learners the operations at the Formulator level are WM-free as they are now under automatic processing. The fact that WM is used not only to cope with higher loads of information, but also with greater demands in processing has an effect in the L2 learner's fluency, which results in lower levels of oral fluency (cf. Temple, 2000; Fortkamp & Bergsleithner, 2007).

This is confirmed with the results found in this study, in which the oral fluency of the intermediate learners is characterized by more dysfluencies (longer pauses and repairs) compared to that of the advanced L2 learners and the native speakers. What is more, the results here found also serve to shed light on how IDs in WMC are significantly related to the intermediate L2Aers' oral fluency; the higher the capacity in WM, the less dysfluencies there are in the L2 learners' oral fluency. Nevertheless, it also serves to demonstrate that the influence of WMC is only significant when the L2 learner is still undergoing the L2 acquisition process since the results here observed show that there are no significant correlations between the advanced learners' WMC and their temporal measures of oral fluency.

This observation is also attested in the work of Mizera (2006), who reports a lack of correspondence between WMC on L2 oral fluency in advanced learners. The absence of correlations between the distinct measures of L2 oral fluency and WMC among in Mizera's

work is in line with what has been pointed out by Temple (2000) and discussed in relation to the findings of the present research. Mizera also makes the interpretation that the L2ers in his study have already automatized fluency processes, since they are in an advanced stage of acquisition. Therefore, working memory capacity might not be a determining factor to develop oral fluency at such level of proficiency.

The latter leads to the second point of the present discussion, in which I deem it necessary to touch on emergentist theory to further interpret the findings obtained in Study 1. In 2.1.1.5, I explained that working memory plays an important role in the L2 acquisition process under emergentist frameworks. In their view, L2 learning requires the perception of input, to be turned into intake, to later be "automatized" by means of extensive practice (Ellis, 2008; McLaughlin & Heredia, 1996). The interim stage in which L2 input turning into intake becomes automatized depends on controlled processing (Anderson, 1992).

Controlled processing, in emergentist theory, is regarded as a the initial stage of the L2 acquisition process; when the L2 learners is in this stage, the demand in their cognitive abilities is high given the difficulty that acquiring an L2 represents. For emergentists, the functions of WM are employed by the L2 learner when undergoing the controlled processing stage of acquisition (see 2.1.1.5 for a detailed discussion on the role of working memory in emergentist theory). Therefore, proponents of this theory assume that automatic processing indicates successful L2 acquisition. In this vein, the findings of the multiple regression showing that working memory capacity can only significantly explain L2 oral fluency in the intermediate group of learners are in line with the emergentist view on L2 acquisition; working memory is an aiding mechanism in the development of L2 fluency while the L2Aers is still in an intermediate stage of acquisition. Nevertheless, the approach to WM under emergentism should be taken with caution. As the results in the present study show, the individual differences in the capacity of working memory is what explain the levels of oral fluency, and not just working memory as a general cognitive mechanism.

Finally, with regards to the third point, the results here found go in accordance with theory that relates L2 fluency to L2 automaticity. As discussed in 2.3.3, L2 oral fluency has been described as an aspect that derives from automaticity. Arguably, automaticity can be defined in DeKeyser's (2001) terms as a "fast, parallel, effortless, capacity-free, unintentional, result of consistent practice, little interference from and with other processes […]" (p. 39). Nevertheless, this conceptualization seems problematic as it is difficult to empirically test if L2Aers have reached an effortless processing, as it is also limiting to characterize the highest

cognitive level of L2 acquisition in terms of speed. Hence, the view by Favreu and Segalowitz (1983) with regards to automaticity is more accurate as they see it as a type of cognitive processing that involves efficiency, which they have been able to demonstrate empirically. Moreover, Segalowitz and Segalowitz (1993) argue that automaticity also implies stability in processing as the result of re-structuring and re-organizing L2 knowledge.

With the latter in mind, the observed results showing that the advanced learners' L2 fluency is not significantly different from that of the native speakers could serve as evidence of their reaching an efficient and stable L2 processing, or L2 automaticity. Consequently, the lack of significant WMC effects on the L2 fluency of this group reiterates that the reaching of automaticity implies less or no reliance on working memory. On the other hand, the results among the intermediate learners demonstrate that higher capacity in working memory can significantly explain increases in speed fluency and decreases in breakdown and repair fluency. Thus, higher capacity in WM might lead to develop L2 automaticity in a faster and more efficient manner; lower capacity in WM, on the other hand, might cause for L2Aers to take a longer period of time to develop automaticity, or to undergo more challenges to achieve it. The latter, therefore, can serve as evidence for how IDs in WMC are a factor that causes variability in L2 attainment among this subset of L2 learners.

All in all, the results show that WMC does have an effect in the development of L2 oral fluency among L2Aers, but only at an intermediate stage of acquisition. This finding also demonstrates that working memory is an individual characteristic and its effects on an aspect as complex as L2 oral fluency depend on the L2Aer's capacity to operate the functions of this cognitive mechanism. Therefore, it can be concluded that those late learners in an intermediate state of L2 acquisition with a higher working memory capacity are able to reach higher levels of oral fluency; that is, their L2 oral fluency will be characterized by a more continuous and denser speech delivery, with fewer instances of disfluency.

6.2 Discussion of results in Study 2

The present section seeks to analyse the results obtained in the statistical analysis of 1) the responses to the sentence conditions in the GJT, and 2) the effects of WMC on the responses to the sentence conditions +RO and -RO among the groups of participants. As it has been pointed out, the importance of such analysis lies in the need to understand if the degree of acceptability of L2 uninterpretable features (Chomsky, 1995; Hawkins & Hattori, 2006) such as object resumptive pronouns, which have been claimed to pose a processing

challenge for the English grammatical parser (Alexopoulou & Keller, 2007, 2013) is influenced by the L2Aers' IDs in WMC.

Thus, to approach the research aims of Study 1, I selected sentences that contained the following grammatical conditions: resumption in object position (+RO), and a gap in object resumptive position (-RO). These sentences were extracted from sources of authentic English and Spanish spoken use. The latter was achieved by using corpora software; specifically, the Corpus of Contemporary American English (COCA) (Davies, 2008), and the Corpus de Referencia del Español Actual (CREA). In this manner, the sentences used were intended to expose the groups of L2Aers to sentences that they could potentially encounter in real-life situations; which, could also force the L2Aer to apply deep grammatical L2 knowledge to parse and emit acceptability judgments towards these particular sentence conditions. However, these sentences were altered to fit the purposes of the objective of Study 2; therefore, the number of syllables were adapted to have a similar length on all the GJT sentences as well as ungrammaticalities were purposely caused on the filler UGRL sentence types by removing pronouns in obligatory positions.

Furthermore, the reason that the object resumptive pronoun was chosen was because there are different grammatical strategies applied in Spanish (L1) and English (L2) when it comes to resumption (cf. Leal-Mendez & Slabakova, 2012). Mainly, given that Spanish is a null-subject language, resumptives are realized as a clitic dislocated to the left (CDDL), whereas in English R pronouns are overtly realized in long-distance sentences where there is no subjacency violation. Some genSLA proponents such as Tsimpli and Dimitrakopoulou (2007) claim that L2Aers are not able to reset those grammatical strategies that involve uninterpretable features (cf. Chomsky, 1995) in the L2, and that are used differently in the learners' L1; such as a resumption.

Nonetheless, Leal-Mendez and Slabakova (2012) have demonstrated that L2Aers can accurately distinguish R pronouns in the L2 (English) even when they are uninterpretable features and the strategy grammatically differs in their L1 (Spanish). Leal-Mendez and Slabakova (2012) explain that the accurate distinction of features such as resumptives in the L2 depends on the L2Aers' L1 individual grammars. However, the latter claims seem insufficient to explain why there is an observable variability among late L2 learners in terms of the properties that constitute their grammars; and, particularly with properties that involve different grammatical strategies from the L1 to the L2.

Given this, I propose that L2 grammatical variability among L2Aers can be explained by IDs in capacity of general cognitive mechanisms that are involved in learning such as WM; given that WM functions are involved in the parsing of sentences that pose both grammatical and semantic difficulty (cf. Just & Carpenter, 1992). Therefore, resumption should be studied from a grammatical parsing perspective. In this sense, I take into consideration the findings in Alexopoulou and Keller (2007, 2013) regarding resumptive pronouns; the authors observed that native speakers of English prefer long-distance sentences in which there is a gap (in the R pronoun position) condition over those that have resumption. Alexopoulou and Keller (2007, 2013) argue that for speakers of English, gap conditions are less costly given that gaps represent less linguistic memory costs (cf. Gibson's, 1989 Syntactic Prediction Long Distance theory); R pronouns add more items to process to the English grammatical parser given that the parser is already prepared to find a gap in R position from the moment that a head antecedent appears in the long-distance structure.

For L2Aers, nonetheless, the presence of an R pronoun in a long-distance sentence can alleviate the parsing pressure of carrying the head antecedent to the end of the sentence structure; which also involves conveying semantic and/or discursive meaning. The latter, therefore, can be related to IDs in WMC since those learners with a lower WMC can benefit from the presence of R pronouns in the sentence. Nonetheless, those L2Aers with a higher WMC may be able to cope with the difficulty of the long-distance structure with a gap condition as they might have more storage and processing tools to carry the head antecedent to the end of the sentence; assimilating in this manner, the parsing strategy applied by the native speaker of English when it comes to resumptive v. gap conditions.

6.2.1 Discussion on the results for acceptability of R pronouns

6.2.1.1 Discussion on the descriptive statistics of the GJT sentence conditions

The first statistical analysis for GJT data was intended to obtain the descriptive statistics of each of the sentences with either an +RO or a -RO condition. Thus, the means and standard deviations were reported per type of sentence (+RO, -RO) and organized by group of participants. The descriptive statistics of each sentence provide an insight on the grammatical behaviors of the L2Aer groups with respect to the group of native speakers of English. It is important to remember that participants marked the acceptability of the sentences with a Likert scale, in which "1" indicated "not natural" and "4", "natural". These numbers were conflated as follows: "1" and "2" were converted to "1", whereas "3" and "4" were turned into "4". The latter was applied in order to have a clearer view of what sentences

had been clearly accepted (closer to 4), and which had been rejected (closer to 1). The results of the descriptive analysis are discussed next.

As it can be seen in the results, the two groups of L2Aers, intermediate and advanced, show acceptability for the sentences containing a resumptive pronoun (+RO), and also for the sentences with a gap in resumptive pronoun position (-RO). However, as Table 5-2 and Table 5-3 show, the acceptability of these sentences among the L2Aer groups varies; certain +RO and -RO sentences have a mean acceptability of *2.5* or *2.6*, which indicates that there is not a clear tendency for rejection or acceptance for any of these sentence conditions. Nevertheless, the degree of acceptance among the group of native speakers of English for the +RO sentences was consistently below 2, which indicates rejection for this sentence condition. The rate of acceptability for the -RO condition varies among the native speakers of English; indicating no tendency (with means ranging from 2.5 to 2.7) for some -RO sentences, but with a clear acceptability for a couple of sentences with this condition.

In general, the means demonstrate that both groups of L2Aers have a tendency to accept both the +RO (intermediate L2Aers *M = 2.95*; advanced L2Aers *M = 2.85*) and -RO conditions (intermediate *M = 2.93*; advanced *M = 3.00*); on the other hand, the native speakers of English consistently reject the +RO condition (*M = 1.75*), but are more accepting of the -RO condition (*M = 2.37*). Given the findings in Alexopoulou and Keller (2002, 2007, 2013), I expected to observe more acceptability for the +RO condition among both groups of L2Aers and a clear rejection for this condition among the native speakers of English. Also, I predicted that the -RO was not going to be accepted among the native speakers of English; but that it was going to be more accepted than the +RO sentences in this group, based on the findings of Alexopoulou and Keller (2002, 2007, 2013).

As mentioned in the review of the literature, Alexopoulous and Keller (2002, 2007, 2013) found that resumptive and gap conditions are not acceptable among speakers of English. The authors observed, however, that sentences with a gap condition (in resumptive position) have a higher acceptability rate than those sentences with a resumptive condition. Although the results of the descriptive statistics are preliminary, they serve to have an insight about the acceptability of these conditions among the three groups of participants. Also, these results serve to provide more details about how L2Aers and native speakers of English react towards the +RO and -RO conditions; and in such manner, these results are useful to confirm the findings reported in Alexopoulou and Keller (2002, 2007, 2013).

Furthermore, Alexopoulou and Keller (2007, 2013) claim that resumptives in English are not used as a "saving device", necessarily; in other words, speakers of English do not integrate R pronouns to save grammatical violations (Alexopoulou, 2009). However, R pronouns have been found to "[…] compensate for the processing difficulty associated" (Alexopoulou, 2009, p. 151) with long-distance sentences in languages like Greek (cf. Alexopoulou & Keller, 2007, 2013), Hebrew (cf. Ariel, 1999; Hawkins, 2004), and given its resemblance to Greek, Spanish (cf. Leal-Mendez & Slabakova, 2012).

Nonetheless, the acceptability of these sentence conditions in English as an L2, as I mentioned in the previous section, might be related with the individual processing strategies that L2Aers possess given the distance between head antecedents and their referents in sentences with +RO and -RO conditions. Although the degree of acceptability might be linked to the role that resumption represents for the L2Aer speaker of Spanish, the R pronoun might aid with the parsing challenge that long-distance structures represent for those L2Aers with a lower WMC; or, the R pronoun could be unnecessary for those L2Aers who can carry the head antecedent information to the end of the sentence structure.

6.2.1.2 Discussion on the chi-square analysis for the GJT sentence conditions

In order to understand if the previous analysis had any significant degree of acceptance among the groups of participants, I conducted a chi-square analysis in SPSS. The following should provide a reference to the most relevant findings in this analysis followed by a discussion of the observed results.

First, for the +RO sentences, there is a significant difference among the three groups *(X2(4) = 29.26, p < 0.05)*. This is observed in the *59.3%* in the advanced and *68.2%* in the intermediate groups who marked acceptance for this type of sentences. Oppositely, *83.3%* of the native speakers rejected sentences with this condition. Second, for the -RO condition there is a significant difference among the three groups (*X2(4) = 13.194, p < 0.05).* In both of L2Aer groups, the majority of participants accept the -RO condition (68.2% of the intermediate L2Aers and 74.1% of the advanced L2Aers), whereas only a 29.2% of the native speakers of English accept this condition. Compared to the +RO sentences, nonetheless, the native speakers of English seem to be more tolerant of the -RO condition.

As I mentioned in the previous section, following the findings of Alexopoulou and Keller (2002, 2007, 2013), I expected to find these results; which also confirm what is observed in the results of the descriptive statistical analysis and in both the descriptive and

chi-square analysis of the pilot. In the pilot, it can be seen that English speakers reject both the resumptive and gap conditions; however, the acceptance rates for the gap condition (-RO) are higher than those of the resumptive condition (+RO). The results of the chi-square analysis of the pilot can be found in appendixes D.2 and D.3, where it can also be observed that Spanish speakers tend to be more accepting of both +RO and -RO conditions; nonetheless, Spanish speakers show more acceptability for sentences with an +RO condition.

The findings in terms of acceptability among the groups of L2Aers in comparison to the group of native speakers of English set the purpose of the present investigation since there is an observable variation in the degree of acceptability of the +RO and -RO sentence conditions among the groups of L2Aers. As I discussed in 2.4.2, the variability observed in L2 grammars must be explored in light of general cognitive mechanisms of learning such as WM; since as Alexopoulous and Keller (2007, 2013) propose, the acceptability of resumption might be linked to processing costs.

Therefore, the claim in Alexopoulous and Keller (2007, 2013) needs to be studied vis-a- vis WM. As I have mentioned before, the parsing of L2 long-distance sentences in English might depend on the L2Aers' IDs in WMC; and, R pronouns might ease the cost that these long-distance structures represent for L2Aers that have a lower WMC.

6.2.2 Question 2: Does working memory capacity have an effect on the acceptability of object resumptives among late L2 learners?

As mentioned in the previous subsection, the main discussion of the present study focuses on the results obtained in exploring the effects of WMC on the degree of acceptability of object resumptives (+RO) among L2Aers. In order to explore the effects of WMC on this feature, I conducted a multiple regression analysis on SPSS. To perform this analysis, I took into account the total measure of WMC and the percentage values given to the sentence conditions +RO and -RO in the GJT. Thus, the following is an interpretation of the statistical results obtained in study 2.

The results show that the only significant equation between WMC and the degree of acceptance of the sentence conditions was observed in the group of intermediate learners *(F (6,15) = 3.504, p < .023)*. Specifically, the statistical model found that WMC can explain the degree of acceptance of sentences with a gap in object resumptive position (-RO) in a significant manner. The analysis demonstrates that as WMC scores increase, the percentage

for the acceptance of -RO sentences are higher. These findings are more detailed in Table 5-9. Moreover, the model shows that there is no collinearity found in the statistical analysis performed for this group. Thus, WMC can explain the degree of acceptability of the -RO sentence condition with an R^2 *of .58,* which represents a large size effect in empirical studies in second language acquisition (Larson-Hall, 2015).

This finding is very relevant since it confirms the claim in Alexopoulou and Keller (2007, 2013) with respect to resumptives. For Alexopoulou and Keller (2007, 2013), R pronouns have repercussions for the English grammatical parser as they add more costs to the processing load imposed by Ā-bound long-distance dependencies (where Rs usually occur). First of all, as it was discussed in the review of the literature, resumptive pronouns in languages like English and Spanish have been thought of as "saving devices" (cf. Ariel, 1999; Hawkins, 2004) given that they aid the parser to overcome the processing complexity involved in carrying the head antecedent information to the end of the long-distance sentence structure (cf. Alexopoulou, 2009). However, in a crosslinguistic study, Alexopoulou and Keller (2002) observed that among English speakers, R pronouns did not help to "rescue" the acceptability of said long-distance constructions. In fact, the authors found that English speakers significantly preferred constructions with a gap in R position (-RO) over those with an R condition (+RO).

As observed in the chi-square analysis in this study and in the pilot of the GJT, the results show that both -RO and +RO sentence conditions have a low degree of acceptability. Nevertheless, the -RO sentences were more accepted than the +RO ones. The latter, though, was not the case for the two groups of L2Aers, who seem to be more accepting of both conditions; however, there is an observed variability in the degree of acceptance of these conditions.

These results seem to indicate that R pronouns are indeed related to individual processing strategies; and therefore, do not necessarily depend on the L2Aer's L1 grammatical repertoires, as argued by Leal-Mendez and Slabakova (2012). The latter is also confirmed by the results in the Pearson covariance analysis on the +RO and -RO conditions in English in Spanish, which show that there is not a significant correlation in the acceptability of these conditions among the intermediate and advanced groups of L2Aers (the results of the Pearson co-variance analysis can be found in Appendix J). The results also contradict the claim by Tsimpli and Dimitrakopoulou (2007) that L2Aers apply the strategies of their L1 to the L2 when it comes to uninterpretable features such as resumptives.

Therefore, the finding observed among the intermediate L2Aers with respect to the significant relation between WMC and the -RO condition demonstrates that L2Aers do rely on processing mechanisms such as WMC to parse long-distance Ā dependencies; the latter serves as supporting evidence for the explanation that Alexopoulou and Keller (2007, 2013) offer regarding resumptive and gap conditions and their implications for the grammatical parser in English. For Alexopoulou and Keller (2007, 2013), R pronouns in English create a recursive syntactic processing load that makes the parser stop and "go back" in the sentence to look for the head antecedent to which the R pronoun is attached; thus, resulting in a higher processing load for the English grammatical parser. The authors base on Gibson's (1998) Syntactic Prediction Locality Theory (SPLT). Under this theory, new discourse referents in a sentence become more costly for the parser as they add a linguistic integration memory cost when processing a long-distance structure.

In Gibson's SLPT theory, the parser makes syntactic predictions about the predicate of a long-distance construction; this prediction includes all morphosyntactic elements that are to occur in the sentence. However, under the Gibson's (1998) SPLT principles, if a discourse referent takes place in a structure, after the syntactic prediction has been made, it causes for the parser to seek, in a cyclical manner (meaning, that it "goes back" in the sentence and then "comes back" to the referent), where in the sentence the head that such referent attaches to is located.

Alexopoulou and Keller (2007, 2013) explain that R pronouns add the processing cost of a new discourse referent to long-distance constructions in English. The authors explain that the English parser predicts that a gap will occur in the long-distance Ā-bound dependency from the moment that the head antecedent is formed, and not a resumptive pronoun. Hence, English speakers accept gap conditions over resumption since the gap is the anticipated outcome that the grammatical parser expects to find once the predicate prediction has been made for the Ā-bound long-distance structure. For L2Aers though, the acceptability of gap conditions (-RO) is related to a higher WMC as the finding of the multiple regression analysis for the intermediate L2Aer group indicates.

Therefore, for those L2Aers whose WMC is higher, the gap condition (-RO) is more acceptable as they do not need the aid of an R pronoun to cope with the parsing load of carrying the information of a head antecedent to the end of the long-distance sentence structure. This might also indicate that L2Aers with a higher WMC are more likely to imitate

the grammatical strategies that native speakers of English apply when parsing long-distance Ā-bound dependencies. Therefore, the finding that IDs in WMC among intermediate L2Aers can significantly explain the degree of acceptability of -RO sentence conditions serves to understand why there is variability in the configuration of the late learner' L2 grammar.

What is more, this finding is in line with studies of a similar nature, which have also found that WMC has significant effects on late learners' L2 grammars. For example, in Sagarra (2017), it is reported that WMC influences the grammatical comprehension of adult learners in initial stages. In order to make this conclusion, Sagarra (2017) measured WM with a span task that allowed to report scores under timed conditions, and in which all components of WM, meaning central executive functions, were tested as well. Even though Sagarra's (2017) study is longitudinal (effects of WMC were measured pre and post L2 instructional exposure periods), the results converge with the ones here found.

Nevertheless, the grammatical features under study in the present investigation are more specific; unlike in Sagarra's (2017) work where L2 grammar was measured in a general manner, and where the specific grammatical features under testing are not reported. Therefore, the result here found attests that WMC has effects not only on ordinary grammatical knowledge, but also on L2 linguistic features that require very specific parsing strategies such as long-distance constructions with a resumptive (+RO) or a gap (-RO) condition.

Nonetheless, it is important to notice that Sagarra (2017) points out that the influence of WM is more present amid learners with a lower L2 proficiency, or in an ab initio state of acquisition. The author reports that this is the case of other empirical studies of similar nature, in which it is concluded that WM has a higher effect in developing L2 grammars (in ab initio states), but not necessarily in advanced states (e.g. Linck & Weiss, 2011, 2015; Santamarina & Sunderman, 2015; Serafini and Sanz, 2016; Sanz et al., 2014; Havik et al., 2009; Miyake & Friedman, 1998).

The findings in the studies described above serve to understand the results observed among the advanced L2Aers in the present work. As it can be seen in the multiple regression analysis, there is not a significant correlation between the WMC scores and the acceptability of the +RO and -RO sentence conditions found among the advanced L2Aers. Other empirical studies that have approached the composition of L2 grammars on adult learners from an

explicit/implicit learning perspective as Denhovska (2014) observed that the mediating effects of WMC are not visible on advanced learners (p. 144).

Conversely, Dussias and Piñar (2010) report significant correlations between WMC scores and the acquisition of L2 -wh movement on an advanced group of learners. Sagarra (2017) argues that the finding in Dussias and Piñar is due to the use of span tasks that measure WM on the learners' L2; testing WM with tasks in the L2, according to the author, conducts learners to employ functions of WM to process L2 input and tasks. Nonetheless, the present study also involves WM span tasks in English, the learners' L2, and the results differ from the ones observed in Dussias and Piñar (2010).

The relevance of the results found in the present study, though, should be considered in terms of the approach taken to measure WMC and the type of L2 grammatical properties that were measured to understand variability in attainment amid L2Aers. First, taking into account Juffs and Rodriguez (2014) who report that previous research on WM and L2 morphosyntactic processing and acquisition among adult learners had not shed any significant relations between these two aspects, the results here observed are an indicator that the instruments used to measure WMC need to be carefully selected.

Therefore, the application of two automated span tasks that measure both phonological and executive functions of WM (complex working memory span tasks as referred to by Wen, 2016) is a more reliable tool to explore the scope that IDs in WMC have on L2 grammatical aspects among L2Aers. Also, it is essential to point out that aside from avoiding unreliable "older methods" to measure WMC (Juffs and Rodriguez, 2014), the present results shed more light on current research trends that demand further specification of where in the developmental process, WMC has more significant effects in the configuration of the L2 grammar.

In conclusion, the results obtained from the statistical analysis on study 2 serve as proof for the influence that IDs in WMC have on the acceptability of L2 grammatical properties amid L2Aers. More specifically, the results show that intermediate L2Aers with a higher capacity in WMC are able to consistently accept sentences with a gap in object resumptive position (-RO). In other words, intermediate L2Aers with a higher capacity in WM prefer sentences in which the object R pronoun is not explicit as they can retain the information of the antecedents involved in the structure, even at the cost of solving distant relativization in the sentence; the latter also might indicate that they are able to apply

grammatical strategies that native speakers of English use for the parsing of long-distance Ā-bound dependencies (cf. Alexopoulou, 2010; Alexopoulou & Keller, 2007, 2013).

Chapter 7 Pedagogical Implications, Limitations and Conclusions

7.1 Pedagogical implications

In the iconic paper by Juffs and Harrington (2011) on working memory and second language acquisition, it is stated that "if WM is a trait that cannot be changed, it is hard to see how it can be altered to improve language learning. However, if WM is linked to attentional capacity, or control of the Central Executive function, which can suppress competing pressure from the L1, […], then learners could be assisted in focusing on L2 forms and attempting to suppress L1 influence" (p. 159). In this sense, two points have to be made with regards to the findings observed in this thesis and the literature in terms of a) WM training in general instances of learning, and 2) the consequences of the latter in L2 instruction.

To the first point, general instances of learning and WM training, the research is taking important shifts. For example, Gathercole et al. (2019) have studied the effects of WM training and its applicability in the learning of new skills. Their findings demonstrate that WM training, including tasks that involve Executive Control (a.k.a. Central Executive) functions, have moderate effects in the learning of new skills in situations in which the new skills have similar characteristics to the tasks involved in the training of WM functions. In order to train WM, the authors considered that the learning of skills, cognitively speaking, is based on constant practice (or The Law of Practice by Newell & Rosenbloom, 1981). However, the authors propose that this practice should involve an extensive and well-structured execution of tasks in which the functions of WM are employed; meaning, in tasks in which information needs to be stored, while at the same time higher order skills stored in long-term memory are executed. All in all, the results observed in the research of Gathercole et al. (2019) has a positive implication for late L2 learners in intermediate stages whose WM capacity is low or lower as it is likely to be trained to obtain the same learning gains as those learners with a higher WMC.

More specifically, this type of training could be applied to L2 instruction settings directed to intermediate L2Aers. Since the findings in the present research demonstrate that a higher WMC in intermediate late learners represents more fluency and acceptability of complex grammatical structures, then instruction for this subset of learners should be based

on strategies in which WM functions are extensively employed; particularly in tasks that foster oral production and grammatical comprehension. In this regard, not only should the tasks involve rehearsal of the input in L2 learning settings, but also meaning comprehension at a deep critical level. The latter is of crucial importance to involve the Executive Control and Episodic Buffer (cf. Baddeley, 2017; Baddeley et al., 2019) functions of WM which increase long-term recall, and thus, the learning of information (cf. Craik et al., 2019).

Hence, L2 teachers of intermediate L2Aers should be able to guide adult intermediate language learners to a practice that is not merely mechanical, but that it also leads the learner to make semantic associations with the L2 input in both grammatical and oral activities. As Baddeley (2004) highlights, "[…], there is no doubt that a word or experience that is processed in a deep way that elaborates the experiences and links it with prior knowledge, is likely to be far better retained than one that receives only cursory analysis" (p. 3). Based on the Levels of Processing by Craik and Lockhart (1972), Baddeley (2004) suggests that the creation of deep semantic associations, linked to WM functions, lead to successful learning gains. In L2 teaching settings, therefore, activities should contain discussions about the context where the L2 forms take place, the implications, the cultural expectations, etc. Also, late L2Aers in intermediate levels should be encouraged to engage in activities that allow them to relate the new L2 information to prior knowledge in a gradual and constant manner.

In this regard, in the Information Processing (IP) framework, (cf. McLaughlin & Heredia, 1996, see 2.1.1 for a detailed discussion on this SLA theory) it is proposed that the L2 input needs to be acquired in a hierarchical manner, in which frequency of exposure, pattern recognition, intrinsic reasoning, and the creation of L2 output opportunities need to occur constantly. The latter implies for the functions of WM to be employed. Namely, the frequency of exposure to the L2 input would trigger the Phonological Loop and, even the Visual Sketchpad functions to retain information. Also, having learners recognize patterns, and applying reasoning skills when processing the L2 input would lead to the employment of Control Executive functions. In addition, since there is a limited capacity of storage in WM, the suggestion of having L2 information delivered hierarchically and frequently in this framework might aid the L2Aer to overcome this burden in WM processing of L2 input. However, this framework needs to further emphasize the importance of utilizing functions of the Control Executive since it has concentrated more on investigating the effects of the functions of Phonological WM for L2 acquisition (e.g. Ellis & Sinclair, 1996).

Nevertheless, what is suggested by Skehan et al. (2012) in terms of the application of pre-, task, and post-task instruction in L2 pedagogical environments can compensate for the limitations in the IP framework to SLA. The authors have found that when L2 learners are involved in pre- and post-tasks in which they need to collaborate, plan, and monitor their use of the second language, they obtain gains in accuracy and L2 knowledge construction (see also Foster & Skehan, 2013 who present more empirical evidence of pre and post tasks in L2 instruction). Therefore, in L2 classrooms, L2Aers can be asked to predict what a conversation in the L2 is about before listening to the actual conversation by focusing on certain phrases, words, a title or even images; this can be done in pairs or groups. Moreover, after listening and/or reading to this conversation, they can discuss with their classmates about the gist, the type of information/interaction that took place, how they relate to the conversation, explain the conversation in their own words, identify L2 forms that were recurrent, how the conversation connects to other topics with which they are familiar, etc. Constant discussion leads to opportunities for L2 oral output production, while it engages the learner in deep reasoning, thinking, elaboration, etc. The latter might result in the employment of Control Executive functions, and the other components of WM.

In terms of L2 knowledge grammar building, it can be suggested to adopt the framework of input processing for L2 instruction introduced by VanPatten (1996). In this theory, the author suggests that L2 learners should be guided to focus on both the grammatical form and the semantic meanings of the L2 input. To do so, a Processing Instruction (cf. VanPatten & Cadierno, 1993) framework for L2 pedagogical environments with the following characteristics is proposed:

> a. It is predicated on sentence-level processing strategies that learners take to the task of comprehending a second language (L2).
> b. It is input oriented.
> c. Input is manipulated (structured) in particular ways to alter processing strategies and increase better intake for acquisition.
> d. It includes [explicit instruction] for the learner on both grammatical structure and processing problems.
> e. It follows certain guidelines for the creation of structured input (SI) activities.
>
> (VanPatten & Borst, 2012, p. 271).

What this implies, broadly speaking, is that the L2 learner should not only be presented with L2 structures, but should also be guided to understand the morphosyntactic composition of said structure (detecting the distinctive orders, forms, inflections, etc.); this can be achieved by presenting them with several examples that target that particular L2 form. Furthermore, the Processing Instruction theory emphasizes that there must be

activities that engage the learner in understanding the L2 form in a semantic manner as well. Hence, L2 instruction should approach grammatical acquisition in a way in which learners are encouraged to extensively focus on both form and meaning comprehension; which could include extracting several concordance lines from a corpus that target a specific complex L2 form (O'Keeffe et al., 2007). The instructor can guide the learners to work in pairs or groups to identify the morphosyntactic patterns, and later have them discuss if they can add more information and/or agree with the topics of such L2 examples. This type of instruction might support the operation of WM functions in the acquisition of L2 complex grammatical properties given that it involves L2 information storage and application of higher-order cognitive skills.

Overall, curriculum developers, trainers, instructors, and all of those concerned with L2 teaching and learning should be more aware of the importance that WMC has in the L2 acquisition process among L2Aers. Thus, the findings of this thesis suggest that L2 instructional environments for adult learners should involve the constant employment of all the functions that comprise working memory; particularly, in intermediate levels. Ultimately, the implementation of activities that foster the learning of a second language through working memory might help L2Aers to develop better levels of oral fluency and a more successful comprehension of morphosyntactic constructions.

7.2 Limitations of this research

The research here developed does not go without limitations. In terms of the methodology, the tasks utilized to measure WMC should have also included span tasks in Spanish. Unfortunately, at the time that I collected the data, automated versions of the RST and LST were not available in Spanish. It is important to make this consideration for future research to avoid that the potential limitations in the L2 proficiency of the learner interfere with the tasks' purpose of measuring their capacity to apply WM functions.

However, the findings of this investigation should not be disregarded owed to this limitation. Appropriate actions were taken to overcome the burden of not having automated versions of the tasks in the learners' native language. These actions included explaining the instructions of the tasks in Spanish, both orally and on paper. Also, there was an extensive revision on the lexicon and structures used in both tasks and it was verified that the learner was not going to encounter unfamiliar and/or highly complex L2 items. Nevertheless, one of future research goals is to develop working memory span tasks, as the ones employed in this

investigation, in Spanish. It is necessary to have these tools available to have reliable measures of the learners' WMC at various levels of proficiency.

Finally, one limitation was not to be able to include reaction times (RTs) on the GJT. Adding RTs would have provided the investigation with more tools to analyse the results of Study 2: the effects of WMC on the acceptability of R pronouns, in terms of processing. This aspect was considered; however, due to technological and economical limitations, it was not possible for me to get access and/or design a GJT that could measure RTs. Also, I would like to suggest for future researchers to find as much guidance as they can to control for the sentence conditions under testing in their research; for me, the sentences that I selected became very challenging as there were some aspects that I should have considered to modify to make sure that the sentences did not have any distractors that could influence the acceptability of the grammatical properties under study. Nevertheless, I feel satisfied that I included sentences extracted from corpora sources as these types of sentences guarantee that the conditions under testing are not used in prefabricated sentences/contexts, which could potentially limit the validity of the acceptability demonstrated by the speakers in the investigation.

7.3 Conclusions

This investigation aimed at studying the effects of WMC (Baddeley, 2017; Daneman & Carpenter, 1980; Cowan, 2005) on L2 oral fluency (Segalowitz, 2010) and the acceptability of object resumptive pronouns (Leal-Mendez & Slabakova, 2012) amid late learners. Although emergentist (e.g. Ellis, 2008; Ellis & Sinclair, 1996; McLaughlin & Heredia, 1996, inter alia) and generativist (e.g. White, 2003, 2007; White & Juffs, 1998; etcetera) researchers have empirically demonstrated that late L2 learners can successfully develop high levels of L2 proficiency, it is evident that there is variability in attainment among this subset of learners (cf. Schmid, 2011). Therefore, this research has focused on exploring if IDs on WMC can explain said variability in attainment by studying two aspects of the L2 that pose a challenge amid L2Aers.

Numerous studies have approached the role of WM in SLA (e.g. Dussias & Piñar, 2010; Ellis & Sinclair, 1996; Fortkamp, 1999; Juffs, 2004; Mizera, 2006; Sagarra, 2017; inter alia); however, there is a lack of a unified criterion in terms of the functions and the influence of WM in the L2 research and literature (Wen, 2015, 2016). Some problematic assumptions in the empirical study of WM have been that 1) only the phonological component has been

considered (e.g. Ellis & Sinclair, 1996), 2) the tasks do not reliably measure WMC (e.g. Juffs, 2004), and/or 3) IDs in WMC have been overlooked (e.g. Temple, 2000; Sagarra, 2017). In order to narrow the gaps in understanding WMC in the L2 attainment process of late learners, this thesis has paid special attention to measure the capacity in WM, including Central Executive functions (Daneman & Carpenter, 1980); predicting that a higher WMC leads to more successful levels of L2 proficiency.

With the latter in mind, the effects of WMC were explored on two aspects of the L2 that are challenging in terms of oral production and comprehension: oral fluency and object resumptive pronouns, respectively. To do so, two studies were conducted. In the first study the effects of WMC on speed, breakdown, and repair fluency (Skehan, 2003; Tavakoli & Skehan, 2005) were calculated. The results indicate that WMC can significantly explain the levels of L2 oral fluency amid intermediate learners. Thus, this finding indicates that, in an intermediate stage of acquisition, adult learners with a higher WMC have higher measures of speed fluency, and lower measures of breakdown and repair fluency. Such finding supports emergentist theories that view WM as a mediating resource that leads to L2 automaticity (as defined by Segalowitz & Segalowitz, 1993), in lower and intermediate stages of the L2 developmental process (cf. Ellis, 2008; McLaughlin, 1996; McLaughlin & Heredia, 1996; inter alia). However, the result of this study emphasizes that 1) there are individual differences in WMC, and 2) all the functions of WM, including Central Executive, are involved in the development of L2 oral fluency. The latter is important since emergentist research and theorists have not highlighted these two aspects in their approach to WM and SLA (e.g. Ellis & Sinclair; Skehan, 2016).

In the second study, WMC was measured in terms of the effects that it has on the degree of acceptability of L2 grammatical conditions (White, 2003; 2007; White & Juffs, 1998); which included sentences with object resumption (cf. Leal-Mendez & Slabakova, 2012; Rouveret, 2011) and gaps in object resumptive position (Alexopoulou & Keller, 2002, 2007, 2013). In this analysis, it can be observed that WMC has a significant effect on the degree of acceptance of sentences with a gap in object resumptive position in the intermediate group of L2Aers. Thus, intermediate learners with a higher WMC are more accepting of sentences in which the R pronouns are not overt. The latter demonstrates that those L2Aers with a WMC do not need to rely on resumption to parse sentences in which there is distance between the head antecedent and its referent (cf. Ariel, 1999; Hawkins, 2004; Alexopoulou & Keller, 2002; 2007, 2013; Tsimpli & Dimitrakopoulou, 2007; Leal-Mendez & Slabakova, 2012). This finding is also relevant because it shows that intermediate L2Aers with a higher

WMC apply similar processing strategies as native speakers of English. The latter is supported by the finding in Alexopoulou and Keller (2007, 2013) who show that native speakers of English predominantly prefer sentences with a gap condition over resumption as the former is less costly to parse than the latter grammatical condition (cf. Gibson's 1998 SPLT). Thus, a higher WMC serves late learners in an intermediate stage of the L2 developmental process to parse L2 grammatical structures in a more "native-like" and efficient manner.

All in all, the findings of this research shed light on the cognitive causes that explain why there is variability in L2 attainment among late learners. More specifically, the results of this thesis are relevant to understand that IDs in WMC significantly influence the development of L2 oral fluency and the acceptability of L2 long-distance structures with or without a resumptive condition when the L2Aer is in an intermediate stage. Therefore, these findings should serve to conduct more research that focuses on the effects that IDs in WMC have in SLA amid late learners. More importantly, more empirical efforts should be made to better comprehend how the conclusions of this research can be transformed into more effective pedagogical approaches that can lead the L2Aer to successfully acquire a second language.

Appendix A **Software details for RST and LST**

A.1 Reading Span Task

To take RST online: http://www.millisecond.com/download/library/rspan/

SCRIPT INFO

Script Author: David Nitz (dave@millisecond.com) for Millisecond Software, LLC.
last updated: 08-18-2016 by K. Borchert (katjab@millisecond.com) for Millisecond Software LLC

A.2 Listening Span Task

To take RST online: http://www.millisecond.com/download/library/listeningspan/

Automated "Listening" Span Task (ALSPAN)

SCRIPT INFO

last updated: 01-20-2016 by K. Borchert (katjab@millisecond.com) for Millisecond Software LLC

This script is based on the Automated Reading Span Task (ARSPAN) Inquisit script
Copyright for the ARSPAN: Task Copyright (c) 2005 by Dr. Randall Engle http://psychology.gatech.edu/renglelab/
Inquisit programming for the ARSPAN:
David Nitz (nitz.david@gmail.com for Millisecond Software, LLC).
Auditory stimuli addition for the ALSPAN:
K. Borchert, Ph.D. (katjab@millisecond.com) for Millisecond Software, LLC.

Appendix B **Sentences in the English GJT**

Note:

a) The number of clause embeddings are highlighted with a particular color.
b) The color "light grey" has been selected to highlight the clause embedding that contains the grammatical condition under testing.
c) The clause embeddings were counted based on whether or not they contained a verb.
d) The number of syllables and embeddings are provided below each sentence or sentence set.
e) The clause that contains the grammatical property under testing has been highlighted **in bold**.
f) For the -RO sentences, the asterisk (_) signals where the resumptive (R) pronoun has been removed in the clause.
g) For the UGRL sentences, the asterisk (_) indicates where the relative pronoun has been removed in the clause causing ungrammaticality.

B.1 Sentences with object resumptives (+RO) and sentences with the R pronoun removed (-RO)

1. Despite the attention given in recent days to allegations by an Okinawan woman **who an American airman raped her**, molestation of women and sexual harassment received little coverage.

No. syllables: 55
No. of embeddings: 2

Despite the attention given in recent days to allegations by an Okinawan woman **who an American airman raped ***, molestation of women and sexual harassment received little coverage.

2. He talked a lot from the speech he gave the other day **that if you've not looked at it** is a good summary of the work the United States has been doing with respect to the Asian regional economy.

No. syllables: 54
No. of embeddings: 3

He talked a lot from the speech he gave the other day **that if you've not looked at *** is a good summary of the work the United States has been doing with respect to the Asian regional economy.

3. The issue is not if gays are in the military. It is if they can be there without lying about it since it is a strict code of conduct **that if they violate it** would lead to dismissal from the service.

No. syllables: 52
No. of embeddings: 5

The issue is not if gays are in the military. It is if they can be there without lying about it since it is a strict code of conduct **that if they violate *** would lead to dismissal from the service.

4. I know the cast felt the same and our producers as it's the best and one of the greatest living playwrights, Jon Robin Baitz **who he wrote this piece with such eloquence and compassion for all the different characters**.

No. syllables: 53
No. of embeddings: 3

I know the cast felt the same and our producers as it's the best and one of the greatest living playwrights, Jon Robin Baitz **who he wrote *** with such eloquence and compassion for all the different characters.

5. We were expecting this foreign man, Maroof Farooq, to show up after having a sexually explicit chat with a decoy **who he told him** that this young girl was a 12-year-old female.

No. syllables: 50
No. of embeddings: 4

We were expecting this foreign man, Maroof Farooq, to show up after having a sexually explicit chat with a decoy **who he told *** that this young girl was a 12-year-old female.

Medavoy was involved in some great films such as "One Flew Over the Cuckoo's Nest" during his tenure at United Artists; "Amadeus" and "Silence of the Lambs" at Orion Pictures, **that he co-founded it**.

No. syllables: 51
No. of embeddings: 1

Medavoy was involved in some great films such as "One Flew Over the Cuckoo's Nest" during his tenure at United Artists; "Amadeus" and "Silence of the Lambs" at Orion Pictures, **that he co-founded ***.

B.2 Sentences containing a relative clause used as sentence fillers

The grammatical relative clause has been highlighted in bold

B.2.1 Fillers with a grammatical relative clause condition (GRL)

1. It's smart, so now the association of People for the Ethical Treatment of Animals is suing on behalf of the monkey claiming **that it owns the photo** and has the right to make money off of it.

No. syllables: 54
No. of embeddings: 3

2. This is something that's being pushed by big banks so they can basically railroad a couple of guys **who they don't want to pay licensing fees on anymore**, says an aide to a Senate Democrat who voted against the patent bill.

No. syllables 53
No. of embeddings: 5

3. Many of the ministers who came to this morning's cabinet meeting had warned Mrs. Thatcher last night that she had lost so much support within her party and **that she would probably lose the leadership race.**

No. syllables: 51
No. of embeddings: 3

4. Therefore, every individual has the choice now as to when they are going to have this type of operation**, who they will go to**, and they are still bombarded by all sorts of commercialism and advertisement.

No. syllables: 53
No. of embeddings: 3

5. I don't think of the abortion **that she decided to get as part of her personality** or, like, the most important thing that had ever happened to her because it turns out being a clear, thoughtful choice that she makes in her late hood.

No. syllables: 53
No. of embeddings: 4

6. Gangiah first caught the attention of Zalebs' creators due to the growing popularity of Jacaranda FM's The Lounge show **that she co-hosts with Cassy Clarke every weeknight between seven and ten**.

No. syllables: 53
No. of embeddings: 1

7. Kercher, a 21-year-old student from Britain, was found dead Nov. 2, 2007, in the apartment **that she shared with Knox in the idyllic hillside town of Perugia** where both women were studying.

No. syllables: 54
No. of embeddings: 2

8. Former Miss South Africa and current Miss World visited the Leeuwkop Correctional Facilities yesterday to check up on the vegetable garden project **that she started during her reign as Miss South Africa.**

No. syllables: 54
No. of embeddings: 2

9. Her work is described as rarely political, with some visceral exceptions such as the five-pointed star of Yugoslav communism **that she carved into her stomach in 1975**.

No. syllables: 51
No. of embeddings: 1

10. A regular and quite recent wardrobe recycler, the duchess, in a generous gesture, donned the same expensive fuchsia pink Mulberry coat **that she wore during the couple's official tour in New York last year**.

No. syllables: 53
No. of embeddings: 1

11. And what he recounted in some of his correspondence, Mussolini found Hitler as almost crazed, **that he would go on** and keep speaking like a phonograph, especially when the question of religion came up.

No. syllables: 53
No. of embeddings: 4

12. As it was explained by one outspoken candidate, "It really did help my student with his reading without him thinking or having the feeling **that he was doing any type of extra or academic work."**

No. syllables: 52
No. of embeddings: 4

B.2.2 Fillers with an ungrammatical relative clause condition (UGRL)

An asterisk (*) has been used to indicate where a pronoun has been removed to cause ungrammaticality in the sentence.

13. And every single individual that is working as part the White House staff will know **who * must check with** before they can speak with the press or anyone else about the matters that take place here.

No. syllables: 50
No. of embeddings: 4

14. The Republican right believes fundamentally that the Democrats **who * believe are to be blamed for the deficits over the years**, and that the Democrats have not compromised enough on spending.

No. syllables: 51
No. of embeddings: 2

15. Kevin's ex- girlfriend Clarita Kendall testified yesterday **that * visited Kevin** and he told her exactly what had happened the night when he was allegedly assaulted and shot by a woman.

No. syllables: 52
No. of embeddings: 5

16. We're planning to bring the community together in applying pressure to Iran through sanction adopted by the UN that will be aimed at those enterprises, **that * believe are supplanting Iran's government.**

No. syllables: 54
No. of embeddings: 5

17. This is just warming up and unless it gets tougher, then it's clear **that * who are running against Hillary Clinton** merely want to pursue her to select them as her vice-presidential candidate.

No. syllables: 54
No. of embeddings: 5

18. For all the thought Simon Thiel has given to how to combat aging, he actually points out **that * does not have a lot of specific ideas** about what he would do if he could live significantly longer.

No. syllables: 55
No. of embeddings: 4

19. Lisa Cohen, a former model, persuaded a New York judge to require Google to reveal the identity of an anonymous blogger **that * felt had defamed her**, and she has filed a suit against him.

No. syllables: 55
No. of embeddings: 4

20. But what do people relish in a Frank Underwood? A guy who gets things done right away and there is a certain deliciousness to it, **that * see in his journey**, but they don't see in the real Washington right now.

No. syllables: 50
No. of embeddings: 4

21. She recalled the many obstacles that she faced along the way, and how her love of Filipino food, **that * used to be ashamed of in grade school**, sustained her until her efforts were rewarded with success.

No. syllables: 55
No. of embeddings: 4

22. Let's take a look over at Janet Reno today, who had a public briefing, or what passes for a briefing at the Department of Justice **that * says some interesting things about this report**.

No. syllables: 54
No. of embeddings: 3

23. American author, Kelly Link loves fiction **that * says** "takes things which are comfortable and familiar and makes them really strange, or else... takes things which are strange and impossible and makes them feel comfortable".

No. syllables: 55
No. of embeddings: 7

24. When my husband's home and may help with some of the chores around, he gets a little bit irritated if I happen to forget to thank him or compliment him **that * did the chore in a very nice way.**

No. syllables: 51
No. of embeddings: 6

Appendix C **Sentences in the Spanish GJT**

C.1 Sentences containing object resumptives (+RO) and sentences with the object pronoun removed (-RO)

13. Aquí pues, un gran sentido de responsabilidad y de logro tiene que inculcarse en cada campesino, cada ejidatario, **para que lo** construya su propio destino, para que él se abra paso al desarrollo y a las oportunidades brindadas. (79)
14. Aquí pues, un gran sentido de responsabilidad y de logro tiene que inculcarse en cada campesino, cada ejidatario, **para que *** construya su propio destino, para que él se abra paso al desarrollo y a las oportunidades brindadas. (79)

15. Ocho ministros fueron objeto de las presiones del Presidente **que los citó a ellos** en Los Pinos, y les dijo, de forma indirecta, que ellos iban a ser responsables de la inestabilidad financiera y del desorden social político y económico. (79)
16. Ocho ministros fueron objeto de las presiones del Presidente **que los citó *** en Los Pinos, y les dijo, de forma indirecta, que ellos iban a ser responsables de la inestabilidad financiera y del desorden social político y económico. (79)

17. No obstante las justificaciones expuestas para dichas medidas, debe tenerse en cuenta **que los afectan a miles de compatriotas**, honestos, trabajadores, responsables y esforzados, para quienes dichos depósitos implican una erogación. (79)
18. No obstante las justificaciones expuestas para dichas medidas, debe tenerse en cuenta **que los afectan ***, honestos, trabajadores, responsables y esforzados, para quienes dichos depósitos implican una erogación. (79)

19. Sin embargo, los que vienen, los turistas franceses que vienen, esos sí se interesan muchísimo por la cuestión arqueológica de todo, no importando **que séanlos ellos** o no, digamos arqueólogos, especialmente arqueólogos, sino simplemente turistas. (80)
20. Sin embargo, los que vienen, los turistas franceses que vienen, esos sí se interesan muchísimo por la cuestión arqueológica de todo, no importando **que séan* ellos** o no, digamos arqueólogos, especialmente arqueólogos, sino simplemente turistas. (80)

21. No porque cambien su manera de pensar las bancadas de la Oposición, sino porque, por la lógica de ese mercado **que ellas lo adoran**, no seguirá siendo posible que compitamos con sueldos tan bajos, como si nuestro aporte fuera explotar trabajadores. (85)
22. No porque cambien su manera de pensar las bancadas de la Oposición, sino porque, por la lógica de ese mercado **que ellas * adoran**, no seguirá siendo posible que compitamos con sueldos tan bajos, como si nuestro aporte fuera explotar trabajadores. (85)

23. Y al comprar un artículo por las cualidades que se magnifican en la propaganda por televisión u otro medio, podría entenderse **que a ellas se las** encuentran incluidas en el contrato, y que en caso de no estarlo procedería entonces reclamar. (87)
24. Y al comprar un artículo por las cualidades que se magnifican en la propaganda por televisión u otro medio, podría entenderse **que * ellas se *** encuentran incluidas en el contrato, y que en caso de no estarlo procedería entonces reclamar. (87)

C.2 Fillers (sentences with relative pronouns)

C.2.1 Grammatical Fillers

1. Creo que es indispensable seguir comentando esto y quiero hacerlo personalmente con la senadora, porque es a partir de este inciso **que ella da** lectura en estos momentos en que insistiríamos en la no conveniencia de este instrumento. (80)

2. Yo insisto en que la información que se intercambia en el ámbito del grupo de contacto es valiosa para conocer los esfuerzos que hacen los estadounidenses para combatir este fenómeno y **para que ellos** conozcan lo que nosotros hacemos. (80)

3. Cuando se dice que no se puede transformar a esas entidades en organismos financieros, porque no cabe olvidar **que ellas** participan en el mercado financiero y que parte de sus ingresos, en los últimos balances, provienen de esa actividad. (80)

4. Las restricciones que impone la normativa, como, la petición del Subsecretario de Salud y la de destinar sólo diez por ciento del presupuesto anual para atención institucional, me parece **que ellas** constituyen limitaciones inadecuadas. (78)

5. Es nuestra convicción que Canadá y México tienen la oportunidad de dar sentido a la globalización a fin de evitar los riesgos **que ella** puede traer a nuestras sociedades y aprovechar equitativamente las oportunidades que nos ofrece. (81)

6. Son esas las razones que determinaron que se aprobara por unanimidad el **que ella** recibiera la Medalla Belisario Domínguez, que se instituyó para reconocer la obra de los que destacan por su actividad en beneficio de la Nación. (80)

7. Ya saben se les enseña también acerca de todos los departamentos del hospital **para que ellas** en caso de que tengan que llevar a un niño a cualquier servicio del hospital fuera de la sala, sepan dónde quedan todos esos lugares, clínicas, rayos X. (80)

8. Con cuáles empresas habló y en qué se basa para decir **que ellos** dicen que son representativos, si el salario ha perdido el quince por ciento de su poder adquisitivo, entonces, por esto es claro que estas empresas son optimistas y unas cuantas. (80)

9. La primera reflexión que me gustaría hacer es identificar la inconsistencia del Gobierno Federal en temas que domina porque se supone **que ellos** saben cómo hacerlo, y me parece que venden la imagen de que ya salimos de la crisis. (78)

10. Lo que queremos es ver papás y mamás con capacidad de generar un ingreso satisfactorio, respetable para ese hogar, y **que ellos** sean los titulares de la responsabilidad de la promoción para el bienestar y satisfacción de los hijos. (80)

11. Señor Presidente, sería conveniente ponerse institucionalmente de acuerdo para suprimir las palabras o engañosa de la indicación número ciento treinta y dos, ya **que ellas** fueron eliminadas de la norma vista anteriormente. (78)

12. En todo caso, conviene dejar en claro que este tipo de negociaciones corresponden al Ejecutivo, quien, en definitiva, debe procurar **que ellas** tengan éxito por el bien de Chile, porque nuestro país necesita estar en el Mercosur. (77)

C.2.2 Ungrammatical fillers

13. Hay municipios que reciben cien mil pesos mensuales y tienen un gasto de cien mil pesos en energía eléctrica, y aún así el Gobierno del Estado quiere que ellos lo paguen, y más aún, que también paguen, incluso, *** que** con anterioridad se adeuda. (80)

14. La pequeñez de la parcela ejidal y el minifundio no ejidal, así como las características de los cultivos **que en *** se practican, impiden absorber la fuerza de trabajo familiar disponible, generando altas tasas de desempleo.

15. ¿Por qué entonces, hace algunas semanas, en el Paso Texas, el General Barry McCaffrey va a invertir dos millones de dólares y colocar * **que él** llama el Zar antidrogas fronterizo, si es que vamos también en este combate al tráfico de drogas. (80)

16. La libertad es del hombre y para el hombre, **y que *** es quien la administra y la usa, no al revés porque si es verdad que países como el nuestro no configuran los paradigmas universales, definamos con visión de Estado, los términos de la adaptación. (81)

17. Y así, si lo tiene se lo participe a la Secretaría de Comunicaciones y Transportes para que **los*** analicen qué Programas Rurales tiene que realizar y utilice el Empleo Temporal de sedesol para hacer sus caminos de mano de obra. (82)

18. De manera tal, podría decirse, no me atrevería, sin embargo, a adjudicarle al chiapaneco Belisario Domínguez, palabras, ideas, o pensamientos **que * hayan surgido** no podría alentar o desalentar, apoyar o sostener. (80)

19. Los peces vienen multiplicados de ocho en diez, de diez en cien, de cero en cero, de diez en miles, y así podría seguir sucesivamente **para decir que** * angustia que nada pese encima de ellos**,** y que puedan continuar su travesía de igual manera. (79)

20. Estaba crecido Sansón y conociendo en un pueblo vecino a una mujer filistea, le gustó tanto **que** se enamoró **de *** y pidió a sus padres que le dieran permiso para casarse, pero sus padres le hicieron notar que tenia otras opciones. (81)

21. René Juárez es un gobierno sumamente legítimo y **que** ** eligió, que el pueblo de Guerrero ha fincado en él muchísimas esperanzas por tener avances, proyectos y soluciones en una gran mayoría de sus problemas ancestrales. (81)

22. Claro, si no hay actividad, también se pierden clases por estas intromisiones desafortunadas de los jóvenes, y entonces pues en cierto modo esto puede ser * **que** es gratificante para ellos, ¿no?, tener un poco de holganza, diríamos, de vacación forzada. (81)

23. Ahora lo curioso es que las objeciones que algún día posiblemente habría puesto Santo Tomás yo, a veces, me ponía a jugar, digamos, mentalmente, tontamente, desde luego, a buscar una objeción aún más fuerte que las **que * mismo** pondría. (81)

24. Esa nota de tipo policiaca, y de **la que * es** el editor responsable del periódico, provocó un escándalo sin precedentes, e incluso se le están adjudicando tintes de tipo politico para influenciar a los indecisos antes de las elecciones. (80)

Appendix D **Results of the pilot for the sentences in the GJT**

The degree of acceptance of the sentence conditions in the GJT in English and Spanish were tested with two pilot groups of participants. The participants in both groups reported to be native speakers of English (for GJT in English) (N = 32) and of Spanish (for the GJT in Spanish) (N = 27). Participants were asked to respond to the same GJTs that were applied to the participants in the Study 2 of the present investigation (the participants also indicated degree of acceptance of the sentence based on a 4-point Likert scale in which 1 indicates that the sentence is "not natural" and 4 indicates "natural"). The age of the participants ranged from 22 to 44 and they all reported to have at least an associate or college degree. The following is the report of the statistical analyses of the data obtained in the pilot GJTs; the analyses were conducted on SPSS and include descriptive statistics reports and the results of chi-square tests.

Sentence conditions included in the GJT=
-RO = Sentences with a gap in object resumptive position (total: 6 sentences)
+RO = Sentences with a resumptive pronoun in object position (total: 6 sentences)
GRL = Sentences with a grammatical relative clause used as fillers (total: 12 sentences)
UGRL = Sentences with an ungrammatical relative clause used as fillers (total: 12 sentences)

D.1 Report of means of the -RO and +RO sentence conditions

Report			
Group of participants		**-RO**	**+RO**
English	**N**	**32**	**32**
Spanish	**N**	**27**	**27**
Total	N	59	59

Group of participants = Native Speakers of English

Descriptive Statistics						
	N	Minimum	Maximum	Sum	**Mean**	**Std. Deviation**
-RO	32	1.0	3.5	65.5	**2.047**	**.7658**
+RO	32	1.0	3.0	54.5	**1.703**	**.6703**
Valid N (listwise)	32					

Group of participants = Native Speakers of Spanish

Descriptive Statistics						
	N	Minimum	Maximum	Sum	**Mean**	**Std. Deviation**
-RO	27	2.0	4.0	88.0	**3.259**	**.5437**
+RO	27	2.0	4.0	83.5	**3.093**	**.6939**

D.1.1 Descriptive statistics of the -RO and +RO sentence conditions in English

Group = Native Speakers of English

Statistics			
		-RO	**+RO**
N	Valid	32	32
	Missing	0	0
Mean		**2.047**	**1.703**
Std. Deviation		**.7658**	**.6703**
Minimum		1.0	1.0
Maximum		3.5	3.0

-RO					
Means of Degree of Acceptability		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
	1.0	5	15.6	15.6	15.6
	1.5	9	28.1	28.1	43.8
	2.0	5	15.6	15.6	59.4
	2.5	6	18.8	18.8	78.1
	3.0	5	15.6	15.6	93.8
	3.5	2	6.3	6.3	100.0
	Total	32	100.0	100.0	

+RO					
Means of Degree of Acceptability		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
	1.0	9	28.1	28.1	28.1
	1.5	12	37.5	37.5	65.6
	2.0	4	12.5	12.5	78.1
	2.5	3	9.4	9.4	87.5
	3.0	4	12.5	12.5	100.0
	Total	32	100.0	100.0	

Group of participants = Native Speakers of Spanish

Statistics			
		-RO	**+RO**
N	Valid	27	27
	Missing	0	0
Mean		**3.259**	**3.093**
Std. Deviation		**.5437**	**.6939**
Minimum		2.0	2.0
Maximum		4.0	4.0

-RO					
Means of Degree of Acceptability		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
	2.0	1	3.7	3.7	3.7
	2.5	3	11.1	11.1	14.8
	3.0	10	37.0	37.0	51.9
	3.5	7	25.9	25.9	77.8
	4.0	6	22.2	22.2	100.0
	Total	27	100.0	100.0	

+RO					
Means of Degree of Acceptability		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
	2.0	3	11.1	11.1	11.1
	2.5	7	25.9	25.9	37.0
	3.0	6	22.2	22.2	59.3
	3.5	4	14.8	14.8	74.1
	4.0	7	25.9	25.9	100.0
	Total	27	100.0	100.0	

D.2 Results of Chi-Square test to determine the significance of the degree of acceptance of the -RO (with a gap in object resumptive position) sentence condition

			Groups of Participants	
			Speakers of English	**Speakers of Spanish**
-RO	**Accept (3-4)**	Count	**8**	**23**
		% within -RO	25.8%	74.2%
		% within Group of Participants	**25.0%**	**85.2%**
		% of Total	13.6%	39.0%
	No Tendency (2.25-2.75)	Count	**5**	**3**
		% within -RO	62.5%	37.5%
		% within Group of Participants	**15.6%**	**11.1%**
		% of Total	8.5%	5.1%
	Reject (1-2)	Count	**19**	**1**
		% within -RO	95.0%	5.0%
		% within Group of participants	**59.4%**	**3.7%**
		% of Total	32.2%	1.7%
Total		Count	32	27
		% within -RO	54.2%	45.8%

-RO= Sentences with a gap in Object Resumptive position
Values considered for acceptance of sentence condition: 3 - 4
Values considered as demonstrating no tendency toward acceptance or rejection of sentence condition: 2.25 – 2.75
Values considered for rejection of sentence condition: 1 – 2

Chi-Square Tests for sentences with -RO condition			
	Value	**df**	**Asymptotic Significance (2-sided)**
Pearson Chi-Square	23.705[a]	2	**.000**
Likelihood Ratio	27.438	2	**.000**
N of Valid Cases	59		

D.3 Results of Chi-Square test to determine the significance of the degree of acceptance of the +RO (with an object resumptive pronoun) sentence condition.

			Group of Participants	
			English Speakers	**Spanish Speakers**
+RO	**Accept (3-4)**	Count	**5**	**17**
		% within +RO	22.7%	77.3%
		% within Group of participants	**15.6%**	**63.0%**
		% of Total	8.5%	28.8%
	No Tendency (2.25-2.75)	Count	**3**	**7**
		% within +RO	30.0%	70.0%
		% within Group of Participants	**9.4%**	**25.9%**
		% of Total	5.1%	11.9%
	Reject (1-2)	Count	**24**	**3**
		% within +RO	88.9%	11.1%
		% within Group of participants	**75.0%**	**11.1%**
		% of Total	40.7%	5.1%
Total		Count	32	27
		% within +RO	54.2%	45.8%

+RO= Sentences with a gap in Object Resumptive position
Values considered for acceptance of sentence condition: 3 - 4
Values considered as demonstrating no tendency toward acceptance or rejection of sentence condition: 2.25 – 2.75
Values considered for rejection of sentence condition: 1 – 2

Chi-Square Tests Results for sentences with +RO condition			
	Value	**df**	**Asymptotic Significance (2-sided)**
Pearson Chi-Square	24.229[a]	2	**.000**
Likelihood Ratio	26.731	2	**.000**
N of Valid Cases	59		

D.4 Descriptive Statistics of the filler sentence conditions: GRL and UGRL

Group = Native Speakers of English

Descriptive Statistics						
	N	Minimum	Maximum	Sum	**Mean**	**Std. Deviation**
GRL	32	1.75	4.00	93.75	**2.9297**	**.57584**
UGRL	32	1.00	3.50	57.00	**1.7813**	**.68906**

Filler Sentences with a grammatical relative clause condition (GRL)					
Means of Degree of Acceptability		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
	1.75	1	3.1	3.1	3.1
	2.00	2	6.3	6.3	9.4
	2.25	3	9.4	9.4	18.8
	2.50	3	9.4	9.4	28.1
	2.75	5	15.6	15.6	43.8
	3.00	7	21.9	21.9	65.6
	3.25	4	12.5	12.5	78.1
	3.50	3	9.4	9.4	87.5
	3.75	2	6.3	6.3	93.8
	4.00	2	6.3	6.3	100.0
	Total	32	100.0	100.0	

Filler Sentences with an ungrammatical relative clause condition (UGRL)					
Means of Degree of Acceptability		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
	1.00	5	15.6	15.6	15.6
	1.25	5	15.6	15.6	31.3
	1.50	7	21.9	21.9	53.1
	1.75	3	9.4	9.4	62.5
	2.00	2	6.3	6.3	68.8
	2.25	7	21.9	21.9	90.6
	3.25	1	3.1	3.1	93.8
	3.50	2	6.3	6.3	100.0
	Total	32	100.0	100.0	

Group = Native Speakers of Spanish

Descriptive Statistics						
	N	**Minimum**	**Maximum**	**Sum**	**Mean**	**Std. Deviation**
Grammatical Filler Sentences (GRL)	27	2.00	4.00	88.25	**3.2685**	**.46989**
Ungrammatical Filler Sentences (UGRL)	27	1.75	3.75	70.00	**2.5926**	**.63982**

Appendix D

Filler Sentences with a grammatical relative pronoun condition					
Means of Degree of Acceptability		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
	2.00	1	3.7	3.7	3.7
	2.50	1	3.7	3.7	7.4
	2.75	3	11.1	11.1	18.5
	3.00	4	14.8	14.8	33.3
	3.25	7	25.9	25.9	59.3
	3.50	4	14.8	14.8	74.1
	3.75	5	18.5	18.5	92.6
	4.00	2	7.4	7.4	100.0
	Total	27	100.0	100.0	

Filler Sentences with an ungrammatical relative pronoun condition					
Means of Degree of Acceptability		**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
	1.75	4	14.8	14.8	14.8
	2.00	4	14.8	14.8	29.6
	2.25	2	7.4	7.4	37.0
	2.50	6	22.2	22.2	59.3
	2.75	3	11.1	11.1	70.4
	3.00	2	7.4	7.4	77.8
	3.25	2	7.4	7.4	85.2
	3.50	1	3.7	3.7	88.9
	3.75	3	11.1	11.1	100.0
	Total	27	100.0	100.0	

Appendix E **Descriptive analysis of filler sentences in the GJT for Study 2.**

E.1 Descriptive statistics of the GRL sentences

Grammatical sentences with relative clauses	**Descriptive Statistics**	**Native Speakers**	**Advanced**	**Intermediate**
13. It's smart, so now the association of People for the Ethical Treatment of Animals is suing on behalf of the monkey claiming **that it owns** the photo and has the right to make money off of it.	**Mean**	1.87	1.88	2.22
	Standard Deviation	1.36	1.36	1.47
14. This is something **that's being pushed** by big banks so they can basically railroad a couple of guys **who don't want to pay** licensing fees on anymore, says an aide to a Senate Democrat who voted against the patent bill	**Mean**	3.37	2.66	2.63
	Standard Deviation	1.21	1.49	1.49
15. Many of the ministers **who came to** this morning's cabinet meeting had warned Mrs. Thatcher last night **that she had lost** so much support within her party and that she would probably lose the leadership race.	**Mean**	3.5	3.55	3.45
	Standard Deviation	1.11	1.06	1.15
16. Therefore, every individual has the choice now as to when they are going to have this type of operation, **who they will go to**, and they are still bombarded by all sorts of commercialism and advertisement.	**Mean**	2.25	2.33	2.63
	Standard Deviation	1.47	1.49	1.49
17. I don't think of the abortion **that she decided to get** as part of her personality or, like, the most important thing **that had ever happened** to her because it turns out being a clear, thoughtful choice that she makes in her late hood.	**Mean**	2.37	2.11	3.04
	Standard Deviation	1.49	1.44	1.39
18. Gangiah first caught the attention of Zalebs' creators due to the growing popularity of Jacaranda FM's The Lounge show **that she co-hosts** with Cassy Clarke every weeknight between seven and ten.	**Mean**	3.75	3.22	2.5
	Standard Deviation	.82	1.31	1.5
	Mean	3.87	3.33	3.45

19. Kercher, a 21-year-old student from Britain, was found dead Nov. 2, 2007, in the apartment **that she shared** with Knox in the idyllic hillside town of Perugia where both women were studying.	**Standard Deviation**	.59	1.24	1.15
20. Former Miss South Africa and current Miss World visited the Leeuwkop Correctional Facilities yesterday to check up on the vegetable garden project **that she started** during her reign as Miss South Africa.	**Mean**	3.75	3.44	3.31
	Standard Deviation	.82	1.16	1.25
21. Her work is described as rarely political, with some visceral exceptions such as the five-pointed star of Yugoslav communism **that she carved** into her stomach in 1975.	**Mean**	4.0	2.88	3.18
	Standard Deviation	0.0	1.44	1.33
22. A regular and quite recent wardrobe recycler, the duchess, in a generous gesture, donned the same expensive fuchsia pink Mulberry coat **that she wore** during the couple's official tour in New York last year.	**Mean**	3.87	2.77	3.18
	Standard Deviation	.59	1.47	1.33
23. And what he recounted in some of his correspondence, Mussolini found Hitler as almost crazed, **that he would go on** and keep speaking like a phonograph, especially when the question of religion came up.	**Mean**	2.87	2.55	3.04
	Standard Deviation	1.45	1.49	1.39
24. As it was explained by one outspoken candidate, "It really did help my student with his reading without him thinking or having the feeling **that he was doing** any type of extra or academic work."	**Mean**	3.12	3.66	3.31
	Standard Deviation	1.36	.94	1.25

E.2 Descriptive statistics of UGRL sentences

Ungrammatical sentences with relative clauses	**Descriptive Statistics**	**Native Speakers**	**Advanced**	**Intermediate**
25. And every single individual that is working as part the White House staff will know **who * must check** with before they can speak with the press or anyone else about the matters that take place here.	**Mean**	1.5	2.66	3.31
	Standard Deviation	1.11	1.49	1.25
	Mean	1.25	2.66	3.45

26. The Republican right believes fundamentally that the Democrats **who * believe are to be blamed** for the deficits over the years, and that the Democrats have not compromised enough on spending.	**Standard Deviation**	.82	1.49	1.15
27. Kevin's ex- girlfriend Clarita Kendall testified yesterday **that * visited Kevin** and he told her exactly what had happened the night when he was allegedly assaulted and shot by a woman.	**Mean**	1.87	2.77	2.77
	Standard Deviation	1.36	1.47	1.47
28. We're planning to bring the community together in applying pressure to Iran through sanction adopted by the UN that will be aimed at those enterprises, **that * believe** are supplanting Iran's government.	**Mean**	1.25	3.22	2.90
	Standard Deviation	.82	1.31	1.44
29. This is just warming up and unless it gets tougher, then it's clear **that * who are** running against Hillary Clinton merely want to pursue her to select them as her vice-presidential candidate.	**Mean**	1.12	2.66	3.18
	Standard Deviation	.59	1.49	1.33
30. For all the thought Simon Thiel has given to how to combat aging, he actually points out **that * does not have** a lot of specific ideas about what he would do if he could live significantly longer.	**Mean**	2.37	2.55	2.72
	Standard Deviation	1.49	1.49	1.47
31. Lisa Cohen, a former model, persuaded a New York judge to require Google to reveal the identity of an anonymous blogger **that * felt had defamed her**, and she has filed a suit against him.	**Mean**	3.37	3.22	3.31
	Standard Deviation	1.49	1.31	1.25
32. But what do people relish in a Frank Underwood? A guy who gets things done right away and there is a certain deliciousness to it, **that * see in his journey**, but they don't see in the real Washington right now.	**Mean**	1.5	2.33	2.63
	Standard Deviation	1.11	1.49	1.49
33. She recalled the many obstacles that she faced along the way, and how her love of Filipino food, **that * used to be ashamed of** in grade school, sustained her until her efforts were rewarded with success.	**Mean**	2.12	2.66	2.77
	Standard Deviation	1.45	1.49	1.47
34. Let's take a look over at Janet Reno today, who had a public briefing, or what passes for a briefing at the	**Mean**	2.62	2.0	3.18
	Standard Deviation	1.49	1.41	1.33

Department of Justice **that * says some interesting things** about this report.				
35. American author, Kelly Link loves fiction **that * says** "takes things which are comfortable and familiar and makes them really strange, or else... takes things which are strange and impossible and makes them feel comfortable.	**Mean**	2.00	2.55	2.90
	Standard Deviation	1.41	1.49	1.44
36. When my husband's home and may help with some of the chores around, he gets a little bit irritated if I happen to forget to thank him or compliment him **that * did the chore** in a very nice way.	**Mean**	2.12	2.66	2.22
	Standard Deviation	1.49	.94	1.47

Appendix F **Results of the chi-square analysis of the filler sentences of the GJT**

Results in percentages per group of participants for the Grammatical Filler Sentences with a Relative Pronoun (GRL).

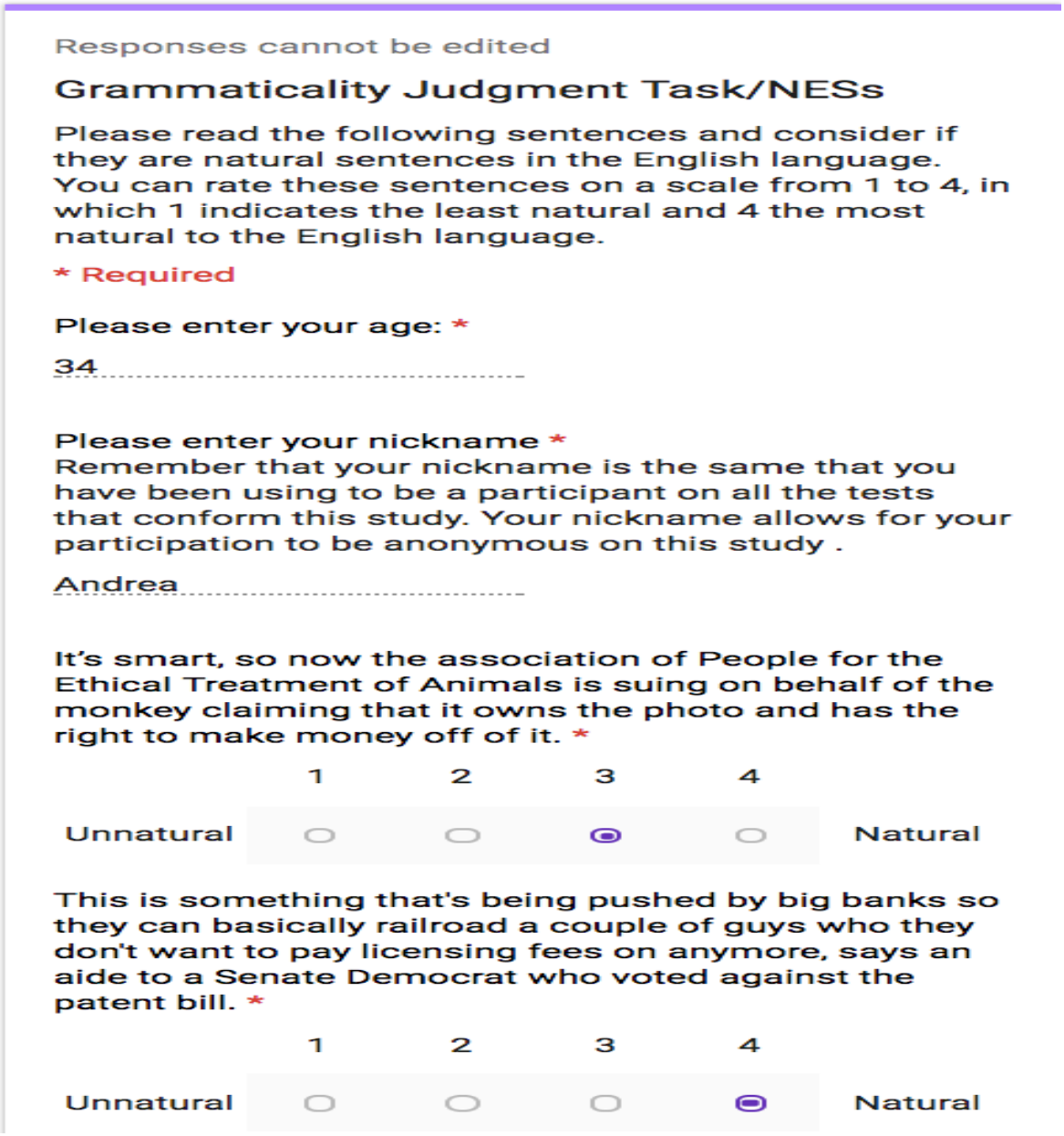
Categorical response to GRL		**Intermediate**	**Advanced**	**Native Speakers**
Reject	Percentage	22.7%	25.9%	4.2%
	Count	5	7	1
No Tendency	Percentage	9.1%	3.7%	8.3%
	Count	2	2	8
Accept	Percentage	68.2%	70.4%	87.5%
	Count	15	19	21
	Total no. of participants per group	22	27	24

Results in percentages per group of participants for the Ungrammatical Filler Sentences with a Relative Pronoun (UGRL).

Categorical response to UGRL		**Intermediate**	**Advanced**	**Native Speakers**
Reject	Percentage	9.1%	33.3%	79.2%
	Count	2	9	19
No Tendency	Percentage	9.1%	14.8%	12.5%
	Count	2	4	3
Accept	Percentage	81.8%	51.9%	8.3%
	Count	18	14	2
	Total no. of participants per group	22	27	24

Appendix G **Screen sample of the GJT in English**

Google Form link to test: https://goo.gl/forms/54LHrKsFK4pvpUWu1



Responses cannot be edited

Grammaticality Judgment Task/NESs

Please read the following sentences and consider if they are natural sentences in the English language. You can rate these sentences on a scale from 1 to 4, in which 1 indicates the least natural and 4 the most natural to the English language.

* Required

Please enter your age: *

34

Please enter your nickname *
Remember that your nickname is the same that you have been using to be a participant on all the tests that conform this study. Your nickname allows for your participation to be anonymous on this study .

Andrea

It's smart, so now the association of People for the Ethical Treatment of Animals is suing on behalf of the monkey claiming that it owns the photo and has the right to make money off of it. *

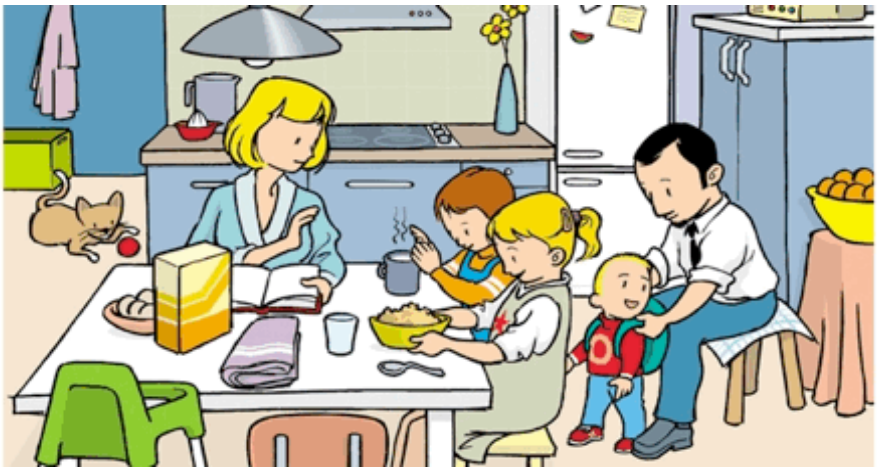
	1	2	3	4	
Unnatural	○	○	◉	○	Natural

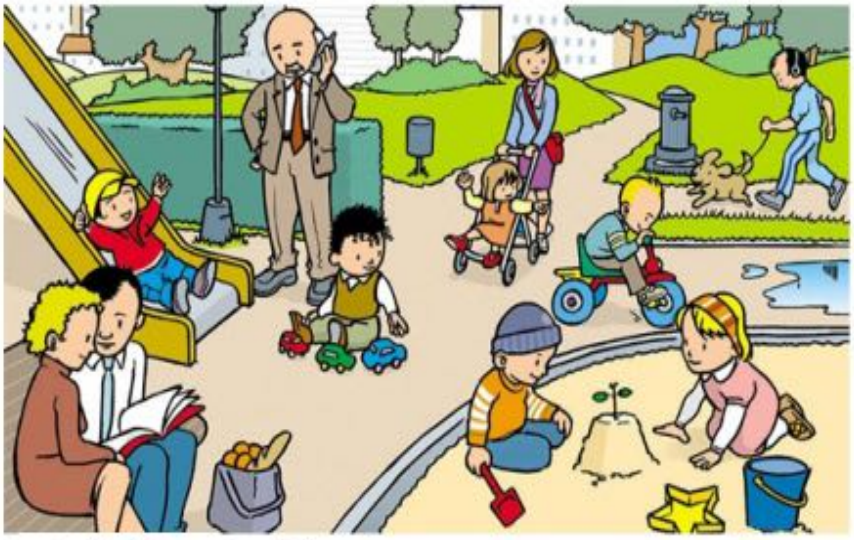
This is something that's being pushed by big banks so they can basically railroad a couple of guys who they don't want to pay licensing fees on anymore, says an aide to a Senate Democrat who voted against the patent bill. *

	1	2	3	4	
Unnatural	○	○	○	◉	Natural

Appendix H **Drawings for SGT**

Drawings retrieved from: http://www.apic.es/imágenes





Fuente: http://www.apic.es/imágenes

Appendix I Speech generation task

I.1 Instructions for SGT

1. You will look at two pictures.
2. You will decide what picture you prefer.
3. You will be asked to describe and make comments about it for 30 seconds.
4. Your description will be recorded.
5. Please, give as much information of the picture as you can. Detail your description and comments.

I.2 Instrucciones para la actividad de generación de discurso

1. Usted observará dos dibujos.
2. Por favor decida cual dibujo prefiere.
3. Se le va a pedir que describa y haga comentarios acerca del dibujo for 30 segundos.
4. Su descripción será grabada.
5. Por favor, provea tanta información como le sea posible acerca del dibujo. De detalles de su descripción y de sus comentarios.

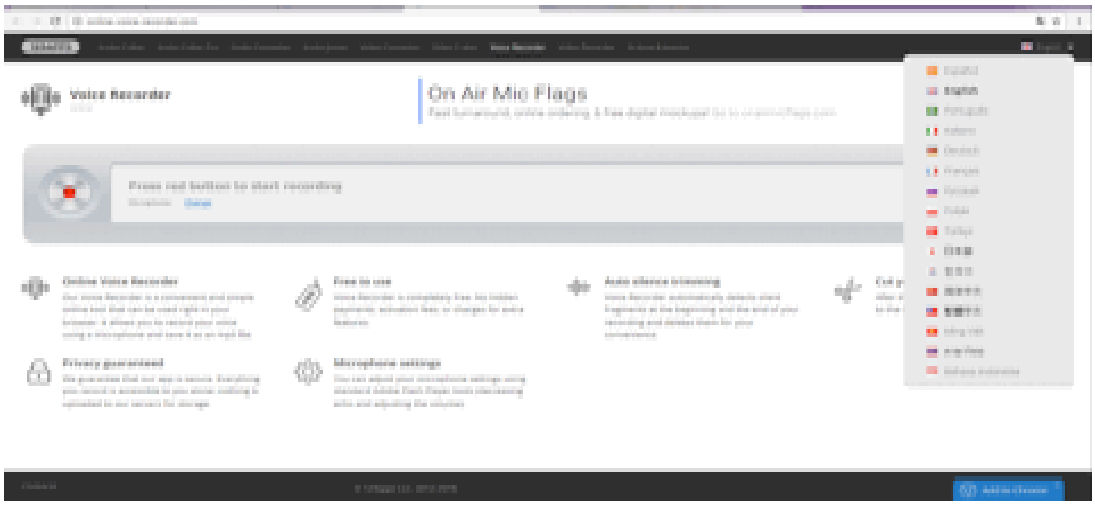
I.3 Intructions to complete the speech generation task (e-mail format)

There are 9 steps for completion of this task. Please make sure that you read all task instructions carefully before starting this task.

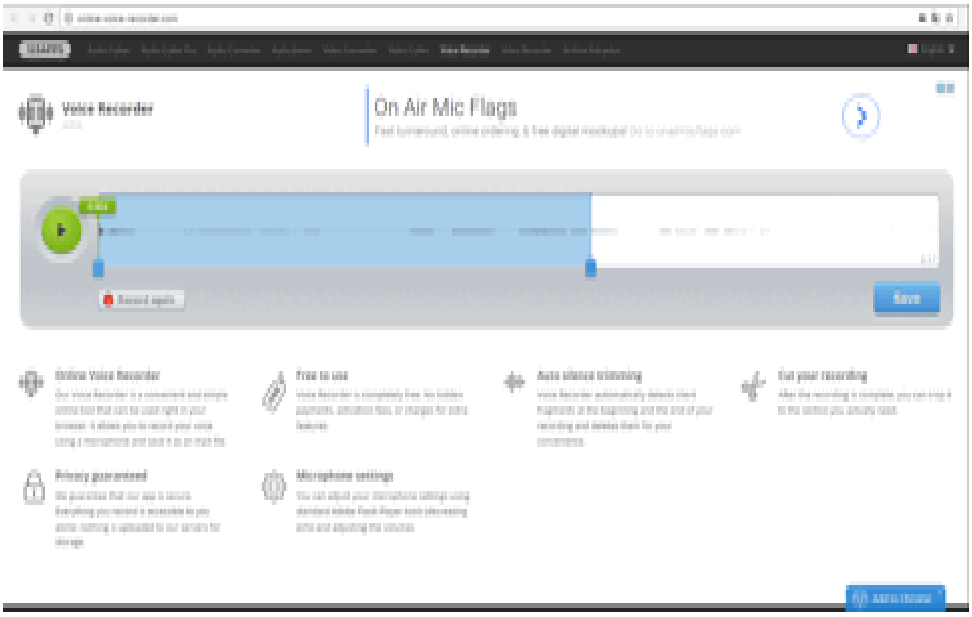
Steps to complete the task:

1. Open the next link, which will lead you to an online recorder named **Voice Recorder.** http://online-voice-recorder.com/

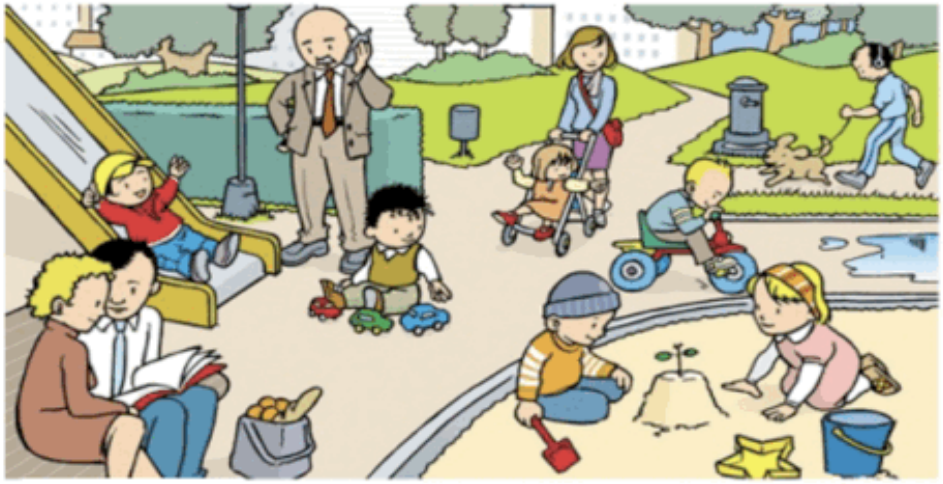
(If the link does not direct you to the recording website, copy the address on your internet browser. There is a copy of this link on the e-mail containing all links to tasks and instructions.) (For language settings, go to the right side of the page.)



2. Try creating a short recording before taking the actual test.



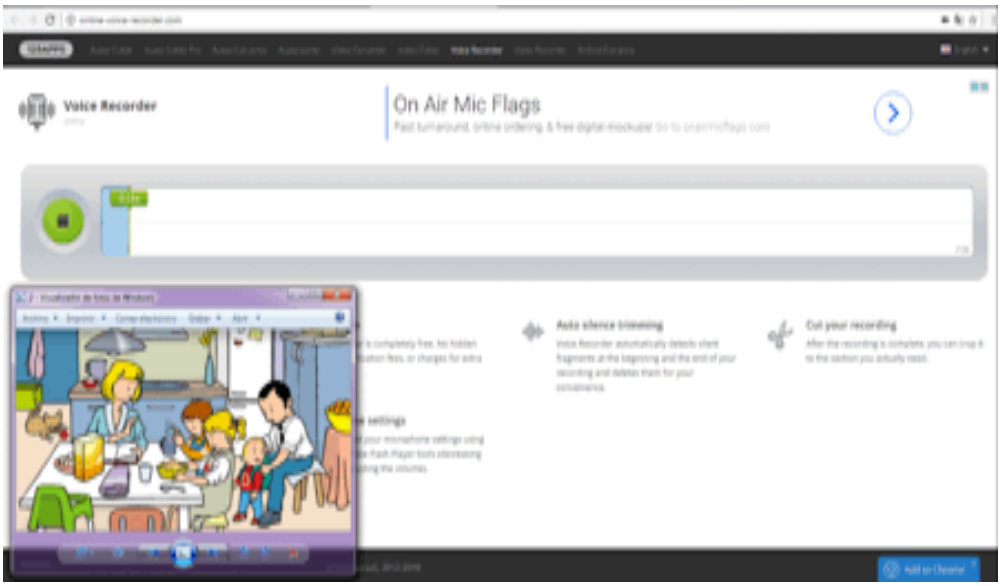
3. Press the save button and save your recording on a folder that is easy to locate in your computer device.
4. Now that you have done the trial, please observe the following image below (take from 20 to 40 seconds to look at all the elements and details therein).



Fuente: http://www.apic.es/imágenes

5. After this, keep the picture at hand and visible as shown below as you prepare to record yourself on the online voice recorder that you have opened on your internet engine.
(There is attachment with the image on the e-mail where you have received all links to tasks and instructions. The attachment is under the name “Image for Speech Generation Task”. This attachment is available for a better access and visibility of the image while recording yourself.)
6. Press the recording button on the online recorder and start describing what you see in the picture for around 40 seconds. Give as much information of the picture as you can.

Appendix I



7. Stop the recording when you have completed around 40 seconds of description of the image above and save it as soon you as you finish.
8. Please save the recording with the nickname that you have chosen to be a participant for this study (remember that this is the same nickname that you will be using for all tests in this study).
9. Once you have saved the recording with your chosen nickname, email it to any of the following email addresses.

luisacejas7@gmail.com
brendagp14@gmail.com
jacquechavira@gmail.com

(emailing your recordings to any of these addresses guarantees that any type of data that you produce for this study remains anonymous and risk-free for the main researcher to link it with your personal information)

Please feel free to contact any of the staff members on the e-mail addresses provided above in case you have any questions on how to complete this task.

Thank you so much for your cooperation!
Please continue taking the rest of the tests for this study.

Appendix J Covariance correlational analysis on the Spanish and English GJTs

The following is a report of the Pearson correlation analyses in SPSS performed to estimate of there was a transfer effect on the responses of the Intermediate and Advanced groups of participants in the GJT. The correlations were made using the participants' responses to the sentences in the GJT in their L1 (Spanish) and their L2 (English).

In order to perform this analysis, the data converted to percentages of acceptance was considered. The results of the covariances are organized by groups of participants and then per sentence condition. (Significant correlations will be highlighted in yellow and marked with an asterisk).

J.1 Intermediate group:

J.1.1 Correlations between the sentences with an Object Resumptive (+RO) Condition:

Correlations			
		English +RO	Spanish +RO
English +OR	Pearson Correlation	1	.317
	Sig. (bilateral)		.151
	N	22	22
Spanish +OR	Pearson Correlation	.317	1
	Sig. (bilateral)	.151	
	N	22	22

J.1.2 Correlations between the sentences without an Object Resumptive (-RO) condition:

Correlations			
		English -RO	Spanish -RO
English -OR	Pearson Correlation	1	.070
	Sig. (bilateral)		.758
	N	22	22
Spanish -OR	Pearson Correlation	.070	1
	Sig. (bilateral)	.758	
	N	22	22

J.1.3 Correlations between the grammatical fillers (GRL)

Correlations			
		English GRL	Spanish GRL
English GRL	Pearson Correlation	1	.558**
	Sig. (bilateral)		.007
	N	22	22
Spanish GRL	Pearson Correlation	.558**	1
	Sig. (bilateral)	.007	
	N	22	22

J.1.4 Correlations between the ungrammatical filler (UGRL):

Correlations			
		English UGRL	Spanish UGRL
English UGRL	Pearson Correlation	1	.655**
	Sig. (bilateral)		.001
	N	22	22
Spanish UGRL	Pearson Correlation	.655**	1
	Sig. (bilateral)	.001	
	N	22	22

J.2 Advanced Group:

J.2.1 Correlations between the sentences with an Object Resumption (+RO) condition:

Correlations			
		English +RO	Spanish +RO
English +RO	Pearson Correlation	1	.246
	Sig. (bilateral)		.215
	N	27	27
Spanish +RO	Pearson Correlation	.246	1
	Sig. (bilateral)	.215	
	N	27	27

J.2.2 Correlations between the sentences without an Object Resumption (-RO) condition:

Correlations			
		English -RO	Spanish -RO
English -RO	Pearson Correlation	1	.375
	Sig. (bilateral)		.054
	N	27	27
Spanish -RO	Pearson Correlation	.375	1
	Sig. (bilateral)	.054	
	N	27	27

J.2.3 Correlations between the filler sentences with a grammatical relativization (GRL) condition:

Correlations			
		English GRL	Spanish GRL
English GRL	Pearson Correlation	1	.260
	Sig. (bilateral)		.191
	N	27	27
Spanish GRL	Pearson Correlation	.260	1
	Sig. (bilateral)	.191	
	N	27	27

J.2.4 Correlations between the ungrammatical filler sentences (UGRL):

Correlations

		English UGRL	Spanish UGRL
English UGRL	Pearson	1	.390*
	Sig. (bilateral)		.044
	N	27	27
Spanish UGRL	Pearson	.390*	1
	Sig. (bilateral)	.044	
	N	27	27

Appendix K Non-significant findings in the multiple regression analysis

The following is a list of the tables the show the non-significant regression equations in the multiple regression analysis performed in SPSS to calculate the effects of WMC on L2 oral fluency (Study 1) and the acceptability of object resumptive pronouns (Study 2). The tables are arranged by study and then per group of participants.

K.1 Study 1

K.1.1 Group of Advanced L2Aers

Summary of the Model

Model	R	R square	R square adjusted	Standard error of estimation	Durbin-Watson
1	.225[a]	.050	-.032	.72949	.513

a. Predictors: (Constant), Sum of Positive Measures of Fluency and Sum of Negative Measures of Fluency
b. Dependent variable: WMC Total Measure

ANOVA[a]

Model		Sum of squares	gl	Square Medium	F	Sig.
1	Regression	.650	2	.325	.611	.551[b]
	Residue	12.240	23	.532		
	Total	12.890	25			

a. Dependent variable: WMC Total Measures
b. Predictors: (Constant), Sum of Positive Measures of Fluency and Sum of Negative Measures of Fluency

Coeficients[a]

Model		Non standardized coefficients		Standardized coefficients	t	Sig.	Statistics of Collinearity	
		B	Standard Error	Beta			Tolerance	VIF
1	(Constante)	3.728	2.063		1.807	.084		
	Sum of positive measures of fluency	.035	.764	.009	.046	.964	1.000	1.000
	Sum of negative measures of oral fluency	-.249	.226	-.224	-1.103	.281	1.000	1.000

a. Dependent variable: WMC Total Measure

K.1.2 Group of Native Speakers

Summary of the Model

Model	R	R square	R square adjusted	Standard error of estimation	Durbin-Watson
1	.402[a]	.161	.081	.56922	.505

a. Predictors: (Constant), Sum of Positive Measures of Fluency and Sum of Negative Measures of Fluency
b. Dependent variable: WMC Total Measures

ANOVA[a]						
Model		Sum of squares	gl	Square Medium	F	Sig.
1	Regression	1.308	2	.654	2.019	.158[b]
	Residue	6.804	21	.324		
	Total	8.112	23			
a. Dependent Variable: WMC Total Measures						
b. Predictors: (Constant), Sum of Positive Measures of Fluency and Sum of Negative Measures of Fluency						

Coefficients								
Model		Non-standardized coefficients		Standardized Coefficients	t	Sig.	Statistics of Collinearity	
		B	Standard Error	Beta			Tolerance	VIF
1	(Constant)	7.021	1.441		4.873	.000		
	Sum of Positive Measures of Fluency	-.820	.462	-.426	-1.774	.091	.692	1.445
	Sum of Negative Measures of Fluency	-.440	.249	-.425	-1.770	.091	.692	1.445
a. Dependent Variable: WMC Total Measures								

K.2 Study 2

K.2.1 Group of Advanced L2Aers

Summary of the Model					
Model	R	R square	R square adjusted	Standard error of estimation	Durbin-Watson
1	.629[a]	.395	.214	1.87842	.934
a. Predictors: (Constant), UGRL, GRL, +RO, -RO					
b. Dependent Variable: WMC Total Measures					

ANOVA[a]						
Model		Sum of squares	gl	Square medium	F	Sig.
1	Regression	46.142	6	7.690	2.180	.089[b]
	Residue	70.569	20	3.528		
	Total	116.711	26			
a. Dependent Variable: WMC						
b. Predictors: (Constant), UGRL, GRL, +RO, -RO						

Coefficients								
Model		Non-standardized coefficients		Standardized Coefficients	t	Sig.	Statistics of Collinearity	
		B	Standard Error	Beta			Tolerance	VIF
1	(Constant)	-.425	3.708		-.115	.910		
	+RO	.617	.330	.608	1.866	.077	.285	3.506

	-RO	-.766	.369	-.737	-2.078	.051	.241	4.157
	GRL	.995	.353	.556	2.824	.010	.778	1.285
	UGRL	.039	.288	.030	.136	.893	.631	1.584
a. Dependent variable: WMC total measure								

WMC Study 2 Advanced L2Aers Group			
Variables	**Betas**	**Effect of Independent Variable**	**Effect of Dependent Variable**
(Constant)	-0.425	Non-significant, it does not affect the variable.	
B2 +RO	0.617	Non-significant, it does not affect the variable.	
B4 -RO	-0.766	If decreased by 1 % B4	It will decrease by 0.766 % in relation to WMC
B5 GRL	0.995	If incremented by 1 % B5	It will increase by 0.995 % in relation to WMC
B6 UGRL	0.039	Non-significant, it does not affect the variable.	

a. Dependent variable: WMC Total Measure

K.2.2 Group of Native Speakers

Summary of the Model					
Model	R	R square	R square adjusted	Standard error of estimation	Durbin-Watson
1	.663[a]	.439	.242	1.91419	1.640
a. Predictors: (Constant) GRL, UGRL, +RO, -RO					

ANOVA[a]

Model		Sum of squares	gl	Square medium	F	Sig.
1	Regression	48.835	6	8.139	2.221	.091[b]
	Residue	62.290	17	3.664		
	Total	111.125	23			

a. Dependent Variable: WMC Total Measures

Tabla 9 Coeficientes[a]								
		Non-Standardized Coefficients		Standardized Coefficients			Statistics of Collinearity	
Model		B	Standard Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	4.853	3.916		1.239	.232		
	+RO	.438	.251	.621	1.743	.099	.260	3.850
	-RO	-.275	.298	-.318	-.923	.369	.279	3.588
	GRL	.408	.485	.206	.840	.413	.549	1.822
	UGRL	-.889	.348	-.771	-2.555	.020	.362	2.760
a. Dependent Variable: WMC Total Measures								

WMC Study 2 Native Speakers

Appendix K

Variables	Betas	Effect of Independent Variable	Effect of Dependent Variable
(Constant)	4.853		
B2+RO	0.438	Non-significant, it does not affect the variable.	
B4 -RO	-0.275	Non-significant, it does not affect the variable.	
B5 GRL	0.408	Non-significant, it does not affect the variable.	
B6 UGRL	-0.889	If increased by 1 % B6	It will decrease by 0.889 % in terms of WMC

a. Dependent Variable: WMC Total Measures

List of references

Alexopoulou, T., & Keller, F. (2007). Locality, cyclicity, and resumption: At the interface between the grammar and the human sentence processor. *Language*, 110-160.

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in cognitive sciences*, *4*(11), 417-423.

Baddeley, A. (2003a). Working memory and language: An overview. *Journal of communication disorders*, *36*(3), 189-208.

Baddeley, A. (2003b). Working memory: looking back and looking forward. *Nature reviews neuroscience*, *4*(10).

Baddeley, A. (2007). *Working memory, thought, and action* (Vol. 45). OUP Oxford.

Baddeley, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 311-324.

Baddeley, A. D. (2017). Modularity, working memory and language acquisition. *Second Language Research*, 0267658317709852.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of learning and motivation*, *8*, 47-89.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*(5), 769-786.

Cowan, N. (2014). Working memory underpins cognitive development, learning, and education. *Educational psychology review*, *26*(2), 197-223.

Cowan, N. (2015). Second-language use, theories of working memory, and the Vennian mind. *Working memory in second language acquisition and processing*, 29-40.

Cowan, N. (2016). *Working memory capacity: Classic edition*. Psychology press.

Daneman, M., & Carpenter, P. A. (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behaviour*, *19*(4), 450-66.

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

de Jong, N. H. (2018). Fluency in second language testing: Insights from different disciplines. *Language Assessment Quarterly*, *15*(3), 237-254.

de Jong, N. H., & Wempe, T. (2009a). Praat script to detect syllable nuclei and measure speech rate automatically. *Behavior research methods*, *41*(2), 385-390.

DeKeyser, R. (2001). Automaticity and automatization. In Robinson, P., editor, Cognition and second language instruction, Cambridge: Cambridge University Press, 125-151.

DeKeyser, R. M. (1997). Beyond explicit rule learning. *Studies in second language acquisition*, *19*(2), 195-221.

Dussias, P. E., & Piñar, P. (2010). Effects of reading span and plausibility in the reanalysis of wh-gaps by Chinese-English second language speakers. *Second Language Research*, *26*(4), 443-472.

Ellis, N. (1999). Cognitive approaches to SLA. *Annual Review of Applied Linguistics*, *19*, 22-42.

Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language learning*, *48*(4), 631-664.

Ellis, N. C. (2007). Dynamic systems and SLA: The wood and the trees. *Bilingualism: Language and Cognition*, *10*(01), 23-25.

Ellis, N. C. (2008). The dynamics of second language emergence: Cycles of language use, language change, and language acquisition. *The Modern Language Journal*, *92*(2), 232-249.

Ellis, N. C. (2009). Constructing a Second Language: Learning as a Function of Frequency, Frequency Distribution, Form & Function. *Cognitive Approaches to Second/Foreign Language Processing: Theory and Pedagogy*, 24.

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology: Section A*, *49*(1), 234-250.

Fortkamp, M. B. M. (1999). Working memory capacity and aspects of L2 speech production. *Communication & cognition. Monographies*, *32*(3-4), 259-296.

Gathercole, S. E., Dunning, D. L., Holmes, J., & Norris, D. (2019). Working memory training involves learning new skills. *Journal of memory and language*, *105*, 19-42.

Hawkins, R. (2007). The nativist perspective on second language acquisition. *Lingua*, *118*(4), 465-477.

Hawkins, R., & Hattori, H. (2006). Interpretation of English multiple wh-questions by Japanese speakers: A missing uninterpretable feature account. *Second Language Research*, *22*(3), 269-301.

Juffs, A. (2004). Representation, processing and working memory in a second language. *Transactions of the Philological Society*, *102*(2), 199-225.

Juffs, A. (2005). The influence of first language on the processing of wh-movement in English as a second language. *Second Language Research*, *21*(2), 121-151.

Juffs, A., & Harrington, M. (1995). Parsing effects in second language sentence processing: Subject and object asymmetries in wh-extraction. *Studies in second language acquisition*, *17*(4), 483-516.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, *44*(02), 137-166.

Juffs, A., & Rodríguez, G. A. (2014). *Second language sentence processing*. Routledge.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological review*, *99*(1), 122.

Kormos, J. (2006). *Speech production and second language acquisition*. Routledge.

Larson-Hall, J. (2016). *A guide to doing statistics in second language research using SPSS and R*. Routledge.

Leal-Méndez, T. L., & Slabakova, R. (2012). The Interpretability Hypothesis again: A partial replication of Tsimpli and Dimitrakopoulou (2007). *International Journal of Bilingualism*

Levelt, W. (1989). Speaking: From intention to articulation. Cambridge, MA: MIT Press.

Levelt, W. J. (1999). Producing spoken language: A blueprint of the speaker. In The neurocognition of language (pp. 83-122). Oxford University Press.

Mackey, A., & Gass, S. M. (2012). Research methods in second language acquisition.

MacWhinney, B. (2006). Emergentism—use often and with care. *Applied Linguistics*, *27*(4), 729-740.

McLaughlin, B., & Heredia, R. (1996). Information-processing approaches to research on second language acquisition and use. *Handbook of second language acquisition*, 213-228.

Rothman, J., & Slabakova, R. (2017). The state of the science in generative SLA and its place in modern second language studies. *Studies in Second Language Acquisition*.

Rouveret, A. (2011). Some issues in the theory of resumption: A perspective on early and recent research. *Resumptive pronouns at the interfaces*, 1-64.

Schmid, M. S. (2011). The theoretical significance of research on language attrition for understanding bilingualism

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in second language acquisition*, *14*(4), 357-385.

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. *Handbook of bilingualism: Psycholinguistic approaches*, 371-388.

Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, *14*(3), 369-385.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.

Segalowitz, N. (2016). Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 79-95.

Skehan, P. (2003). Task-based instruction. *Language teaching*, *36*(1), 1-14.

Temple, L. (2000). Second language learner speech production. *Studia linguistica*, *54*(2), 288-297.

Tavakoli, P. (2019) Automaticity, fluency and second language task performance. In: Wen, Z. E. and Ahmadian, M. J. (eds.) Researching L2 Task Performance and Pedagogy. John Benjamins, Amsterdam, pp. 39-52.

Tavakoli, P., Nakatsuhara, F., & Hunter, A. M. (2020). Aspects of fluency across assessed levels of speaking proficiency. *The Modern Language Journal*, *104*(1), 169-191.

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), Planning and task performance in a second language (pp. 239–273). Amsterdam, the Netherlands: John Benjamins.

Tsimpli, I. M., & Dimitrakopoulou, M. (2007). The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research*, *23*(2), 215-242.

Wen, Z. (2015). Working memory in second language acquisition and processing: The phonological/executive model. In: Z. Wen, M. Mota, & A. McNeill (Eds.). Working memory in second language acquisition and processing (pp. 41–62). Bristol: Multilingual Matters.

Wen, Z. E. (2016). *Working memory and second language learning: Towards an integrated approach* (Vol. 100). Multilingual matters.

White, L. (2003). *Second language acquisition and universal grammar*. Cambridge University Press.

White, L. (2007). Linguistic theory, universal grammar, and second language acquisition. *Theories in second language acquisition: An introduction*, 37-55.

White, L., & Juffs, A. (1998). Constraints on wh-movement in two different contexts of nonnative language acquisition: Competence and processing. *The generative study of second language acquisition*, *111129*.

Bibliography

Alexopoulou, T. (2010). Truly intrusive: resumptive pronominals in questions and relative clauses. *Lingua*, *120*(3), 485-505.

Allard, F., & Starkes, J. L. (1991). Motor-skill experts in sports, dance, and other domains. *Toward a general theory of expertise: Prospects and limits*, 126-152.

Allen, R., Baddeley, A., & Hitch, G. (2017). Executive and perceptual distraction in visual working memory. *Journal of Experimental Psychology: Human*

Anderson, J. R. (1992). Automaticity and the ACT theory. *The American journal of psychology*, 165-180.

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

Ariel, M. (1990). Accessing Noun Phrase Antecedents. Routledge, London.

Ariel, M. (1999). Cognitive universals and linguistic conventions: the case of resumptive pronouns. Studies in Language 23 (2), 217–269.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of learning and motivation*, *2*, 89-195.

Baddeley, A. D. (2004). The psychology of memory. *The essential handbook of memory disorders for clinicians*, 1-13.

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2019). From short-term store to multicomponent working memory: The role of the modal model. *Memory & cognition*, *47*(4), 575-588.

Baddeley, A., & Lewis, V. (1984). When does rapid presentation enhance digit span? *Bulletin of the Psychonomic Society*, *22*(5), 403-405.

Baddeley, A., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of memory and language*, *27*(5), 586-595.

Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long-and short-term memory. *Journal of verbal learning and verbal behavior*, *9*(2), 176-189.

Bates, E. (1999). Plasticity, localization and language development. *The changing nervous system: Neurobehavioral consequences of early brain disorders*, 214-253.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and brain Sciences*, *22*(1), 77-94.

Carlson, R., Sullivan, M., & Schneider, W. (1989). Component fluency in a problem-solving context. *Human Factors*, *31*(5), 489-502.

Chomsky, N. (1965). Aspects of the Theory of Syntax. Cambridge: MIT Press.

Chomsky, N. (1995). The Minimalist Program (Current Studies in Linguistics 28). *Cambridge et al*.

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic bulletin & review*, *8*(2), 331-335.

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

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological bulletin*, *104*(2), 163.

Craik, F. I., Eftekhari, E., Bialystok, E., & Anderson, N. D. (2018). Individual differences in executive functions and retrieval efficacy in older adults. *Psychology and aging*, *33*(8), 1105.

Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of verbal learning and verbal behavior*, *11*(6), 671-684.

Creswell, C. (2002). Resumptive pronouns, wh-island violations, and sentence production. *TAG+ 6*, 40.

Cucchiarini, C., Strik, H., & Boves, L. (2002). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *the Journal of the Acoustical Society of America*, *111*(6), 2862-2873.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge.

Daneman, M. (1991). Working memory as a predictor of verbal fluency. *Journal of Psycholinguistic Research*, *20*(6), 445-464.

Davies, M. (2008-) *The Corpus of Contemporary American English (COCA): 520 million words, 1990-present*. Available online at https://corpus.byu.edu/coca/.

de Jong, N. H., Schoonen, R., & Hulstijn, J. H. (2009b). Fluency in L2 is related to fluency in L1. In *7th International Symposium on Bilingualism, Utrecht, The Netherlands*.

de Jong, N. H., & Bosker, H. R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In *The 6th workshop on disfluency in spontaneous speech (diss)* (pp. 17-20).

de Jong, N. H. (2016). Predicting pauses in L1 and L2 speech: The effects of utterance boundaries and word frequency. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 113-132.

de Garavito, J. B. (2006). Knowledge of clitic doubling in Spanish Evidence against pattern learning. *Inquiries in Linguistic Development: In honor of Lydia White*, 305.

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language learning*, *54*(4), 655-679.

Dewaele, J.–M. (1996). How to measure formality of speech? A model of synchronic variation. In K. Sa- javaara & C. Fairweather (Eds.), *Approaches to sec- ond language acquisition* (pp. 119–133).

Eckman, F. R. (1977). Markedness and the contrastive analysis hypothesis. *Language learning*, *27*(2), 315-330.

Ejzenberg, R. (2000). The juggling act of oral fluency: A psycho-sociolinguistic metaphor. In: H. Riggenbach (ed.), Perspectives on Fluency (pp. 287-313). Ann Arbor: University of Michigan Press.

Elliott, D., & Lyons, J. (1998). Optimizing the use of vision during motor skill acquisition. *Motor behavior and human skill: a multidisciplinary approach*, 57-72.

Ellis, N. C., & Larsen-Freeman, D. (2006). Language emergence: Implications for applied linguistics—Introduction to the special issue. *Applied Linguistics*,*27*(4), 558-589.

Ellis, N. C., & Schmidt, R. (1998). Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language and cognitive processes*, *13*(2-3), 307-336.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: a test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 972.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex.

Engle, R.W. (2002). Working memory capacity as executive attention. Current Directions in Psychological Science, 11, 19–23.

Ericsson, K. A., & Delaney, P. F. (1999). Long-term working memory as an alternative to capacity models of working memory in everyday skilled performance.

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first-and second-language reading of fluent bilinguals. *Memory & Cognition*, *11*(6), 565-574.

Fillmore, C. J. (1979). On Fluency. In C. I. Fillmore, D. I. Kempler, W. i. Wang (Eds.) *Individual Differences in Language Ability and Language Behaviour* (pp. 85-101). New York: Academic.

Finardi, K., & Weissheimer, J. (2008). On the relationship between working memory capacity and L2 speech development. *Signótica*, *20*(2), 367-391.

Foster, P., & Skehan, P. (2013). Anticipating a post-task activity: The effects on accuracy, complexity, and fluency of second language performance. *Canadian modern language review*, *69*(3), 249-273.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied linguistics*, *21*(3), 354-375.

Fortkamp, M. B. M., & Bergsleithner, J. M. (2007). Relationship among individual differences in working memory capacity, noticing, and L2 speech production. *Signo*, *32*(52), 40-53.

Freed, B.F. (2000). Is fluency, like beauty, in the eyes (and ears) of the beholder? In H. Riggenbanch (Ed.), Perspectives on Fluency (pp. 243-265). Ann Harbor; University of Michigan Press.

Frost, J. (2017). Multicollinearity in regression analysis: problems, detection, and solutions. *Statistics by Jim. Available online also at: https://statisticsbyjim.*

com/regression/multicollinearity-in-regression-analysis/[accessed in Jakarta, Indonesia: December 22, 2018].

Gatbonton, E., & Segalowitz, N. S. (1988). Crea- tive automatization: Principles for promoting fluency within a communicative framework. TESOL Quarterly, 22, 473–492.

Gatbonton, E., & Segalowitz, N. S. (2005). Rethinking communicative language teaching: A focus on access to fluency. Canadian Modern Language Review.

Gathercole, S. E., & Alloway, T. P. (2007). Understanding working memory: A classroom guide. *Lontoo: Harcourt Assessment*.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1-76.

Goad, H., & White, L. (2004). Ultimate attainment of L2 inflection: Effects of L1 prosodic structure. *EUROSLA yearbook*, *4*(1), 119-145.

Goldman-Eisler, F. (1968). Psycholinguistic experiments in spontaneous speech. London: Academic Press.

Halford, G.S., Wilson, W.H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. Behavioral and Brain Sciences, 21, 723–802.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193-225). San Diego, CA: Academic Press.

Havik, E., Roberts, L., Van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning*, *59*(1), 73-112.

Hawkins, J.A. (2004). Efficiency and Complexity in Grammars. Oxford University Press.

Hawkins, R., & Chan, C. Y. H. (1997). The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. *Second language research*, *13*(3), 187-226.

Hebb, D. O. (1949). *The organization of behaviour: A neuropsychological approach*. John Wiley & Sons.

Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, *9*(5), 220-225.

Hulstijn, J. H. (1990). A Comparison Between the Information-Processing and the Analysis/Control Approaches to Language Learning1. *Applied Linguistics*,*11*(1), 30-45.

Hulstijn, J. H., & De Graaff, R. (1994). Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *Aila Review*, *11*, 97-112.

James, W. (1890). The principles of psychology. New York: Henry Holt & Co.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive psychology*, *21*(1), 60-99.

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of experimental psychology: General*, *130*(2), 169.

Kane, M.J., & Engle, R.W. (2001). *Individual differ- ences in executive attention and the Stroop*. Manu- script submitted for publication.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189.

Kahng, J. (2014). Exploring utterance and cognitive flu- ency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learn- ing*, *64*, 809–854.

Kempe, V., & MacWhinney, B. (1998). The acquisition of case marking by late learners of Russian and German. *Studies in Second Language Acquisition*, *20*(04), 543-587.

Kline, R. B. (2004). Beyond significance testing. Reforming data analysis methods in behavioral research. Washington, DC: American Psychological Association.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, *32*(2), 145-164.

Lahmann, C., Steinkrauss, R., & Schmid, M. S. (2017). Speed, breakdown, and repair: An investigation of fluency in long-term second-language speakers of English. *International Journal of Bilingualism*, *21*(2), 228-242.

Lardiere, D. (1998). Dissociating syntax from morphology in a divergent L2 end-state grammar. *Second language research*, *14*(4), 359-375.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language learning*, *40*(3), 387-417.

Lennon, P. (2000). The Lexical Element in Spoken Second Language Fluency. In H. Riggenbach, R. (. Schmidt (Eds.), *Perspectives on Fluency* (pp. 25-42). Ann Arbor, MI: U of Michigan P.

Linck, J. A., & Weiss, D. J. (2011). Working memory predicts the acquisition of explicit L2 knowledge. *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism*, 101-114.

Linck, J. A., & Weiss, D. J. (2015). Can working memory and inhibitory control predict second language learning in the classroom?. *Sage Open*, *5*(4), 2158244015607352.

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language Learning*, *60*(3), 501-533.

MacWhinney, B. (1986). Competition and teachability.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk. transcription format and programs* (Vol. 1). Psychology Press.

MacWhinney, B. (2001). The competition model: The input, the context, and the brain. *Cognition and second language instruction*, 69-90.

MacWhinney, B. (2004). A unified model of language acquisition. *Handbook of bilingualism*, 49.

MacWhinney, B., Leinbach, J., Taraban, R., & McDonald, J. (1989). Language learning: Cues or rules? *Journal of Memory and language*, *28*(3), 255-277

Manikandan, S. (2011). Measures of central tendency: The mean. *J Pharmacol Pharmacother*, *2*(2), 140-2.

Matessa, M., & Anderson, J. R. (2000). Modelling focused learning in role assignment. *Language and Cognitive Processes*, *15*(3), 263-292.

McCloskey, J. (2006). Resumption. In: Everaet, M., van Riemsdijk, H. (Eds.), The Blackwell Companion to Syntax. Blackwell Publishing, pp. 94– 117.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language learning*, *33*(2), 135-158.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of verbal Learning and verbal Behaviour*, *2*(1), 1-21.

Meisel, J. M. (1997). The acquisition of the syntax of negation in French and German: Contrasting first and second language development. *Second language research*, *13*(3), 227-263.

Miyake, A. (2001). Individual differences in working memory: Introduction to the special section. *Journal of Experimental Psychology: General*, *130*(2), 163.

Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. *Foreign language learning: Psycholinguistic studies on training and retention*, 339-364.

Mizera, G. J. (2006). *Working memory and L2 oral fluency* (Doctoral dissertation, University of Pittsburgh).

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly journal of experimental psychology*, *11*(1), 56-60.

O'keeffe, A., McCarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge University Press.

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural computation*, *18*(2), 283-328.

O'Grady, W. (2007). The emergentist program. Lingua, 118(Current emergentist and nativist perspectives on second language acquisition), 447-464. doi: 10.1016/j.lingua.2006.12.001

Omaggio-Hadley, A. (2000). Teaching Language in Context (3rd Ed.) Thomson-Heinle

Oswald, F. L., McAbee, S. T., Redick, T. S., & Hambrick, D. Z. (2014). The development of a short domain-general measure of working memory capacity. *Behaviour research methods*, *47*(4), 1343-1355.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language*, *30*(3), 331-347. *Perception and Performance*, *43*(9), 1677-1693

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In Ellis, N. (Ed.), Implicit and explicit learning of languages (pp. 393–419). London: Academic Press.

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of experimental psychology*, *58*(3), 193.

Prebianca, G. V. V. (2009). Communication strategies and proficiency levels in L2 speech production: A systematic relationship. *Revista de Estudos da Linguagem*, *17*(1), 7-50.

Prebianca, G. V. V., Finardi, K. R., & Weissheimer, J. (2014). Working memory capacity across L2 speech proficiency levels. *Revista Brasileira de Linguística Aplicada*, *14*(2), 441-462.

Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second language research*, *16*(2), 103-133.

Prince, E. F. (1990, August). Syntax and discourse: A look at resumptive pronouns. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 16, No. 1, pp. 482-497).

REAL ACADEMIA ESPAÑOLA: Banco de datos (CREA) [online]. *Corpus de referencia del español actual.* <http://www.rae.es> [September 2016]

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*.

Rehbein, J. (1987). On fluency in second language speech. In H. W. Dechert, M. Raupach, H. W. Dechert, M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 97-105). Westport, CT, US: Ablex Publishing.

Roberts, B., & Kirsner, K. (2000). Temporal cycles in speech production. *Language & Cognitive Processes*, *15*(2), 129-157. doi:10.1080/016909600386075

Roberts, R., & Gibson, E. (2002). Individual differences in sentence memory. *Journal of psycholinguistic research*, *31*(6), 573.

Ross, J. R. (1967). Constraints on Variables in Syntax. Ph. D. thesis, MIT. Published as (Ross 1986).

Rossiter, M. J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *Canadian Modern Language Review*, *65*(3), 395-412.

Rumelhart, D. E. & McClelland, J. L. (1986). On learning the past tense of English verbs. In D.E. Rumelhart, J.L. McClelland & the PDP Research Group. Parallel Distributed Processing, Volume II (pp. 216-71). Cambridge, MA: MIT Press.

Ruxton, G. D., & Beauchamp, G. (2008). Time for some a priori thinking about post hoc testing. *Behavioral ecology*, *19*(3), 690-693.

Sagarra, N. (2017). Longitudinal effects of working memory on L2 grammar and reading abilities. *Second Language Research*, *33*(3), 341-363.

Sajavaara, K. (1987). Second Language Speech Production: Factors Affecting Fluency. In H. W. Dechert, M. Raupach (Eds.), *Psycholinguistic Models of Production* (pp. 45-65). Norwood, NJ: Ablex.

Santamaria, K., & Sunderman, G. (2015). 12 Working Memory in Processing Instruction: The Acquisition of L2 French Clitics. *Working memory in second language acquisition and processing*, *87*, 205.

Sanz, C., Lin, H. J., Lado, B., Stafford, C. A., & Bowden, H. W. (2014). One size fits all? Learning conditions and working memory capacity in ab initio language development. *Applied linguistics*, *37*(5), 669-692.

Schneider, W. X. (1999). Visual-spatial working memory, attention, and scene representation: A neuro-cognitive theory. *Psychological Research*, *62*(2-3), 220-236.

Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in second language acquisition*, *26*(2), 173-199.

Serafini, E. J., & Sanz, C. (2016). Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Studies in Second Language Acquisition*, *38*(4), 607-646.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of experimental psychology: General*, *125*(1), 4.

Shallice, T., & Warrington, E. K. (1970). Independent functioning of verbal memory stores: A neuropsychological study. *The Quarterly journal of experimental psychology*, *22*(2), 261-273.

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

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied linguistics*, *30*(4), 510-532.

Skehan, P. (2014). The context for researching a processing perspective on task performance. *Processing perspectives on task performance*, *5*, 1-26.

Skehan, P. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. *Cognitive individual differences in second language processing and acquisition*, 17-40.

Skehan, P., Xiaoyue, B., Qian, L., & Wang, Z. (2012). The task is not enough: Processing approaches to task-based performance. *Language Teaching Research*, *16*(2), 170-187.

Skehan, P., & Foster, P. (2008). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research. *Complexity, accuracy, and fluency in second language use, learning, and teaching*, *207*, 226.

Squire, L. R., & Knowlton, B. J. (2000). The medial temporal lobe, the hippocampus, and the memory systems of the brain. In M. S. Gaz- zaniga (Ed.), The new cognitive neurosciences (pp. 765–779). Cambridge, MA: MIT Press.

Smith, N. V., & Tsimpli, I. M. (1995). *The mind of a savant: Language learning and modularity*. Blackwell Publishing.

Taube-Schiffnorman, M., & Segalowitz, N. (2005). Within-language attention control in second language processing. *Bilingualism*, *8*(3), 195.

Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. *ELT journal*, *65*(1), 71-79.

Tavakoli, P. (2016). Fluency in monologic and dialogic task performance: Challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 133-150.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied linguistics*, *17*(1), 84-119.

Towse, J. N., Hitch, G. J., & Hutton, U. (1998). A reevaluation of working memory capacity in children. *Journal of Memory and Language, 39,* 195-217.

Towse, J. N., Hitch, G. J., & Hutton, U. (2000). On the interpretation of working memory span in adults. *Memory & Cognition, 28,* 341-348. Turner, M., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28,* 127-154.

Tsimpli, I. M., & Mastropavlou, M. (2007). Feature interpretability in L2 acquisition and SLI: Greek clitics and determiners. *The role of formal features in second language acquisition*, 143-183.

Tsimpli, I. M., & Roussou, A. (1991). Parameter-resetting in L2. *UCL Working Papers in Linguistics*, *3*, 149.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of memory and language*, *28*(2), 127-154.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, *5*(2), 207-232.

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. Bilingual- ism: Language and Cognition, 4, 105–122.

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1), 231-270.

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behaviour research methods*, *37*(3), 498-505.

Vallar, G., & Baddeley, A. (1987). Phonological short-term store and sentence processing. *Cognitive Neuropsychology*, *4*(4), 417-438.

Van Lier, L. (2000). 11 From input to affordance: Social-interactive learning from an ecological perspective. *Sociocultural theory and second language learning*, 245.

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Greenwood Publishing Group.

VanPatten, B., & Borst, S. (2012). The roles of explicit information and grammatical sensitivity in the processing of clitic direct object pronouns and word order in Spanish L2. *Hispania*, 270-284.

VanPatten, B., & Cadierno, T. (1993). Input processing and second language acquisition: A role for instruction. *The Modern Language Journal*, *77*(1), 45-57.

Witton-Davies, G. (2010). The role of repair in oral fluency and disfluency. In *Selected papers from the 19th international symposium on English teaching* (pp. 119-129).

Wright, C., & Tavakoli, P. (2016). New directions and developments in defining, analyzing and measuring L2 speech fluency. *International Review of Applied Linguistics in Language Teaching*, *54*(2), 73-77.