The Interaction of Individual Working Memory Capacity with Cognitive Linguistics-Based and Translation-Based Instructional Treatments During the Acquisition of Polysemous L2 Spanish Spatial Prepositions

Joseph F. LeTexier

Follow this and additional works at: https://scholarcommons.sc.edu/etd

Part of the Linguistics Commons

Recommended Citation

This Open Access Dissertation is brought to you by Scholar Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact dillarda@mailbox.sc.edu.
THE INTERACTION OF INDIVIDUAL WORKING MEMORY CAPACITY WITH COGNITIVE LINGUISTICS-BASED AND TRANSLATION-BASED INSTRUCTIONAL TREATMENTS DURING THE ACQUISITION OF POLYSEMOUS L2 SPANISH SPATIAL PREPOSITIONS

by

Joseph F. LeTexier

Bachelor of Arts
Crown College, 1992

Master of Arts
Columbia International University, 1996

Master of Divinity
Columbia International University, 2004

Submitted in Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy in
Linguistics
College of Arts and Sciences
University of South Carolina
2019

Accepted by:

Nina Moreno, Major Professor
Paul Malovrh, Committee Member
Mila Tasseva-Kurktchieva, Committee Member
W. Lindsay Hislop, Committee Member
Cheryl L. Addy, Vice Provost and Dean of the Graduate School
DEDICATION

I dedicate this research project to my wife, Susana. I love you very much.

Thanks for introducing me to the Spanish language. And to my four children, Lucas, Marcus, Hannah and Paul. Thanks for being patient with your dad as I worked long and hard on this project. I love you all very much.
ACKNOWLEDGEMENTS

There were many individuals who supported me in the long endeavor of writing this dissertation. I want to first thank my loving wife Susana and my four kids, Lucas, Marcus, Hannah and Paul for supporting me over the past nine years as I worked long hours on research. I want to thank my dissertation director, Dr. Nina Moreno, for all her wise counsel and encouragement. I also want to thank each of my dissertation committee members, Dr. Mila Tasseva-Kurtchieva, Dr. Paul Malovrh and Dr. W. Lindsay Hislop for their insightful feedback at key moments throughout my dissertation. I want to extend tremendous gratitude to Matt Decker for his help in running my data through SPSS software and for his assistance in interpreting the statistics. I would like to thank Columbia International University for supporting my studies financially. I spent over a year gathering more than 100 participants for the present research, and to each one who participated in any way, I want to give a huge thank you. Finally, I want to thank my Lord and Savior, Jesus Christ, for sustaining me through to the end of this long academic journey. May the education that I have received bring you glory.
The present study investigated three areas in SLA related to the acquisition of polysemous L2 Spanish spatial prepositions. These three areas were (1) the effect of instructional method on the acquisition of productive knowledge of polysemous L2 Spanish spatial prepositions, (2) the effect of working memory capacity on the acquisition of productive knowledge of polysemous L2 Spanish spatial prepositions, and (3) the effects resulting from the interaction of working memory capacity with instructional method on the acquisition of productive knowledge of polysemous L2 Spanish spatial prepositions. The target learners were adult L1 English speakers 18 years of age or over with no prior knowledge of Spanish or any cognate language (*ab initio* learners).

These target learners and these three areas of inquiry motivated three research questions and related hypotheses. The first research question and hypothesis examined the effectiveness of two techniques commonly used in the teaching of L2 polysemes. These two instructional methods were (1) Translation-based instruction (TBI), which treats the multiple meanings of polysemes as arbitrary, discreet and unrelated; (2) Cognitive linguistics-based instruction (CLBI), which treats the multiple meanings of polysemes as interrelated and motivated by an association to a common conceptual base via the processes of metaphor and metonymy. Immediate post-test scores suggest
that these two instructional methods are equally effective in developing short-term productive knowledge, but delayed post-test scores suggest that learners under CLBI acquire a greater level of long-term productive knowledge.

The second research question and hypothesis examined the predictive nature of working memory in the acquisition of L2 polysemes. Immediate and delayed post-test results suggest that higher scores in working memory capacity directly correlate to higher scores in productive knowledge of the four target prepositions.

Finally, the third research question and hypothesis examined effects resulting from the interaction of working memory with the two instructional treatments, CLBI and TBI. Immediate and delayed post-test results suggest that learner working memory capacity does interact with the instructional treatment. High working memory learners under TBI outscored their high working memory counterparts under CLBI, but low working memory learners under CLBI outscored their low working memory counterparts under TBI on the same immediate post-test.
PREFACE

This dissertation is an original, unpublished, independent work by the author, Joseph LeTexier.
# Table of Contents

DEDICATION ........................................................................................................... iii  
ACKNOWLEDGEMENTS ............................................................................................ iv  
ABSTRACT .............................................................................................................. v  
PREFACE .................................................................................................................. vii  
LIST OF TABLES ...................................................................................................... xi  
LIST OF FIGURES ................................................................................................... xiii  
LIST OF ABBREVIATIONS ......................................................................................... xiv

CHAPTER 1: INTRODUCTION ..................................................................................... 1  
   1.1 PROBLEM STATEMENT .................................................................................. 4  
   1.2 PURPOSE OF THE STUDY .......................................................................... 7  
   1.3 RATIONALE AND SIGNIFICANCE .............................................................. 9  
   1.4 OVERVIEW OF THE DISSERTATION .......................................................... 11

CHAPTER 2: LITERATURE REVIEW ............................................................................. 13  
   2.1 VARIABILITY IN ADULT SLA ...................................................................... 13  
   2.2 INDIVIDUAL DIFFERENCES, APTITUDE AND WORKING MEMORY ............ 19  
   2.3 MEASURING WORKING MEMORY CAPACITY .......................................... 31  
   2.4 VOCABULARY IN SLA ................................................................................ 36  
   2.5 THE NATURE OF L2 VOCABULARY KNOWLEDGE ................................... 38
CHAPTER 2: VOCABULARY ACQUISITION ................................................................. 48
2.6 L2 VOCABULARY ACQUISITION ................................................................. 48
2.7 L2 VOCABULARY INSTRUCTION................................................................. 52
2.8 PREVIOUS COGNITIVE LINGUISTICS-BASED RESEARCH.............................. 53
2.9 MEASURING L2 VOCABULARY KNOWLEDGE ........................................... 58
2.10 RESEARCH QUESTIONS AND HYPOTHESES ............................................. 69

CHAPTER 3: METHODOLOGY ........................................................................... 73
3.1 EXPERIMENTAL DESIGN ........................................................................... 73
3.2 TARGET PREPOSITIONS ........................................................................... 78
3.3 PARTICIPANTS ........................................................................................ 101
3.4 MATERIALS .......................................................................................... 103
3.5 PILOT TESTING ...................................................................................... 107
3.6 DATA CODING AND ANALYSIS ............................................................... 108

CHAPTER 4: RESULTS .................................................................................... 109
4.1 EFFECTS OF INSTRUCTIONAL TREATMENT ON SHORT AND LONG-TERM
PRODUCTIVE KNOWLEDGE ........................................................................ 113
4.2 EFFECTS OF WMC ON SHORT AND LONG-TERM PRODUCTIVE KNOWLEDGE
.................................................................................................................. 121
4.3 EFFECTS OF THE INTERACTION BETWEEN WMC AND INSTRUCTIONAL
TREATMENT ON BOTH SHORT AND LONG-TERM PRODUCTIVE KNOWLEDGE ..... 136

CHAPTER 5: DISCUSSION ............................................................................... 154
5.1 DISCUSSION OF THE EFFECTS OF INSTRUCTIONAL TREATMENT ON
PRODUCTIVE KNOWLEDGE ........................................................................ 155
5.2 DISCUSSION OF THE EFFECTS OF WMC ON SHORT AND LONG-TERM
PRODUCTIVE KNOWLEDGE ........................................................................ 171
5.3 DISCUSSION OF THE EFFECTS RESULTING FROM THE INTERACTION BETWEEN
WMC AND INSTRUCTIONAL TREATMENT ON THE ACQUISITION OF PRODUCTIVE
KNOWLEDGE ............................................................................................ 183
LIST OF TABLES

TABLE 2.1 CARROLL’S (1981) SUBCOMPONENTS OF LANGUAGE APTITUDE .................. 22
TABLE 2.2 SLA PROCESSING STAGES ........................................................................... 24
TABLE 2.3 SLA PROCESSING STAGES AND POTENTIAL APTITUDE COMPONENTS ........ 25
TABLE 2.4 THE ROLE OF COGNITIVE ID FACTORS IN SLA PROCESSING .................. 31
TABLE 2.5 THE THREE AREAS OF WORD KNOWLEDGE ........................................... 40
TABLE 2.6 ASPECTS OF WORD KNOWLEDGE FOR TESTING .................................... 60
TABLE 2.7 TEST TYPE ITEMS ..................................................................................... 62
TABLE 2.8 EIGHT VOCABULARY TESTING FORMATS ................................................ 63
TABLE 3.1 EXPERIMENTAL TIMELINE ......................................................................... 78
TABLE 3.2 SPECIFIC TARGET PREPOSITIONS ............................................................ 87
TABLE 3.3 CORE TEACHING PRINCIPLES OF TBI AND CLBI ................................... 105
TABLE 4.1 CLBI VS. TBI IMMEDIATE POST-TEST RESULTS ...................................... 114
TABLE 4.2 CLBI VS. TBI IMMEDIATE POST-TEST PREPOSITION RESULTS ............... 115
TABLE 4.3 CLBI VS. TBI IMMEDIATE POST-TEST SELECTION PATTERNS ................. 117
TABLE 4.4 CLBI VS. TBI DELAYED POST-TEST RESULTS ........................................ 117
TABLE 4.5 CLBI VS. TBI DELAYED POST-TEST PREPOSITION RESULTS .................... 118
TABLE 4.6 CLBI VS. TBI DELAYED POST-TEST SELECTION PATTERNS ..................... 120
TABLE 4.7 CB2 VS. CB1 IMMEDIATE POST-TEST RESULTS ........................................ 122
TABLE 4.8 CB2 VS. CB1 IMMEDIATE POST-TEST PREPOSITION RESULTS ................. 123
TABLE 4.9 CB2 VS. CB1 IMMEDIATE POST-TEST SELECTION PATTERNS .................... 124
TABLE 4.10 TB2 vs. TB1 IMMEDIATE POST-TEST RESULTS ........................................ 125
TABLE 4.11 TB2 vs. TB1 IMMEDIATE POST-TEST PREPOSITION RESULTS ................ 126
TABLE 4.12 TB2 vs. TB1 IMMEDIATE POST-TEST SELECTION PATTERNS .................. 128
TABLE 4.13 CB2 vs. CB1 DELAYED POST-TEST RESULTS ...................................... 130
TABLE 4.14 CB2 vs. CB1 DELAYED POST-TEST PREPOSITION RESULTS ................. 130
TABLE 4.15 CB2 vs. CB1 DELAYED POST-TEST SELECTION PATTERNS .................. 132
TABLE 4.16 TB2 vs. TB1 DELAYED POST-TEST RESULTS ....................................... 133
TABLE 4.17 TB2 vs. TB1 DELAYED POST-TEST PREPOSITION RESULTS ................. 134
TABLE 4.18 TB2 vs. TB1 DELAYED POST-TEST SELECTION PATTERNS .................. 136
TABLE 4.19 MANOVA RESULTS FOR INTERACTIONS ............................................. 138
TABLE 4.20 CB2 vs. TB2 IMMEDIATE POST-TEST RESULTS .................................... 139
TABLE 4.21 CB2 vs. TB2 IMMEDIATE POST-TEST PREPOSITION RESULTS ............ 140
TABLE 4.22 CB2 vs. TB2 IMMEDIATE POST-TEST SELECTION PATTERNS ............... 141
TABLE 4.23 CB1 vs. TB1 IMMEDIATE POST-TEST RESULTS .................................... 142
TABLE 4.24 CB1 vs. TB1 IMMEDIATE POST-TEST PREPOSITION RESULTS ............ 144
TABLE 4.25 CB1 vs. TB1 IMMEDIATE POST-TEST SELECTION PATTERNS ............... 145
TABLE 4.26 CB2 vs. TB2 DELAYED POST-TEST RESULTS ..................................... 146
TABLE 4.27 CB2 vs. TB2 DELAYED POST-TEST PREPOSITION RESULTS ............... 148
TABLE 4.28 CB2 vs. TB2 DELAYED POST-TEST SELECTION PATTERNS ................... 149
TABLE 4.29 CB1 vs. TB1 DELAYED POST-TEST RESULTS ..................................... 150
TABLE 4.30 CB1 vs. TB1 DELAYED POST-TEST PREPOSITION RESULTS ............... 152
TABLE 4.31 CB1 vs. TB1 DELAYED POST-TEST SELECTION PATTERNS ................... 153
TABLE 5.1 CLBI VERSUS TBI IMMEDIATE POST-TEST STANDARD DEVIATIONS ...... 161
LIST OF FIGURES

FIGURE 2.1 LINEAR RELATIONSHIP OF WORDS KNOWN AND READING COMPREHENSION ................................................................. 38

FIGURE 5.1 PARA -PROTO-SCENE ........................................................................... 165

FIGURE 5.2 A -PROTO-SCENE ............................................................................. 165
LIST OF ABBREVIATIONS

AO ................................................................. Age of Onset
CB1 ............................................................. Low WMC CLBI Experimental Group
CB2 ............................................................. High WMC CLBI Experimental Group
CLBI ............................................................ Cognitive Linguistics Based Instruction
CST ................................................................. Complex Span Task
FDH ............................................................. Fundamental Differences Hypothesis
FonF ............................................................. Focus on Form
GJT ............................................................. Grammaticality Judgment Task
HWM ............................................................. High Working Memory
IDs ............................................................. Individual Differences
L1 ..................................................................... First Language
L2 ..................................................................... Second Language
LWM ............................................................. Low Working Memory
OST ............................................................. Operation Span Task
RST ............................................................. Reading Span Task
SLA ............................................................. Second Language Acquisition
TB1 ............................................................. Low WMC TBI Experimental Group
TB2 ............................................................. High WMC TBI Experimental Group
TBI ............................................................. Translation-Based Instruction
WAT .................................................................................................................. WORD ASSOCIATION TASK
WM .................................................................................................................. WORKING MEMORY
WMC ............................................................................................................ WORKING MEMORY CAPACITY
CHAPTER 1: INTRODUCTION

Second language acquisition (SLA) research continues to seek answers to the question of why learners demonstrate such great variation in both their rates of acquisition and their level of ultimate attainment. So far, studies show that individual differences (IDs) are implicated as one of the most important predictors of L2 acquisition success (Dörnyei, 2005). As a result of this increasing evidence for the importance of IDs, it has become apparent that they need to be accounted for both in SLA theory and in the practical development of second language teaching methodologies. Perhaps the most important area of research lies in understanding how specific categories of IDs predict and interact with various stages of language learning processes (e.g., Dörnyei, 2005; Granena & Long, 2013; Robinson, 2002).

There are numerous empirical studies demonstrating strong correlations between individual differences and overall achievement in L2 acquisition (e.g., Dörnyei & Skehan, 2003; Robinson 2002; Skehan 2002). There is also growing evidence that individual learners will interact differently with various techniques employed by second language teachers (Lam, 2009; Touplikioti, 2007). In citing three different language learning experiences, Robert Sternberg (2002, p. 13) states that “I was being taught in different ways and responding differently to each of these ways. My aptitude was not internal to me, but in the interaction between my abilities and the way I was being taught.” It is important for SLA
theory to explain the causes of variation in language learning success that is due to IDs interacting differently under different instructional conditions. Individual differences such as aptitude, age and motivation are important variables to be researched in the field of SLA (Dörnyei, 2005; Robinson, 2002). Our understanding of SLA is only partial until we have clear explanation as to how IDs interact with specific instructional methods.

For years, SLA researchers have sought to determine the route that second language acquisition takes, as well as what factors affect the rate of SLA (e.g., Gass, 2013; Gass & Mackey, 2012; VanPatten & Williams, 2007). A great deal of knowledge has been developed, but there are still many questions that are left unanswered. One of the most pressing issues facing researchers is determining the factors that lead to such great variation in final attainment among adult second-language learners. This is in contrast to first language acquisition where very little variation is observed in final attainment. It is also not just an issue of inter-learner variability, but also intra-learner variability. That is to say, there exists tremendous variability in L2 acquisition success both between learners and within the experiences of the individual adult learner. A learner may experience a slow rate of acquisition during the study of a language at one point in life, but have great success in the same language at another point in life (Granena & Long, 2013; Robinson, 2002). The present study explored (1) how individual differences, specifically working memory capacity, affect vocabulary acquisition, (2) how the method of instruction affects vocabulary acquisition, and
finally, (3) how ID’s and method of instruction interact to affect vocabulary acquisition.

Just as IDs have been demonstrated to be strong predictors of L2 acquisition success, previous research shows strong correlations between breadth of vocabulary knowledge and overall proficiency in a foreign language (see Alderson, 2005). Wilkins (1972, p. 111) notes in this regard that “without grammar very little can be conveyed, without vocabulary nothing can be conveyed.” Schmitt (2010) argues that there is strong evidence for the importance of vocabulary in all facets of language proficiency. This is evident from the typically high correlations between vocabulary knowledge and various measures of language proficiency. (e.g., Albrechtsen, Haastrup & Henriksen 2008; Laufer & Goldstein 2004). Although great strides have been made in the development of more effective techniques for developing the L2 learner’s lexicon, nonetheless, there continues to be a high degree of variability between learners in terms of both overall breadth as well as depth of vocabulary knowledge. Of course, variability between L2 learners is not limited to vocabulary knowledge, but rather, it is one of the most distinguishing characteristics separating adult L2 learners in all aspects of L2 knowledge.

One of the most comprehensive explorations of the relationship between vocabulary knowledge and language proficiency occurred as part of the development of the DIALANG tests. DIALANG is described in the Common European Framework of Reference Guidelines (p. 226) as an assessment system intended for language learners who want to obtain diagnostic information
about their proficiency (see Alderson, 2005, for a detailed account). From the evidence, he states that the “DIALANG analysis would appear to show that the size of one’s vocabulary is relevant to one’s performance on any language test, in other words, that language ability is to quite a large extent a function of vocabulary size” (2005, p. 88).

1.1 **Problem Statement**

There is growing evidence from SLA research that demonstrates strong correlations between individual differences and overall achievement in L2 acquisition (e.g., Dörnyei & Skehan, 2003; Robinson 2002; Skehan 2002). There is also evidence that the individual learner’s general cognitive abilities will interact differently with learning under different instructional strategies and learning conditions employed during second language acquisition (Canner, 2013).

The theoretical framework for the present study was Robinson’s (2002) aptitude-treatment interaction framework which was designed to explore the interaction between IDs and instructional treatments/learning conditions. Based on this framework, the present study attempted to fill some of the gaps in our understanding of when L2 vocabulary acquisition occurs as a result of an interaction between working memory (WM) and instructional treatments and if WM predicts the rate of vocabulary acquisition. In addition, this study provides information on the relative effectiveness of two instructional treatments commonly used in the teaching of polysemous words, cognitive-linguistics based instruction (CLBI) and translation-based instruction (TBI).
In addition to improving our understanding in these specific aspects of SLA, these areas of inquiry also helped to explain some of the causes of variation in final attainment in L2 vocabulary acquisition. Robinson (2002, p. 122) claimed that a major challenge for ID condition interaction research is trying to explain why patterns of abilities lead to learning outcomes in any one context in terms of proposed SLA processes and mechanisms. Furthermore, he argued that in order to explain how when and why these general cognitive abilities are employed during adult SLA, it will be necessary to link patterns of abilities to particular acquisition processes, and information processing demands of learning contexts. Robinson argued that it is this link that causes their effects in SLA. One of the primary purposes of the present study was to determine the link between learners’ working memory and L2 vocabulary acquisition success under two different instructional techniques. Adult L2 vocabulary knowledge can be divided into two main aspects, productive knowledge and receptive knowledge. The present study focused exclusively on the acquisition of productive knowledge, as learners could negotiate receptive meaning of the L2 target prepositions simply by observing the pictures and apply their L1 knowledge to the elicitation tasks.

There are a number of individual learner variables that can affect the acquisition of L2 vocabulary; however, working memory was chosen because of its recognition as the primary cognitive variable affecting adult SLA (e.g. Miyake & Friedman, 1998). Robinson (2002) argues that there exists a set of cognitive abilities, or aptitude complexes, that relate differently to language learning under
different psycholinguistic processing conditions. He describes these conditions as the situational level of classroom instruction and the specific pedagogic tasks that learners perform in classrooms. He also describes these conditions as the cognitive level of implicit, explicit, and incidental learning processes. Robinson argues that purposely matching learners’ strengths in particular aptitude complexes with specific learning conditions and instructional techniques will be an important element in the delivery of optimally effective classroom exposure and practice for L2 learners.

Snow (1994) argues that a theory of cognitive abilities contributing to IDs in specific aptitudes for learning (learner variables) must be developed in conjunction with a theory of contextual or learning constraints (context variables) (see also Corno, Cronbach, Kupermintz, Lohman, Mandinach & Porteus, 2002). Snow describes the interactionist perspective on ID research as follows: the “relevant aspects of person and situation are specified, their interaction is demonstrated empirically, and some process explanation of how and why this occurs is offered” (1994, p. 4).

Robinson (2002) argues that by using Snow’s approach, correlations between cognitive variables and learning outcome measures can be examined in relation to the differing information processing demands of learning conditions and instructional methods. He also argues that when interactions between cognitive variables or aptitude components such as working memory and learning conditions occur, these are attributed to acquisition processes which utilize the cognitive variables or aptitude components of interest. Robinson
claims that the “extent to which IDs in cognitive abilities differentially affect second language acquisition under different conditions of exposure is an issue of theoretical, and practical importance” (2002, p. 211).

1.2 Purpose of the Study

The primary purposes of the present study were to (1) examine how vocabulary words, specifically polysemous spatial prepositions, are most effectively taught and learned, (2) how learner WMC may differentially affect vocabulary acquisition under differing instructional strategies and (3) if WMC has predictive power in understanding a learner's final attainment in L2 vocabulary knowledge. There is growing theoretical evidence related to how language aptitude affects SLA, but our understanding is incomplete until we can answer how and why learning happens under certain conditions but not under others. Robinson (2002, p. 113) argues that identifying individual differences in cognitive abilities, and matching these individual differences with the most effective instructional methods is the primary aim of SLA related language aptitude research. Therefore, it is important in SLA theory to explain the causes of variation in language learning success under different instructional conditions. Individual differences such as aptitude, age and motivation are important variables to be researched, especially in regards to how they may affect or predict SLA success.

Robinson (2002, p. 122) writes “of essential theoretical interest, then, in this research is a characterization of how these variously specified learning conditions constrain information processing, and how the information processing
resource demands of these conditions are affected by IDs in the extent of resource availability and the structure of abilities.”

In addition to language aptitude, there is a growing body of evidence supporting the important role that vocabulary knowledge has on overall second language proficiency. Vermeer (1992, p. 147) claims that “knowing words is the key to understanding and being understood. The bulk of learning a new language consists of learning new words.” However, even though vocabulary has been shown to be of high importance in SLA, there still is no understood best practice for the teaching of vocabulary. Schmitt (2010) details several prominent knowledge gaps in the field of vocabulary studies. He argues that there is yet to be developed an overall theory of vocabulary acquisition. He describes the development of this theory as the “Holy Grail” of vocabulary studies. He argues that although we have grown in our understanding of the development of specific aspects of vocabulary knowledge, the overall acquisition system is too complex and varied to be understood in its entirety. No one study will answer all the questions related to vocabulary acquisition, rather it will take a large number of separate studies that can be used in combination to better understand how vocabulary is acquired. Schmitt also argues that we have yet to understand how the development of vocabulary knowledge moves from no knowledge to receptive mastery and finally productive mastery. The purpose of the present study was to specifically add to the body of knowledge related to second language vocabulary acquisition.
1.3 RATIONALE AND SIGNIFICANCE

There are two significant contributions that the present study made towards current SLA theory and pedagogy. The first contribution is to our understanding of how individuals may differ in how they best acquire vocabulary in a second language. This is important as numerous studies show high correlations between vocabulary knowledge and various measures of language proficiency (e.g., Albrechtsen, Haastrup & Henriksen, 2008; Alderson, 2005; Laufer & Goldstein, 2004), but none has provided specific reasons for success or lack thereof in acquiring L2 vocabulary.

The second contribution of the present study was to increase our understanding of how IDs may interact with vocabulary teaching methods, as well as predict overall success in acquiring vocabulary in a second or foreign language. Robinson (2002, p. IV) writes, “learning is a result of the interaction between learner characteristics, and learning context. Describing and explaining these patterns of ID intervention interactions are fundamentally important to theories of instructed second language acquisition, and for effective pedagogy.” He also describes the purpose of research in IDs as trying to determine the fit between the second language learner and learning condition in second language classrooms. Numerous researchers (Snow, 1987; Sternberg & Wagner, 1994) have pointed out the fact that it is only through research that the connection between learners’ individual differences and specific learning conditions can be determined.
Robinson (2002) argues that research into aptitude treatment interactions also has theoretical relevance and potential explanatory value, in addition to pedagogic utility. He claims that an explanation of SLA requires a *transition theory* as well as a *property theory*. He describes a transition theory as one which specifies relationships between cognitive abilities, acquisition processes and the mechanisms that move a learner’s knowledge from one point to another along a continuum. He describes a property theory as simply a characterization of the properties of knowledge at different points of learning. In summary, how do a learner’s cognitive abilities interact with acquisition processes in specific teaching methodologies in order to move the learner from one clearly defined point of knowledge to another more advanced point? Robinson (2002) argues the following:

“Accounting for the findings from research into the effects of individual differences in cognitive abilities on acquisition process should therefore form an important part of any *transition theory* and the causal relationships it proposes between cognitive resource allocation and learning mechanisms; illuminating how these mechanisms are integrated in cognitive architecture.” (p. 115)

He goes on to argue that in the aptitude-treatment interaction framework, it will be important to study the effects of individual differences in the cluster of abilities (aptitude complexes) in order to support language learning under
particular conditions of exposure. Aptitude-treatment interaction studies are theoretically important, since they can help to cast light on the cognitive correlates and components of implicit, incidental, and explicit L2 acquisition processes. Robinson claims that matching both instructional treatments and the different techniques for focus-on-form that those treatments may make use of, to aptitude complexes is an important part of effective classroom L2 instruction and is therefore an area of needed further research.

1.4 OVERVIEW OF THE DISSERTATION

The remainder of the present study is divided into five chapters. Chapter two provides a review of the current literature related to the following topics: 1) variability in adult SLA; 2) individual differences, aptitude and working memory; 3) measuring working memory; 4) vocabulary in SLA; 5) the nature of L2 vocabulary knowledge; 6) L2 vocabulary acquisition; 7) L2 vocabulary instruction; 8) measuring L2 vocabulary knowledge; 9) research questions. Chapter three provides a description of the research methods that were used in the present study, including the experimental design, participants, experimental and testing materials used, pilot testing, and data coding and analysis procedures. Chapter four reports the results for each of the four experimental groups from their immediate and delayed post-tests. Chapter five provides a discussion of the results for both the general patterns seen in the data and for the two post-tests. For greater ease in reading, Chapters four and five follow the same general format with a report and discussion of the results in terms of each of the three
research questions. Finally, chapter six provides a conclusion with study limitations and recommendations for future research.
CHAPTER 2: LITERATURE REVIEW

The goal of the present study was to fill the gaps in the current body of research findings in three important areas related to adult L2 vocabulary acquisition. In order to better understand the background of the research that was conducted, the following literature review examined what is currently known about the following areas of adult L2 acquisition: (1) variability in adult SLA; (2) IDs in language aptitude and working memory effects in adult SLA; and (3) L2 vocabulary acquisition.

2.1 VARIABILITY IN ADULT SLA

This section of the literature review examined current research related to the enduring gap in our understanding of the causes of variability between adult L2 learners, in terms of their rate of acquisition as well as final attainment in specific measures of proficiency. This gap in our understanding is due in large part to the fact that the underlying causes for variability in adult SLA are so numerous and diverse. Factors such as the age of onset (AO), motivation, instructional techniques, cognitive abilities and learning conditions can all greatly affect L2 proficiency. As a result, no single study can possibly provide answers for inter-learner variability; rather, it will take the combination of numerous studies to provide a fuller picture.

Robinson, (2002) argues that rigorous empirical studies of the effects of IDs on instructed and naturalistic learning add to our knowledge of how
these individual differences interact with the conditions of instructed language exposure. He goes on to argue that research in the interaction between individual differences and instructed language learning is necessary not only to gain new insights into the cognitive correlates of SLA processes, but also to help researchers and teachers to design more effective instructional practices which aim to facilitate language acquisition.

**Variability in Final Attainment in Adult L1 versus L2 Proficiency**

There are a number of researchers who have attempted to explain the great variability in final attainment in adult L2 proficiency which is in sharp contrast to observed differences in final attainment in L1 proficiency. Bley-Vroman’s (1990) Fundamental Differences Hypothesis (FDH) argues that general cognitive abilities are employed during adult SLA, in contrast with L1 acquisition which develops under pre-wired learning mechanisms that guide L1 acquisition. Bley-Vroman contends that the differences in observed adult L2 attainment are therefore the result of differences in individual cognitive resources and abilities that are drawn on in L2 acquisition.

Skehan (1998) also argues for a general cognitive basis for L2 acquisition stating that there is evidence for two modules connected to the acquisition of the L1, the syntax and semantics modules. The syntax module is prewired or innate and aids in the acquisition of syntax and is subject to maturational constraints. On the other hand, the semantics module is not subject to maturational constraints, and continues to operate throughout life. By comparison, Skehan argues for the existence of three modules related to L2 acquisition, each of which
is in turn, connected to an aptitude component. These three modules are (1) auditory processing, (2) language processing and (3) memory. Skehan states that the auditory processing module is connected to the operation of phonemic coding ability. The language processing module is connected to inductive language ability, and the memory module to the extended conceptualization of the functioning of memory. The important distinction between the L1 modules and L2 modules is that in the acquisition of the L2, learners no longer have access to the prewired system which enables them to acquire syntactic features; rather, in the acquisition of the L2 learners have to rely on general learning mechanisms. The variance in general learning mechanisms between learners accounts for a lot of the variability between learners in ultimate L2 attainment.

**DEPENDENCE ON FOCUS-ON-FORM IN ADULT L2 ACQUISITION**

Another factor that may contribute to inter-learner variability is the adult language learner’s dependence upon a measured amount of Focus-on-Form (FonF) (Long and Robinson, 1998). VanPatten and Benati (2010) state that FonF generally refers to any intervention in which a teacher draws simultaneous attention to both meaning and how that meaning is encoded. A good example of this is the case of recasts. In a recast, a learner produces something that is not quite native-like, and the native speaker interlocutor recasts what was said to show both that they understood the learner and to also draw brief attention to the formal features of the language.

VanPatten (1990, 2007) argues that when L2 learners are put under processing pressure, they will attend first to meaning, and it is only when they
have spare processing capacity available that they will also attend to form. The challenge of adult L2 acquisition is therefore vitally connected to the learner’s having the spare processing capacity available to actually attend to form in the input. Where learners differ in their processing capacity during different stages of L2 acquisition, they will also differ in the levels of attainment that they reach.

**Intra-learner Variability**

In addition to inter-learner variability, adult L2 acquisition is also characterized by *intra-learner* variability. The present study understands intra-learner variability as the phenomenon in which the language learner differs in the level of proficiency attained based on such factors as learning contexts (i.e., formal vs. informal) and instructional methods employed. Sternberg (2002) notes that in the study of three different languages, he found varying degrees of success in the acquisition of each of these. From this, he hypothesized that learner’s may have varying levels of language learning aptitude dependent upon the specific language being learned. He contends that although learners may find one language easier to learn than another, there is likely additional mechanisms involved that would explain varying levels of L2 acquisition success. He goes on to describe these mechanisms as involving multiple aspects of language aptitude or even multiple intelligences in general ability theories. Gardner (1983, 1999) proposed a model of multiple intelligences involving eight distinct intelligences: (1) linguistic, (2) logical-mathematical, (3) spatial, (4) musical, (5) bodily-kinesthetic, (6) interpersonal, (7) intrapersonal and (8) naturalistic. Sternberg argues that success in language acquisition may vary as
a function of the specific language being taught, the way it is taught, how progress is assessed and the interaction of all of these with the individual learner’s intelligence in the eight categories listed above.

Some researchers have attempted to explore the causes of intra-learner variability. Wesche (1981) found strong connections between language learner aptitude profiles and instructional treatments. Wesche reported that not only did learners perform better when matched with methods that aligned with their aptitude profile, but they also reported greater satisfaction with instruction. Canner (2013) revealed correlations between language learner aptitude and levels of proficiency in L2 Russian attained in naturalistic versus formal learning contexts. The focus of this study was the learner’s oral proficiency in three distinct areas related to both fluency and accuracy: (1) a measure of fluency in terms of the number of meaningful syllables uttered per minute, (2) an assessment of accuracy in morphology and syntax measured by eliciting specific and frequently-uttered constructions that contain the major elements necessary for effective, native-like speech, and (3) overall ability in terms of both command of vocabulary and effective use of major structures commonly used in Russian. These studies not only point to the need for matching students with appropriate methods based on IDs, but also to the fact that intra-learner variability may be the result of the interaction between learner characteristics and learning methods and contexts.

Few studies have investigated the effects of vocabulary teaching methods on vocabulary acquisition. There have also been few studies that have
examined the interaction of IDs with vocabulary teaching methods in the retention of vocabulary. One study that attempted to shed light on these areas of vocabulary acquisition was conducted by Levine and Reves (1990). In this study, the authors researched the extent to which differences in vocabulary retention were related to different methods of vocabulary presentation. They also explored how different methods of vocabulary presentation interact with different learner factors such as personality, L1 background, word-processing habits and language attitudes. The methods of vocabulary presentation used were (1) written presentation of L2 word and its L1 translation, (2) written presentation of L2 word in sentential context (3) L2 word with a picture, (4) written presentation of L2 word with its meaning, (5) auditory presentation of L2 word and its L1 translation, (6) auditory presentation of L2 word in sentential context and (7) three-fold computer presentation (word and its definition, word presented in analogy, word in context). The findings from this study show that the method of presenting new vocabulary leads to varying degrees of vocabulary retention. They argue that the retention of vocabulary seems to be related to the learner’s general learning patterns and/or cognitive styles of visual, auditory and contextual associations. They found that visual presentation of vocabulary led to higher levels of attainment. They also found that various learner factors or IDs combined differently with various methods of vocabulary presentation. As a result, they argue that the processing of learning new vocabulary is a multifarious challenge; therefore, no single method should be imposed on learners.
2.2 INDIVIDUAL DIFFERENCES, APTITUDE AND WORKING MEMORY

Research shows that the influence of IDs on L1 acquisition is very different from L2 acquisition (see discussions in DeKeyser, 2000; Harley & Hart, 1997). There is great variation in rates of acquisition and final attainment among L2 learners. In contrast, L1 acquisition virtually always leads to proficient native like ability, and the rate at which this ability is attained is much less varied than is found among L2 learners. This section of the literature review begins with a general discussion of IDs. It is followed by a section on language aptitude, and concludes with an exploration of working memory which is one the components of language aptitude.

INDIVIDUAL DIFFERENCES

Research shows that IDs are one of the primary causes of observed variability in adult L2 proficiency (see discussion in Canner, 2013; Levine & Reves, 1990; Wesche, 1981). The major categories of IDs that have been demonstrated to affect SLA include cognitive and learning styles, language learning strategies, motivation and language aptitude. Dörnyei and Skehan (2003) argue that the last two, language aptitude and motivation, show the greatest potential to generate a promising SLA research program.

Dörnyei and Skehan adopt Keefe and Perrell's (1990) definition of learning style which reads, "A complexus of related characteristics in which the whole is greater than its parts. Learning style is a gestalt combining internal and external operations derived from the individual's neurobiology, personality and development, and reflected in learner behavior" (1990, p. 16). Dörnyei and
Skehan use this definition to bring out a contrast between *cognitive* style and *learning styles*. They define cognitive style as the predisposition to process information in a characteristic way and learning styles as a general preference for how to approach learning. Dörnyei and Skehan argue that cognitive style is related more to information-processing preferences, while learning styles connects to all aspects of learning.

Dörnyei and Skehan combine research by O'Malley and Chamot (1990), Oxford (1990), and Wenden (1991) to define the concept of a *language learning strategy* as the learner's active contribution to improving the overall effectiveness of his or her own learning. They go on to argue that research (e.g., Naiman, Frohlich, Stern, & Todesco, 1978; Rubin, 1975; Wong Fillmore, 1979) demonstrates that some learners excel due to their own active participation in the learning process through the application of personalized learning techniques. This demonstrates the importance of matching individual learners with individualized learning techniques in order to optimize learning outcomes.

The concept of motivation and its influence on SLA has been extensively researched (Clement and Gardner, 2001; Dörnyei, 1998, 2001). Dörnyei argues that motivation is concerned with the affective characteristics of the learner, and refers to the “direction and magnitude of learning behavior in terms of the learner’s choice, intensity, and duration of learning” (2009, p. 231). Dörnyei and Skehan (2003) reports that IDs in second language learning, primarily foreign language aptitude and motivation, have generated the most consistent predictors of SLA proficiency. They report that studies which provide data correlations of
aptitude or motivation with language achievement consistently range between 0.20 and 0.60, with a median value a little above 0.40. They argue that since aptitude and motivation do not show particularly high correlations with one another, they combine to yield multiple correlations which are frequently above 0.50.

In contrast to the overview of IDs given above, Robinson (2002) broadly categorizes learner variables or IDs as being either cognitive or affective. Cognitive abilities would include such things as intelligence, language learning aptitude, working memory capacity (WMC) and speed. Affective variables include factors such as emotions and motivation. VanPatten and Benati (2010) describes individual differences as a set of personality and psycho-emotive characteristics that learners bring to the task of learning. Dörnyei (2009) probably best summarizes IDs as “the background learner variables that modify and personalize the overall trajectory of the language acquisition processes” (231).

**APTITUDE & SLA PROCESSING STAGES**

This section provides an overview of the definitions and core constructs of language aptitude, along with the putative connections between these aptitude constructs and SLA processing stages. Snow (1992) proposes that aptitude has several meanings, including readiness, suitability, susceptibility and proneness for learning in particular situations. He also argues that aptitude is not a constant and prewired intellectual capacity; rather, it is a complex of individual characteristics that interact dynamically with the situation in which the learning
takes place. From this view of aptitude, it follows that different sets of abilities can enhance learning under various learning conditions.

In any discussion about language learning aptitude, it is probably best to start with the work of John Carroll (1958, 1962) who established the method for studying aptitude as well as its component parts. Through his research, he was able to develop a large number of tests which could be used to determine fundamental capacities involved in second language acquisition. From these numerous tests, the modern language aptitude test (MLAT) was developed. The MLAT provided a measure which produced a reasonably high correlation with language course performance, and it has been a cornerstone of aptitude research ever since. Carroll (1981) suggests that there are four subcomponents that make up the broader construct of language aptitude. In table 2.1 below, I’ve named and given a brief description of each of these four sub-components.

**TABLE 2.1 CARROLL’S (1981) SUBCOMPONENTS OF LANGUAGE APTITUDE**

<table>
<thead>
<tr>
<th>Name of sub-component</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonetic coding ability</td>
<td>The ability to encode unfamiliar sounds into long-term memory</td>
</tr>
<tr>
<td>Grammatical sensitivity</td>
<td>The ability to identify the grammatical function of words in inductive sentences</td>
</tr>
<tr>
<td>Inductive language learning ability</td>
<td>The ability to notice and identify patterns to create new sentences</td>
</tr>
<tr>
<td>Associative memory</td>
<td>The ability to form links in memory</td>
</tr>
</tbody>
</table>

Carroll (1973) argued that aptitude is either genetically determined or becomes fairly stable early in life. Politzer and Weiss (1969) is an example of a failed attempt to train aptitude, further supporting the notion that aptitude is fixed. However, Kormos (2013) claims that there is converging evidence that certain
components of aptitude, specifically phonological sensitivity and metalinguistic awareness, that might improve in the course of language learning. Bialystok and Majumder (1998) demonstrates the cognitive advantages of bi- and multilingualism. In other words, having knowledge of more than one language may be beneficial in developing aptitude in phonological sensitivity and metalinguistic awareness; however, there is currently no research evidence that supports the possibility of developing working memory. Working memory is a largely fixed component of the language learner’s cognitive abilities, therefore, for the present study, it was assumed that participants would be unable to improve their working memory at any point during the study.

Language learners differ in individual cognitive abilities as well as combinations of abilities. Robinson (2002) utilizes the interactionist framework of Snow (1987, 1994) to identify a number of aptitude complexes or combinations of cognitive abilities of the language learner. He argues that these aptitude complexes or combinations of cognitive abilities are differentially related to processing under different conditions of instructed language learning. As a result, Robinson contends that strengths in one or another of these aptitude complexes or cognitive abilities can be expected to be more important for learning under one instructional method, or in one learning condition, as compared to another.

In Table 2.2 (Skehan, 2002, p. 88-89) names and briefly explains the various processing stages associated with SLA. These stages represent the
development of an L2 structure from the initial stage of simply giving attention to it to being able to use it accurately and fluently.

**TABLE 2.2 SLA PROCESSING STAGES**

<table>
<thead>
<tr>
<th>SLA Processing Stage</th>
<th>Nature of Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noticing</td>
<td>Learner directs attention to some aspects of the language system, or as led to direct attention in this way</td>
</tr>
<tr>
<td>Pattern identification</td>
<td>On the basis of the focal attention, the learner makes a hypothesis about the target language based on perceived pattern or regularity</td>
</tr>
<tr>
<td>Extending</td>
<td>The learner extends the domain of the hypothesis, without changing it fundamentally in-kind</td>
</tr>
<tr>
<td>Complexifying</td>
<td>The learner apprehends the limitations of the identified pattern, and restructures it, as new aspects of the target language are noticed</td>
</tr>
<tr>
<td>Integrating</td>
<td>The learner takes the output of this process of complexification and integrates the new sub area of inter-language into a larger structure</td>
</tr>
<tr>
<td>Become an accurate, avoiding error</td>
<td>The learner becomes able to use the inter-language area without making errors, although this use may be slow and effortful</td>
</tr>
<tr>
<td>Creating a repertoire, achieving salience</td>
<td>Not only can error be avoided, but the inter-language form can be accessed at appropriate places, becoming part of a salient language repertoire</td>
</tr>
<tr>
<td>Automatising rule-based language, achieving fluency</td>
<td>The domain is now used not simply without error, but with reasonable speed, and the role has become, to some degree, proceduralized</td>
</tr>
<tr>
<td>Lexicalising</td>
<td>The learner, at this stage, is also able to produce the inter-language form in question as a lexical item element.</td>
</tr>
</tbody>
</table>

Robinson (2002) argues that in order to explain how, when and why Carroll's (1981) sub-components of language aptitude (see Table 2.1) are employed during adult SLA, it will be necessary to link patterns of abilities to particular acquisition processes and information processing demands of learning contexts. Robinson claims that it is this link that causes their effects on SLA, and
it is this link that helps explain the causes of inter-learner and intra-learner variability discussed previously.

Skehan (2002) then argues that if the stages shown in Table 2.2 are accepted, then the question that needs to be asked is if there is variation in the speed of learning in each of these areas? If the answer to this question is yes, then researchers need to consider that the differences in learning at each of these stages are the result of some component of aptitude. In Table 2.3, Skehan (2002, p.90) proposes potential aptitude components connected with specific SLA processing stages.

**TABLE 2.3 SLA PROCESSING STAGES AND POTENTIAL APTITUDE COMPONENTS**

<table>
<thead>
<tr>
<th>SLA Processing Stage</th>
<th>Potential Aptitude Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noticing</td>
<td>Auditory segmentation</td>
</tr>
<tr>
<td></td>
<td>Attention management</td>
</tr>
<tr>
<td></td>
<td>Working memory</td>
</tr>
<tr>
<td></td>
<td>Phonemic coding</td>
</tr>
<tr>
<td>Pattern identification</td>
<td>Fast analysis/working memory</td>
</tr>
<tr>
<td></td>
<td>Grammatical sensitivity</td>
</tr>
<tr>
<td>Extending</td>
<td>Inductive language learning ability</td>
</tr>
<tr>
<td>Complexifying</td>
<td>Grammatical sensitivity</td>
</tr>
<tr>
<td></td>
<td>Inductive language learning ability</td>
</tr>
<tr>
<td>Integrating</td>
<td>Restructuring capacity</td>
</tr>
<tr>
<td>Becoming accurate, avoiding error</td>
<td>Automatisation</td>
</tr>
<tr>
<td></td>
<td>Proceduralisation</td>
</tr>
<tr>
<td>Creating a repertoire, achieving salience</td>
<td>Retrieval processes</td>
</tr>
<tr>
<td>Automatizing rule-based language, achieving fluency</td>
<td>Automatisation</td>
</tr>
<tr>
<td></td>
<td>Proceduralisation</td>
</tr>
<tr>
<td>Lexicalising, dual-coding</td>
<td>Memory, chunking, retrieval processes</td>
</tr>
</tbody>
</table>

Even though SLA research has shown strong correlations between aptitude and ultimate L2 attainment, it has garnered only moderate interest.
Dörnyei and Skehan (2003) argue that one of the reasons aptitude has received so little interest is because of its perceived irrelevance to language acquisition within communicative contexts. Krashen (1981) argues that measures of language learning aptitude tend to only predict the effects of instruction in specific kinds of methods in educational settings such as the audiolingual method or grammar focused instruction. From this, Krashen claimed that aptitude could only predict learning not acquisition in formal classroom settings. However, numerous researchers (Dörnyei, 2005; Dörnyei and Skehan, 2003; Harley & Hart 2002; Ranta, 2002; Skehan, 1989, 2002) have shown that this dismissal of aptitude and its importance in formal classroom instruction is without merit.

DeKeyser (2000) explored the effects of age on language acquisition success. He reports a steady decline in the ability to acquire a foreign language through age 17. The important factor from his study for the current study is the fact that DeKeyser used aptitude tests with his subjects. The use of aptitude tests allowed him to determine correlations between aptitude, age and foreign language proficiency. He reports that for learners who began to learn English in the United States before the end of the critical period there was no significant correlation between their aptitude for language learning and their proficiency in English. However, for those who began to learn English after the end of the critical period, there were strong correlations between their aptitude and proficiency in English. These results provide evidence for the importance of language learner aptitude after age 17.
Robinson (2002, p. 129) argues “that sets of abilities, or aptitude complexes, may be differentially related to learning outcomes that result from learning under different processing conditions.” Robinson brings together various described hypotheses, and shows how they are related, defining the Aptitude Complexes Ability/Differentiation Framework. His summary includes the following:

1. There are child-adult differences in language learning; adults rely heavily on general problem solving abilities and exhibit much greater variation in levels of attainment (the Fundamental Difference Hypothesis). But the FDH alone does not explain why variation in levels of attainment should be so great.

2. The information processing demands of tasks draw differentially on cognitive abilities, or aptitude complexes (the Aptitude Complex Hypothesis).

3. Therefore, adult learning under any condition is fundamentally similar, since it is a result of the interaction between a pattern of cognitive abilities and the consciously mediated processing demands of the task (the Fundamental Similarity Hypothesis).

4. Points two and three together help explain variation in adult L2 learning outcomes:
   
   (i) patterns of abilities need to be matched to learning tasks and conditions to be effective, and they often are not.
(ii) some learners may have differentiated abilities (a number of high highs, as well as high lows, and low lows in aptitude complexes) (i.e., exhibit multiple aptitudes), whereas others have less differentiated abilities (high lows, and low highs) (i.e., exhibit a stronger general aptitude factor) (the Ability Differentiation Hypothesis).

(iii) it follows that for groups of learners with more differentiated abilities, there will be more variation in learning in any one environment, or on any one task (i.e., less matching of abilities and processing demands) than for groups of learners with less differentiated abilities.

**Working Memory**

The specific component of aptitude that served as an independent variable in the present study was *working memory* (WM), with the learner’s ability in WM characterized as the individual’s “working memory capacity” (WMC). WM is used here according to the definition by Gathercole & Baddeley (1993) and Baddeley (2003), who describe this component of aptitude as a system of “temporary storage and manipulation of information that is assumed to be necessary for a wide range of complex cognitive activities” (Baddeley, 2003, p. 189). VanPatten and Benati define WM as a “psychological construct referring to the processing space in the mind when a person is computing information” (2010, p. 167). In regards to SLA, working memory is what language learners
use to briefly store and process new linguistic input in order to analyze it for comprehension.

There is a growing body of research that demonstrates that adult L2 learners rely on certain types of cognitive resources, especially WM, to attain high levels of proficiency in an L2 (e.g. DeKeyser, 2000; Harley and Hart, 1997, 2002; Ross, Yoshinaga and Sasaki, 2002). This is in contrast with child language acquisition which develops under more innate processes and resources. The results of these studies give support to Bley-Vroman’s (1989) Fundamental Difference Hypothesis (FDH), which argues for a fundamental difference between adult and child acquisition processes. As such, IDs in working memory have a direct impact on adult acquisition. Since differences in WM exist between learners, these differences naturally cause variability in levels of proficiency and final attainment.

VanPatten and Benati (2010) note that although there are multiple theories and models of working memory, they all hold in common that WM has a limited capacity. Baddeley and Hitch (Baddeley, 1986; Baddeley & Hitch, 1974; 2003a; 2003b) developed the most recognized model of working memory. In contrast with other models of WM which focused on the storage function of memory, they developed a more comprehensive approach. Their model of WM simultaneously combines storage with the processing and manipulation of information. This understanding of WM more closely links it with cognitive activities such as comprehension, reasoning and learning than previously
understood. It also links it more closely to the cognitive processes employed during L2 acquisition.

Baddeley and Hitch’s model conceptualizes WM as a system consisting of multiple components. These components include (1) the *central executive*, which coordinates two modality-specific or slave subsystems, (2) the *phonological loop*, which stores phonological information and prevents its decay by continuously articulating its contents and (3) the *visuo-spatial sketchpad* which stores visual and spatial information. In other words, the visuo-spatial sketchpad manipulates and retains visual and spatial information, while the phonological loop is specialized for the manipulation and retention of speech. Baddeley (2000) extended the model by adding a fourth component, the *episodic buffer*, which uses multidimensional coding to integrate phonological, visual, and spatial information, and possibly information not covered by the two slave systems such as semantic and musical information.

The conceptualization of the central executive has evolved since the original model. Baddeley (2003) argues that it is the most important and least understood component of working memory. This is probably due to the fact that it is more complex and performs several functions, including attentional control, and directing the flow of information through the system (Gathercole, 1999). The most widely researched component of working memory is the phonological loop. This subsystem consists of a phonological store, which holds information for a few seconds, as well as an articulatory rehearsal process, which refreshes decaying information. The rehearsal process is analogous to sub-vocal speech
and takes place in real-time. As a result, there is limited span of immediate memory. For each individual, there will be a limited number of items they can hold in immediate memory before the first item fades and before it can be rehearsed again. The phonological loop is important for the present study as it plays a crucial role in the learning of new words by storing unfamiliar sound patterns so that long-term representations can be formed (Baddeley, 1986).

Kormos (2013, p. 142) provides a list of cognitive IDs demonstrated to influence identified language learning processes. The important fact to note in regards to the present study is that WM has demonstrated an influential role in each of the language learning processes (e.g., Juffs & Harrington, 2011; Kane & Engle, 2003; Mackey, Philp, Egi, Fujii & Tatsumi, 2002; Mackey & Sachs, 2012; Revesz, 2012). A summary of this data is provided in table 2.4 below.

**Table 2.4 The Role of Cognitive ID Factors in SLA Processing**

<table>
<thead>
<tr>
<th>Input processing</th>
<th>Noticing</th>
<th>Integrating new knowledge</th>
<th>Automatization</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Working memory</em></td>
<td><em>Working memory</em></td>
<td><em>Working memory</em></td>
<td><em>Working memory</em></td>
</tr>
<tr>
<td><em>Phonological short-term memory</em></td>
<td><em>Processing speed</em></td>
<td><em>Inductive ability</em></td>
<td><em>Perceptual speed</em></td>
</tr>
<tr>
<td><em>Phonological sensitivity</em></td>
<td><em>Metalinguistic awareness</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Inductive ability</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Metalinguistic awareness</em></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 2.3 Measuring Working Memory Capacity

Working memory capacity can be measured through a variety of tasks that vary in how much demand they put on the system. As a result, Shipstead, Lindsey, Marshall, & Engle (2014) argue that it is reasonable to understand that
different working memory tasks actually reflect different mechanisms of working memory, and thus provide slightly different perspectives on the cognitive processes that define this construct. They focus on three specific mechanisms of WM: (1) primary memory, (2) attention control, and (3) retrieval from secondary memory. In regards to working memory, primary memory is typically understood to be a type of limited capacity storage that can maintain between 3–5 items at a time (Cowan, 2001; Luck & Vogel, 1997; Rouder, Morey, Morey, & Cowan, 2011; Unsworth & Engle, 2007). In regards to WM, Shipstead et.al argue that attention control is critical when the environment or a memory search activates conflicting information. Engle’s (2002) executive attention account equates this mechanism of working memory capacity with the ability to use attention to select relevant information from the environment and to retain access to memories that are outside of conscious awareness (Kane, Conway, Hambrick & Engle, 2007). That is, WMC is seen to be driven by ability to focus on critical information and resist distractions to one’s attention. Finally, secondary memory refers to the ability to retrieve displaced information from longer-term storage (Unsworth & Engle, 2007). Secondary memory is important because regardless of the scope of a person’s primary memory, or attention control abilities, some information is displaced.

Shipstead et.al argue that each of these mechanisms is important to explaining individual differences in working memory capacity, and are reflected in three different WM tasks, (1) the visual arrays performance, (2) the running memory span and (3) the complex span. Of particular importance for the present
study is that although these tasks differ in their demands, they all predict reasonably similar variation in working memory capacity.

Shipstead and Engle (2013) describe the *visual arrays task* as a simple, but effective, computer-based measure of WM capacity. The task begins with the brief presentation of several randomly arranged objects such as colored shapes. This is followed by a brief blank screen, after which, the objects reappear. On this second presentation, one object has been circled, and the test-taker is required to decide whether this object has changed in any way, relative to the first presentation (e.g., Has the triangle’s color changed?).

Broadway (2008) describes *running memory span* tasks as being similar to simple span tasks in that there is no processing task in between to-be remembered items, but like complex span tasks, running span tasks have been proposed to uniquely tap into executive cognitive functions by virtue of a special procedure in which more items are presented than can be, or are instructed to be, remembered. For example, a participant may be required to report the *last four* items from a series of presentations, but the total items presented varies from just four to more than four.

Finally, in a *complex span task* (CST), subjects are asked to remember an item and also perform another attention demanding task at the same time. The most commonly used CSTs are the reading, operation and backward digit span tasks. All three of these are argued to be useful for measuring more than just phonological short-term memory, but also assess the capacity of complex verbal working memory including the functioning of the central executive, which is
responsible for regulating attention (Gathercole, 1999; Hale, Hoeppner & Fiorello, 2002).

The reading span task (RST) was the first complex span task used to study WMC and its relationship with higher-order cognition (Daneman & Carpenter, 1980). In a typical RST, participants are required to read aloud a series of sentences and try to remember the last word of each sentence for later recall. Participants may also be asked to remember random words that are presented in 'bold' lettering rather than just the last word of each sentence. The requirement to read aloud is in order to reduce the participant's ability to rehearse the to-be-remembered items. Because participants must remember a chosen word, as well as read aloud, the task requires participants to store information (words) over a short time span while at the same time engaging in a processing activity (reading). The idea was that this task measured the working memory system that gives rise to complex behavior better than a simple memory span task in which participants are required to remember items without a secondary processing task (e.g., word span).

The backward digit span task is also part of the Wechsler IV intelligence test for children (Gathercole & Alloway, 2008) which is hypothesized to be strongly related to general fluid intelligence (Engle, Kane and Tuholsky, 1999). In a typical backward digit span task, subjects are given a series of numbers, one at a time about three seconds apart. The total number in the series typically runs from two to eight. After the list of numbers has been given, subjects are asked to recite or write the numbers back in reverse order. For example, if a subject is
given the number series 3-7-2-9-5, they must write or say them backwards, 5-9-2-7-3.

Engle (2002) describes the operation-span task as being similar in format to the reading-span task. In a typical operation span, subjects read aloud a series of operation-word strings such as “Does 4/2+3=6? (yes or no) DOG.” They respond ‘yes’ or ‘no’ as to whether or not the equation is correct and then read the capitalized word aloud. After a set of two to seven such operation word strings, participants are required to write each of the words in the same serial order as presented.

It’s important to note that there is a distinction made between tasks that are designed to measure working memory and those designed to measure short-term memory (STM). This distinction is made between the complex span tasks described above, and simple span tasks (SSTs). Unsworth and Engle (2006) report that complex (working memory) span tasks have generally shown larger and more consistent correlations with higher-order cognition than have simple (or short-term memory) span tasks. They also argue that SSTs measure memory storage only versus CSTs that measure memory storage as well as higher-order cognition. Complex span tasks like simple span tasks, require participants to recall a set of items, often times in a specified order. However, CSTs differ from SSTs in that a separate processing activity is interwoven between the to-be-remembered items. CSTs came about in order to test a more dynamic memory system based on the Baddeley and Hitch’s (1974) model. Phonological loop capacity or short-term memory is often measured by tasks involving immediate
recall of a series of numbers (digit span) or words (word span) task (e.g., Baddeley, 2003; Morra, 1994; Osaka & Osaka, 1992). One of the most popular tests for measuring phonological short-term memory capacity is the non-word repetition test. Non-words are not real words, but they do conform to the phonotactic constraints of the language of the test taker. In this test, participants listen to, and then try to repeat non-words of varying length. Participants’ short-term memory capacity is then scored in terms of the non-word length, which is the highest number of syllables the participant could repeat accurately in at least 50% of the cases.

2.4 Vocabulary in SLA

In regards to SLA, the present study focused specifically on the acquisition of L2 vocabulary. This section highlights research related to vocabulary knowledge, acquisition and testing. Schmitt (2010) contends that the issues which attract the most attention in the field of vocabulary concern the nature of the lexis, its employment in language use, and the best ways of facilitating its acquisition. One of the most systematic explorations of the relationship between vocabulary knowledge and language proficiency occurred as part of the development of the DIALANG tests. DIALANG is an online diagnostic system designed to assess a person's proficiency in a foreign language. It was designed primarily for European citizens to assess their language abilities in adherence to Europe’s Common European Framework of Reference (CEFR) guidelines. CEFR guidelines are a widely recognized framework used to describe and measure the language proficiency level of a learner in a particular language.
Alderson (2005) worked with a research team in charge of the vocabulary section of the DIALANG. As part of their research, they compared scores on various vocabulary test scores with other language components of DIALANG. The result of the comparison showed strong correlations between vocabulary knowledge and the level of proficiency in other language skills, such as reading, listening and writing. Alderson writes “What the DIALANG analysis would appear to show is that the size of one’s vocabulary is relevant to one’s performance on any language test, in other words, that language ability is to quite a large extent a function of vocabulary size” (p. 88).

Laufer and Sim (1985) investigated the knowledge needed to successfully comprehend the English for Academic Purposes Cambridge Certificate English examination. Their study showed that vocabulary knowledge is the most important area of knowledge required for comprehension. It’s more important than knowledge of the subject and even syntactic knowledge. Later studies have estimated the percentage of vocabulary necessary for second language learners to understand written texts as ranging from 95% (Laufer, 1989) to 98% (Hu & Nation, 2000).

In a more recent study, Schmitt, N., Jiang, X., & Grabe, W. (2011) focused on the relationship between percentage of vocabulary known in a text and level of comprehension of the same text. In this study, 661 participants from eight countries completed a vocabulary measure based on words drawn from two texts. Participants were asked to read the texts, and then complete a reading comprehension test for each text. The results showed a relatively linear
relationship between the percentage of vocabulary known and the degree of reading comprehension. In terms of the percentage of known words necessary to accurately comprehend a text, results suggest that the 98% estimate of Hu and Nation is a more reasonable coverage target for readers of academic texts. Figure 2.1 below depicts the relationship between words known in a given text, and the level of comprehension of that same text.

![Figure 2.1 Linear relationship of words known and reading comprehension](taken from Schmitt, Jiang & Grabe, 2011, p. 29)

### 2.5 The Nature of L2 Vocabulary Knowledge

There are many challenges to overcome for researchers investigating the acquisition of vocabulary. First and foremost among these challenges is developing a clear understanding of the nature of vocabulary knowledge. Schmitt (2010) argues that vocabulary researchers need to be aware of the various characteristics of a lexical item in order to make conscious and principled decisions about which characteristics to control for in their studies. He notes that careful planning at the beginning stages of research design is the best insurance
against a study being later contaminated by unwanted lexical behavior which confounds interpretation of the results.

**FORM, MEANING AND USE**

One of the most important factors in researching vocabulary is for the researcher to have a clear understanding of what it means to know a word. Nation (2001) argues that words are not isolated units but rather they fit into many interlocking systems and levels. In order to clearly understand these interlocking systems, he developed a detailed understanding of the various aspects of word knowledge. This information is provided in Table 2.5 below.

Nation notes that it is possible to know the form of a word but have no concept of its meaning. It is also possible to be familiar with the form to have the appropriate concept but not connect the two. He goes on to claim that “the strength of the connection between form and meaning will determine how readily the learner can retrieve the meaning when seeing or hearing the word form, or retrieve the word form when wishing to express the meaning” (2001, p. 48).

VanPatten and Benati state that the “form-meaning connection refers to the correspondence between the formal properties of language and the meaning they encode” (2010, p. 86). Nation (2001) argues that this aspect of knowing a word tries to separate recognizing the form and knowing the meaning from being able to connect a particular form to a meaning. He argues that the strengthening of the form-meaning connection involves having to recall a meaning when seeing or hearing a particular word, or having to produce the spoken or written form when wanting to express a meaning.
### Table 2.5 The Three Areas of Word Knowledge

<table>
<thead>
<tr>
<th></th>
<th>Form</th>
<th>Written</th>
<th>Word Parts</th>
<th>Meaning</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken</td>
<td>R What does the word sound like?</td>
<td>P How is the word pronounced?</td>
<td>R What parts are recognizable in this word?</td>
<td>P What word parts are needed to express meaning?</td>
<td></td>
</tr>
<tr>
<td>Written</td>
<td>R What does the word look like?</td>
<td>P How is the word written and spelled</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word Parts</td>
<td>R What parts are recognizable in this word?</td>
<td>P What word parts are needed to express meaning?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Form and Meaning</td>
<td>R What does this word form signal?</td>
<td>P What word form can be used to express this meaning?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concepts and Referents</td>
<td>R What is included in the concept?</td>
<td>P What items can the concept refer to?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Associations</td>
<td>R What other words does this word make us think of?</td>
<td>P What other words could we use instead of this one?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammatical Functions</td>
<td>R In what patterns does the word occur?</td>
<td>P In what patterns must we use this word?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Collocations</td>
<td>R What words or types of words occur with this one?</td>
<td>P What words or types of words must we use with this one?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constraints on use</td>
<td>R Where, when and how often would we meet this word?</td>
<td>P Where, when and how often can we use this word?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Nation, 2001, p. 27) R=Receptive, P=Productive

In regards to concept and reference, Nation (2001) claims that this aspect of word knowledge involves having a clear idea of the underlying meaning running through its related uses, and also involves being aware of its range of uses. It is this knowledge which contributes to being able to understand a word when it is used in a new situation and being able to use a word in creative ways.

**Fluency versus Accuracy**

In addition to the systems and levels in the chart above, word knowledge can also be looked at in terms of fluency and accuracy. Milton (2009, p.16)
defines fluency as the ease and speed with which learners access and use the words they know, from simply recognizing a word and knowing how to use it (accuracy). In terms of accuracy, Schmitt (2010, p. 21) argues that knowledge of a lexical item ranges from zero to partial to precise. In other words, this means that all word knowledge ranges on a continuum, rather than being known versus unknown. For the present study, it could mean that learners have learned some of the meanings related to a polysemous form, but not all of them.

**BREADTH VS. DEPTH OF VOCABULARY KNOWLEDGE**

Milton (2009) argues that a further distinction in vocabulary knowledge can be made by separating vocabulary knowledge into three distinct categories. These three categories are (1) depth of vocabulary knowledge, (2) width of vocabulary knowledge and (3) breadth of vocabulary knowledge.

Milton argues that depth of vocabulary knowledge refers to the relationship between the various forms and meaning components of a word. He further defines it as “the wide variety of word characteristics, including the shades of meaning a word may carry, its connotations and collocations, the phrases and patterns of use it is likely to be found in, and the associations the word creates in the mind of the user” (2009, p. 149). He argues that this understanding of vocabulary depth implies that a word will be linked to other words and ideas in the lexicon. Milton also argues that the concept of vocabulary depth is even more difficult to determine than the concept of vocabulary breadth. This is due to the fact that it might involve knowledge of word associates, collocation, colligation or word function. By comparison, the second category, width of
vocabulary knowledge, primarily refers to the number and degree of the relationships between the word and other words in the lexicon.

In regards to the third category, breadth of vocabulary knowledge, Anderson and Freebody (1981) also argues that a distinction needs to be made between breadth of vocabulary knowledge and depth of vocabulary knowledge. Breadth of knowledge refers to the number of words that are known, and depth of knowledge refers to what a learner knows about these words. Nation (2001) makes a similar distinction between how well a particular word is known, depth of knowledge, with how many words are known, breadth of knowledge. The majority of L2 vocabulary studies have focused on breadth of word knowledge due to the great interest in building up a large lexicon in the mind of the L2 learner.

Schmitt (2010) notes that in addition to needing a large vocabulary size to function in a language, a person must also know a great deal about each individual lexical item in order to use it well. This often is referred to as the quality or depth of vocabulary knowledge, and it is as important as vocabulary size. In other words, depth of word knowledge should be seen as equally important as breadth of word knowledge, as it allows the learner to convey numerous concepts using a single form. These are very important distinctions for the present study as the focus is on depth and width of word knowledge rather than breadth of word knowledge which is the focus of a majority of vocabulary studies.
THE RECEPTIVE / PRODUCTIVE DISTINCTION

Researchers in L2 vocabulary acquisition also make a clear distinction between receptive and productive knowledge. Nation argues that essentially “receptive vocabulary use involves perceiving the form of a word while listening or reading and retrieving its meaning. Vocabulary use involves wanting to express the meaning through speaking or writing and retrieving and producing the appropriate spoken or written word form” (2001, p. 24-25). Nation (2001, p. 26) applies the scope of the receptive/productive distinction to the word underdeveloped. He writes that knowing the word underdeveloped involves:

1. Being able to recognize the word when it is heard or seen.
2. Recognizing that it is made up of the parts, under-, -develop-, and -ed and being able to relate these parts to its meaning.
3. Knowing that underdeveloped signals a particular meaning.
4. Knowing its meaning dependent on the context in which it is used.
5. Knowing the concept behind the word which will allow understanding in a variety of contexts.
6. Knowing that there are related words like overdeveloped, backward and challenged.
7. Being able to recognize that words such as territories, areas and ideas are typical collocations.
8. Knowing how often underdeveloped is used.
From the point of view of productive knowledge and use, knowing the word *underdeveloped* involves:

1. Being able to correctly pronounce it.
2. Being able to correctly write it.
3. Being able to form it with the right combination of morphemes.
4. Being able to produce the word in various contexts to express the range of meanings of underdeveloped.
5. Being able to produce synonyms and antonyms.
6. Being able to use it correctly in an original sentence.
7. Being able to produce words that commonly occur with it.
8. Being able to decide to use or not use it depending on the sociolinguistic constraints of the language. (Developing is slightly more acceptable than underdeveloped in many situations.)

There are also differences between receptive and productive knowledge in terms of the ease with which learners develop these two different modalities of knowledge. Milton (2009) claims that productive vocabulary knowledge is generally less than receptive, and he places the estimate at around 50 to 80% of receptive knowledge. Waring (1997) suggests that there are a number of factors outside of the learners knowledge of the language that affect the differences in receptive and productive vocabulary knowledge. He argues that a listener or reader can employ a number of contextual clues and other information to help reach meaning. However, when called upon to produce speaking or writing
learners may be under time pressure, and will be lacking the cues for drawing meaning that are available in receptive language.

Milton (2009) argues that the receptive and productive division of lexical knowledge is helpful in two ways; as a means to distinguish the ability to recognize some aspect of a word, and the ability to both recognize and use a word. Schmitt (2010) equates the idea of reception with language comprehension which defines as learners demonstrating how well they understood words provided to them in the context of a written text. In regards to production, some researchers define it as language use. Schmitt gives the example of language use as having a measurement of the vocabulary produced in a language task designed to elicit the targeted lexical items.

**POLYSEMOUS FORM-MEANINGS RELATIONSHIPS**

The present study investigated the acquisition of L2 polysemes, specifically L2 Spanish prepositions. These are lexical items that have multiple interrelated meanings or senses, and acquiring them requires the development of a much greater depth of vocabulary knowledge. Schmitt (2010) notes that knowledge of rarer meanings or senses can indicate a more comprehensive knowledge of a lexical item. Taylor defines polysemy as the, “association of two or more related senses with a single linguistic form” (1995, p. 99). Polysemy reflects the possibility of language change, as well as “the coexistence of more general and more specific, more literal and more figurative, more ancient and more innovative meanings” (Gragg, 1984, p. 140). In most cases, there is one
sense of the word which appears to be dominant, and the other senses appear to be derived from this sense.

Nation (2001, p. 49) notes that a “feature of words that is especially striking when they are looked up in a dictionary is that they can have a lot of different meanings.” He notes that this is particularly true of words that are used more frequently in a language. For example, the English word run which is a very high frequency word, has over 30 distinct senses or meanings listed in the Oxford English Dictionary. The benefit of knowing a polysemous word such as run is that the learner can convey many different concepts with a single form.

However, Nation notes that when you examine the range of meanings which may be associated with a single form, you may notice that sometimes they are completely unrelated with each other. An example he gives is the word bank, which may refer to the bank of a river or a financial institution. Words such as this, which have the same form but have completely unrelated meanings are called homonyms. Homonyms can be either homographs (words with identical written forms) or homophones (words with identical spoken forms). The various meanings of homonyms need to be counted and since their various meanings are unrelated, they must be learned separately, one at a time. By comparison, cases of polysemy involve words with multiple interrelated meanings.

Touplikioti (2009) cites catch as an example of polysemy in English. She argues that there are at least seven distinct concepts that can be expressed by catch. They are (1) getting hold of, (2) capturing, (3) perceiving, (4) becoming affected by, (5) overtaking, (6) meeting and (7) becoming alight. According to
Touplikioti, all seven of these meanings can potentially be introduced in a single lesson via Cognitive linguistics-based instruction (CLBI). CLBI will be introduced later in the vocabulary instruction section.

Nagy (1997) argues that there are two main ways that language users can handle words with multiple meanings. First, they may have a permanent internal representation of each related meaning and simply select the appropriate sense of the word stored in the brain. He refers to this process as sense selection. Second, language users may have underlying concept that is appropriate for the range of meanings with which the word is used. When language users encounter the word, they have to work out during the comprehension process what particular real-world items the word is referring to. He refers to this process as referenced specification.

Ruhl (1989) argues that rather than following dictionaries in seeing words as having multiple meanings we should assume that each word has a single parent lexical meaning. There are two major sources of meaning when we comprehend a word in context: (1) its inherent lexical meaning which refers to its meaning when used in isolation, and (2) inferential meaning which we infer when used alongside other words in the immediate context, and from our knowledge of the world.

The challenge for researchers is the fact that the lexical meaning may be very abstract. This is especially true of polysemes where a word has more than one sense. We should assume that the senses are related to each other by general rules that apply to other words. These rules include the idea that words
can have a range of senses concrete and abstract and these differences in concreteness and abstractness are inferred from the context. Ruhl produces evidence in support of his position by examining many examples of use and shows that apparent variations in meaning can be accounted for by inferential meaning that is stable and can be seen running through all senses of the word.

2.6 L2 Vocabulary Acquisition

For the present study, it was necessary to examine the major factors that can affect the acquisition of L2 vocabulary. This is especially true since the target structures, polysemes, are particularly challenging for learners of L2 vocabulary.

Vocabulary Learning Burden

There are a number of factors that can affect how easily a word is learned in an L2. This concept is referred to as the vocabulary learning burden, and it is important for the present study as polysemes bear a heavy learning burden. Nation (2001) claims that each word in the lexicon of language bears a unique learning burden for the learner. He defines learning burden as “the amount of effort required to learn it” (p.23). Nation (1990) argues that the learning burden of a word decreases the more a word represents patterns and knowledge that the learner is already familiar with. Kwon (2005) advances a case for two additional factors that affect the learnability of words. One is semantic complexity and, in particular, a hierarchy of semantic complexity where words with multiple meanings are acquired later than those with fewer meanings. The
second is language transfer, where forms and structures in the L1 are likely to impact the performance in the L2.

In regards to learning burden, Schmitt (2010) notes that there are several aspects of meaning that need to be given attention in vocabulary research. The first is a concept that he refers to as ‘imagineability’. Imagineability refers to how easy it is to imagine a concept. De Groot (2006) refers to the concept of imagineability as its level of concreteness. Concreteness is a variable that expresses the degree to which a word, or rather the entity the word refers to, can be experienced by the senses. Studies have shown that the degree of concreteness greatly affects how well words are learned (DeGroot 2006; Ellis and Beaton, 1993). The more concrete a word is the easier it is to learn. Therefore, Schmitt argues that when comparing the acquisition of different groups of words, it is necessary to make sure that each of the groups consists of equivalent degrees of concreteness. This will ensure that any advantage in learning one group of words will not be due to the fact that those particular lexical items were more concrete than another group of words. For the present study, there was one set of words, all L2 Spanish prepositions, each with equal levels of concreteness in order to lessen any advantage of learning one word over another. For the present study, the main way the level of concreteness was controlled for was through controlling the parts of speech and levels of abstractness of the words included in the target preposition list.
THE INCREMENTAL NATURE OF L2 VOCABULARY ACQUISITION

Schmitt (2010, p. 20) argues that vocabulary learning is incremental because some aspects of word knowledge are learned before others. Schmitt (1998) found evidence for partial to precise degrees of knowledge of lexical items in a study made of advanced L2 learners at university level. In this study, Schmitt followed the mastery of a number of word knowledge aspects for eleven words over the course of an academic year. At the end of the study, the students rarely knew all the words’ derivational forms or meaning senses. They only knew the word class of the stimulus word in one derivation, but rarely all four main forms nouns, verbs, adjectives, adverbs. Likewise, they normally knew the core meaning sense, but almost never all of the possible senses.

Schmitt (2010) reports that through examination of a range of vocabulary studies that some word knowledge aspects will reach a productive level of mastery sooner than others. For example, advanced learners may be able to produce the spelling of the base form of target words that may not be able to produce the words derivative forms and meaning senses (Schmitt 1998).

LEARNING NEW FORM-MEANING RELATIONSHIPS VS. ‘RELABELING’

Schmitt (2010) makes a distinction between learners simultaneously learning a new concept and its L2 form, and simply learning a new L2 form for a concept that already exists in the mind of the learner. He argues that for adult learners the majority of L2 forms are linked to concepts that they already have in their L1. Schmitt refers to learning tasks where students learn to attach the correct L2 form to a known concept in their L1 as relabeling. He argues that it is
important to control for the ‘concept plus label’ versus ‘relabeling only’ items in target vocabulary. This is due to the fact that relabeling is simpler than a learning task that involves both learning the label and the concept.

Milton (2009) also describes the task of learning a foreign language as being far less a question of developing a whole new structure for a new lexicon, and more about learning to relabel the concepts and connections that already exist in the lexicon, so that they can be used for communication. In regards to the present study, the concept of relabeling was very important. This is due to the fact that all the concepts associated with a single polysemous form in an L2, may also be concepts that can be conveyed with the lexicon of the L1; however, often there is a separate form for conveying each of those concepts. In other words, when a learner is tasked with learning a polysemous L2 word, the conceptual challenge is learning to associate various concepts that are considered separate in the L1, with a single form in the L2. Robinson (2001, p. 51) reports that from an Indonesian point of view, ‘fork’ is defined mainly by its function – something to push food on to your plate. From an English point of view, fork is defined by its shape. Treating the relationship between a form and its various meanings in one language as if they are the same for a near equivalent form in another language can obscure these differences. It also adds to the learning burden of these lexical items as new associations and collocations have to be learned.
2.7 L2 VOCABULARY INSTRUCTION

There are numerous vocabulary teaching techniques employed by L2 teachers, but there are three important processes that should be kept in mind with any vocabulary teaching activity (Nation, 2001). These processes are noticing, retrieval and creative use.

The first process, noticing, involves making learners aware of the usefulness of a given lexical item (e.g., McLaughlin, 1990; Schmidt, 1990). Nation argues that noticing may be affected by several factors, including the salience of the word in the input, and/or any previous contact the learner has had with the word in being aware of the gap that the word can fill in their interlanguage. He claims that teaching activities that involve noticing would include having learners look up a word in the dictionary, or intentionally study its definitions, guessing meaning from context or simply having the word explained to them.

Nation argues that the second process, retrieval, involves having learners retrieve the word from memory through tasks that involve both receptive and productive knowledge. Receptive retrieval involves recalling a word’s meaning when encountered in either aural or reading input. Productive retrieval involves retrieving and then producing the word in either spoken or written form when it is necessary to communicate its meaning.

Nation argues that the third process, creative use, involves having learners use the word in ways that force them to reconceptualize the extent of
their knowledge of the word. This is commonly done through metaphorical extensions of already known meanings.

Glossing words in a written text is one common way for presenting new vocabulary. A gloss is a brief synonym, either in the L1 or L2, which is provided with the text that they are reading. Nation (2001) argues that there are certain attractions to glossing. First, it allows text to be used that may be too difficult learners to read without glosses. This means that unsimplified and unadapted texts can be used. Nation contends that first language translations are an efficient and effective means for teaching and testing of vocabulary.

2.8 Previous Cognitive Linguistics-based Research

In regards to words that are polysemous, Nation (2001) notes that the interesting question is which process accounts for most of the allocation of meaning to a word. Do learners have to develop strategies for storing multiple meanings for a word, or do they need to learn how to use a single underlying meaning to work out a particular related sense? Should teachers show the meaning underlying different senses of a word or should the teacher treat the different senses as different items to be learned one at a time? Nation argues that the best way to explain the meaning of a polyseme is to define the word by looking for the concept that runs through all its senses or uses, thereby reducing the learning burden.

Cognitive linguistics-based instruction is designed to start the process of learning a polyseme by focusing on a common conceptual core that runs through all its other meanings. Touplikioti (2007) describes CLBI as a motivated
approach to vocabulary teaching that provides a systematic model for linking the multiple meanings of words. This is in contrast to translation-based instruction (TBI) that she defines as instruction based on the technique of memorization and repetition where students are simply guided to learn lists of words without any guidance as to how their various meanings interrelate and behave.

Touplikioti (2009) cites catch as an example of polysemy in English. She argues that there are seven distinct concepts that can be expressed by catch, and all seven of these meanings can potentially be introduced in a single lesson via Cognitive linguistics-based instruction (CLBI). Touplikioti provides a schematic demonstration of how this would be done in a lesson in which the word catch would be taught (see appendix C). What this schematic shows is that the various meanings or senses associated with catch can all be related back to one core meaning - ‘get a hold of’. This one core meaning is in turn shown to interrelate to the other six core meanings through a visual schematic. The argument behind CLBI is that through this visual presentation, learners more quickly learn and begin to understand all the various meanings of a polyseme by seeing that they are not an unrelated list of discrete meanings, rather, they are all conceptually interrelated.

There have been a number of studies that have attempted to compare cognitive linguistics-based approaches to polysemy instruction that employ image-schemas with translation-based approaches that employ translation and memorization. Khodadady and Khaghaninizhad (2011) reports a comparison of CLBI versus TBI in the teaching of the polysemous French verb arriver and the
polysemous French preposition *sur*. The participants in their study were 49 L1 Farsi speakers who had studied French for at least six semesters at a foreign language institute. The results show that cognitive linguistics-based approaches tend to be as effective as TBI for acceptability judgment tasks, but far more effective for productive knowledge tasks.

Touplikioti (2007) reports a comparison of CLBI versus TBI in the teaching of the polysemous English verbs *make* and *do* to L1 Greek speakers who were at the low intermediate level in English at the start of the study. Participants were given a pre-test to measure their knowledge prior to instruction followed by a post-test. The results show greater gains pre-test to post-test for participants taught under CLBI compared to participants taught under TBI.

Makni (2013) compared CLBI and TBI approaches in the teaching of the English polysemes *hand, break, head, over, burn, push, beyond* and *root*. The participants in her study were 40 L1 Arab speakers who were at the low-intermediate level in English at the start of the study. These participants were divided into an experimental group and a control group, each made up of 20 participants. Makni reports that statistically analyzed results confirm the primacy of techniques inspired by cognitive linguistics approaches versus those based on translation in learning polysemous words. She argues that “such findings give pedagogic support to the tenets of cognitive linguistics and prototype theory within cognitive linguistics (2013: 190).”

Morimoto and Loewen (2007) reports on the effectiveness of two types of vocabulary instruction, image-schema-based instruction (ISBI) and translation-
based instruction (TBI) in the acquisition of two L2 English polysemes (i.e., break and over). Participants in the study were fifty-eight L1 Japanese speakers in a high school English language program. They were divided into two treatment groups (ISBI and TBI) and a control group. The results showed that ISBI was equally effective as TBI for both acceptability judgment test and production test scores, except in the case of acceptability judgment test for the preposition over where ISBI was significantly more effective than TBI. As a result, they argue that image-schema from the field of cognitive semantics can serve as a highly effective pedagogical technique in teaching L2 polysemes.

Beréndi, Csábi, and Kovecses (2008) reports on three separate experiments in which cognitive linguistic principles were employed to teach a range of vocabulary including idioms and metaphors. Participants in the first experiment were two parallel groups of 13 students enrolled in a Budapest secondary school who were assessed at level B1 on the CEFR scale. The first experiment focused on the teaching of the English polysemes hold and keep as well as idioms related to anger. The results of this study support the idea that the retention of the multiple senses and uses of a polyseme can be enhanced by employing insights from cognitive semantics. The second experiment involved first year English majors at a college in Hungary who were assessed at the B2 level according to the CEFR guidelines. In this experiment, participants were separated into a control group and an experimental group, and were then taught 22 idioms using either translation-based instruction or cognitive-linguistics based instruction involving explicit instruction of conceptual metaphors. The
The experimental group outperformed the control group by 87.5% to 78% (p = .03, Mann-Whitney U test, one-tailed) in their understanding of the idioms. They argue that this shows that explicating the conceptual metaphors behind idioms can help learners understand them better. The third experiment also involved first year English majors in a Hungarian university with a slightly lower level of proficiency in English than participants in experiment 2. In this experiment participants were taught metaphors, however; there was not a statistically significant result between the control and experimental groups. They reported an especially high standard deviation (SD 6.9) in the scores of the experimental group who were taught according to cognitive linguistic principles. They argue that this high SD suggests that there may be considerable individual differences when it comes to successfully applying cognitive-linguistic techniques autonomously.

Finally, Lam (2009) compared the effectiveness of teaching the Spanish prepositions por and para according to translation-based techniques versus cognitive linguistics-based techniques. Participants in the study were two groups of intermediate-level university Spanish students. The most significant results of the study were that participants who were taught according to cognitive linguistics-based principles scored significantly better on the post-delayed test than participants instructed according to principles of translation and memorization. This supports the idea that CLBI does a better job of developing long-term knowledge of these polysemous prepositions.
**TRANSLATION-BASED INSTRUCTION**

Unlike Cognitive linguistics-based instruction, Translation-based instruction involves having learners memorize a list of the various ways in which an L2 polyseme can be translated into their L1 without any understanding or attempt to find a unifying conceptual core. TBI requires learners to memorize a list of all the definitions that are possible for an L2 word, but with the mindset that the various meanings are arbitrary and unrelated. TBI is closely linked with the process of *noticing* previously discussed. The key difference between CLBI and TBI is not in the list of meanings applied to a target L2 form, rather, the way in which that list of meanings is presented to the learner.

**2.9 MEASURING L2 VOCABULARY KNOWLEDGE**

One of the greatest challenges of L2 vocabulary acquisition research is the development of valid and reliable tests to measure the aspect of vocabulary knowledge in question. Milton argues that “language knowledge is not a directly accessible quality and we rely on learners to display their knowledge in some way so it can be measured” (2009, p. 6). The challenge for vocabulary researchers is devising effective ways for learners to display their knowledge. Milton contends that a single test cannot possibly measure every aspect of word knowledge since it is impossible to test every aspect of word knowledge simultaneously. However, a lack of knowledge in one aspect of word knowledge can adversely affect performance measures in other aspects of word knowledge as well.
There are two very common concepts important in testing, including the creation of a vocabulary tests. One is reliability and the other is validity. Reliability is the ability of a test to measure something consistently and accurately. Validity addresses the issue of whether it measures what it is supposed to measure. Milton (2009) separates validity into two different types. Content validity takes into account whether a test has the necessary and appropriate content to measure what it is supposed to. Construct validity is often closely associated with content validity, but it also takes into account whether or not the test actually measures the content or skill it is supposed to. These common concepts will be taken into account later in the development of the actual testing instruments.

Nation (2001) cites a study by Ellis and Beaton (1993) in which they always tested receptive knowledge before productive knowledge, and Nation argues that by testing in this order, productive scores would have been given a boost. He therefore argues that a productive test should always be given first, followed by receptive tests. The development of vocabulary test should first start with an understanding of the aspect of word knowledge that will be tested. Table 2.6 below contains a detailed summary of the aspects of word knowledge that a researcher should be aware of when designing vocabulary tests.
Table 2.6 Aspects of Word Knowledge for Testing

<table>
<thead>
<tr>
<th>Form</th>
<th>spoken</th>
<th>Receptive</th>
<th>Can the learner recognize the spoken form of the word?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Productive</td>
<td>Can the learner pronounce the word correctly?</td>
</tr>
<tr>
<td></td>
<td>written</td>
<td>Receptive</td>
<td>Can the learner recognize the written form of the word?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Productive</td>
<td>Can the learner spell and write the word?</td>
</tr>
<tr>
<td></td>
<td>word parts</td>
<td>Receptive</td>
<td>Can the learner recognize known parts in the word?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Productive</td>
<td>Can the learner produce appropriate inflected and derived forms of the word?</td>
</tr>
<tr>
<td>Meaning</td>
<td>form and meaning</td>
<td>Receptive</td>
<td>Can the learner recall the appropriate meaning for this word form?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Productive</td>
<td>Can the learner produce the appropriate word form to express this meaning?</td>
</tr>
<tr>
<td></td>
<td>concept and referents</td>
<td>Receptive</td>
<td>Can the learner understand a range of uses of the word and its central concept?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Productive</td>
<td>Can the learner use the word to refer to a range of items?</td>
</tr>
<tr>
<td></td>
<td>associations</td>
<td>Receptive</td>
<td>Can the learner produce common associations for this word?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Productive</td>
<td>Can the learner recall this word when presented with related ideas?</td>
</tr>
<tr>
<td>Use</td>
<td>grammatical functions</td>
<td>Receptive</td>
<td>Can the learner recognize correct uses of the word in context?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Productive</td>
<td>Can the learner use this word in the correct grammatical patterns?</td>
</tr>
<tr>
<td></td>
<td>collocations</td>
<td>Receptive</td>
<td>Can the learner recognize appropriate collocations?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Productive</td>
<td>Can the learner produce the word with appropriate collocations?</td>
</tr>
<tr>
<td></td>
<td>constraints on use</td>
<td>Receptive</td>
<td>Can the learner tell if the word is common, formal, infrequent, etc.?</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Productive</td>
<td>Can the learner use the word at appropriate times?</td>
</tr>
</tbody>
</table>

(Nation, 2001, p. 347) R=Receptive; P=Productive

Nation argues for the benefit of multiple-choice item vocabulary tests.

Nagy, Herman and Anderson (1985) argues that it is possible to design multiple-choice items of different degrees of difficulty by varying the closeness of meaning
between the distracters and the correct answers. A distracter is an answer in a multiple-choice question that closely approximates the correct answer, and by design, forces the learner to clearly think and identify the best choice. One of the challenges of multiple-choice test questions is how to overcome the possibility of participants simply guessing at the answer. Paul, Stallman and O’Rourke (1990) claim that guessing is not a major problem with multiple-choice items and that learners’ responses are generally not random but largely driven by some knowledge of the word.

In regards to polysemes, Nation (2001) argues that an advantage of multiple-choice items is that they can focus on particular meanings where words have more than one meaning. He goes on to note that it is important to be consistent about the closeness of the relationship between the distracters and the correct answers for form and meaning as this has a major effect on the difficulty of the item. Nation also suggests the use of matching as one way of reducing the amount of work involved in making multiple-choice test. He developed a list of tests item types according to the aspects of word knowledge, both receptive and productive. These test type items can be found in Table 2.7 below. Nation arranges the test type items in Table 2.7 according to the parts of Table 2.6 which is itself based on Table 2.5.

Nation argues that in experimental research it is very useful to test the same word in several different ways. He claims that by doing this, researchers are able to give a ‘strength of knowledge’ score for each word. McKeown, Beck, Omanson and Pople’s (1985) found that if only one measure of vocabulary
knowledge was used, important differences from the effects of the treatments might not have been revealed. As a result, Nation developed a table of a test formats classified according to three distinction affecting difficulty (Table 2.8).

**Table 2.7 Test Type Items**

<table>
<thead>
<tr>
<th>Test Type</th>
<th>R</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spoken form</td>
<td>Word or sentence dictation / Hear the word and choose the L1 translation.</td>
<td>Reading aloud / crued oral recall.</td>
</tr>
<tr>
<td>Written form</td>
<td>Say these written words.</td>
<td>Word or sentence dictation.</td>
</tr>
<tr>
<td>Word parts</td>
<td>Break the word into parts.</td>
<td>What do you call someone who paints houses?</td>
</tr>
<tr>
<td>Form and meaning</td>
<td>Translate these words into L1.</td>
<td>Translate these words into L2.</td>
</tr>
<tr>
<td>Concept and referents</td>
<td>Translate the underlined words into L1.</td>
<td>'It was a hard frost.'</td>
</tr>
<tr>
<td>Associations</td>
<td>Choose the words that you associate with this word.</td>
<td>Choose the words to translate this L1 word.</td>
</tr>
<tr>
<td>Grammatical functions</td>
<td>Is this sentence correct?</td>
<td>Use this word in a sentence.</td>
</tr>
<tr>
<td>Collocations</td>
<td>Is this sentence correct?</td>
<td>Produce collocations to go with this word.</td>
</tr>
</tbody>
</table>

(Nation, 2001, p. 355) R=Receptive; P=Productive

For the present study, it was necessary to measure learners’ L2 knowledge through tests that translated all words back into their L1 except the original L2 form of the target words (i.e., *por, para, a, en*). Basically, every word except for the target words were translated back into English in order to make it possible to test learner’s depth of vocabulary knowledge. Nation (2001) argues that there are several reasons why glossing or translation is useful in second language acquisition and research. Glossing and translation allow more difficult text to be read, provide accurate meanings for words that might not be known or
guessed correctly, and allows readers to continue uninterrupted. Other research supports this view as well (Goldstein 1992 and Haynes 1993)

**TABLE 2.8 EIGHT VOCABULARY TESTING FORMATS**

<table>
<thead>
<tr>
<th></th>
<th>Recognition</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Receptive</strong></td>
<td>Imprecise</td>
<td>Sensitive multiple-choice, for example, <em>fertilizer</em></td>
<td>a. growing plants</td>
<td>b. medicine</td>
<td>c. history</td>
</tr>
<tr>
<td></td>
<td>Precise</td>
<td>Non-sensitive multiple-choice, for example,</td>
<td>There was no <em>response</em></td>
<td>a. movement</td>
<td>b. answer</td>
</tr>
<tr>
<td><strong>Recall</strong></td>
<td>Imprecise</td>
<td>Recalling a related meaning</td>
<td>Does this word remind you of anything?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Precise</td>
<td>Meaning recall</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Productive</strong></td>
<td>Imprecise</td>
<td>Sensitive multiple-choice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Precise</td>
<td>Non-sensitive multiple-choice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Recall</strong></td>
<td>Imprecise</td>
<td>Cued recall, for example, <em>an additional part suppl</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Precise</td>
<td>Form recall</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Schmitt, 2010) notes that a third gap in our knowledge of vocabulary studies is how to best measure the various word knowledge aspects. He contends that in order to understand how a learner acquires vocabulary, it is necessary to develop measures for the different aspects of word knowledge. He claims that it is really not possible or even desirable to attempt to measure all aspects of word knowledge in a single study, but rather researchers should focus on developing measurements of specific aspects of word knowledge. The challenge facing the current study was the lack of accepted measurement
instruments for certain aspects of word knowledge such as register, collocation, frequency use and associations.

Schmitt (2010) reiterates the fact that vocabulary knowledge is multifaceted and contains a number of interrelated yet severable aspects. The word knowledge framework developed by Nation (2001) is helpful for understanding the various aspects of vocabulary knowledge. However, because vocabulary knowledge involves several distinct aspects, it is virtually impossible to measure all aspects of word knowledge with a single test instrument. Schmitt argues for three reasons why a single vocabulary test is insufficient. The first and most important fact is that many of the word knowledge aspects do not yet have accepted methods of measurement. He argues that this necessitates that the researcher develop his or her own new methodology. The second reason is connected to time. Instrument that measures all aspects of word knowledge would be extremely time-consuming and difficult to manage. The third reason is the fact that various types of word knowledge are interrelated necessitating the organization of the different word knowledge tests in such a way as to not affect the other tests.

Schmitt (2010) argues that measurement problems often stem from an unclear conceptualization of the vocabulary knowledge aspect to be measured. Henriksen (1999) gives three components of vocabulary knowledge. They are (1) partial to precise knowledge of word meaning, (2) depth of knowledge of the different word knowledge aspects and (3) receptive knowledge verse productive knowledge.
MEASURING RECEPTIVE VS. PRODUCTIVE L2 VOCABULARY ACQUISITION

Read (2000, p. 155) argues that “the difficulty with measurement stems from the lack of an adequate conceptual definition of the difference between reception and production.” He claims that one source of confusion about the distinction between receptive and productive measures is the fact that many researchers use two different definitions of reception and production. They also use these definitions interchangeably. Because these definitions can greatly affect assessment, Read found it necessary to explain each one using other terms.

Schmitt (2010) claims that discrete in context independent test formats tend to focus on what Read refers to as recognition and recall. Read (2000) states that recognition involves presenting a word to test takers and having them demonstrate their understanding of its meaning. By comparison, in the case of recall they are given some sort of stimulus designed to elicit the target word from the memory. Read gives a simple example of this where recognition means that the subject gives the L1 translation of the L2 word, and recall refers to the reverse process; they give the L2 word in response to the L1 translation. He summarizes the difference as basically being able to recognize the word when you are presented with it and being able to recall it when prompted to do so.

For example, some researchers equate the idea of reception with language comprehension. Schmitt (2010) argues that comprehension could involve learners demonstrating how well they understood words provided to them in the context of a written text. In regards to production, some researchers define
it as language use. Schmitt gives the example of language use as having a measurement of the vocabulary produced in a language task designed to elicit the targeted lexical items.

Milton (2009) argues that measuring the productive vocabulary that learners possess poses methodological problems for the researcher in how best to capture this variable. He claims that the problem is not in deciding how to devise a test but rather using the best approach that other researchers have used. He claims that no single definitive method for measuring productive vocabulary knowledge has emerged. Milton organizes the measuring of productive vocabulary knowledge into four approaches: (1) translation and elicitation methods, (2) statistical analysis of free production in speech or writing, (3) association tests and (4) measures of automaticity.

Milton (2009) details measuring productive vocabulary using translation and elicitation. He describes a basic example of this type of productive vocabulary test as providing a list of words in the learner’s L1 and asking them to provide a foreign language equivalent. This form of testing is not favored among many researchers, particularly those in favor communicative approaches to language acquisition. In order to avoid mechanical nature of translation tests, researchers interested in measuring productive vocabulary knowledge have to find a way to elicit the words they are interested in testing. Laufer and Nation (1995) constructed a productive measure of vocabulary knowledge based on Nation’s (1990) receptive vocabulary placement test. The test procedure
presents the learner with a series of sentences with a missing word. The learner is then called upon to fill the gap with the appropriate word.

**Measuring Breadth vs. Depth of L2 Vocabulary Acquisition**

Milton (2009) notes that one of the difficulties in testing vocabulary is that even though it is a single word that is being tested, a knowledge of a wider range of words is actually needed to be successful. This was an especially important point for the present study, as the target words, polysemes, require an even greater knowledge of a large number of L2 words than to words with a limited number of possible L2 senses. In other words, it takes a breadth of vocabulary knowledge in order to adequately test a learner’s depth of vocabulary knowledge. Milton argues that confusion and ignorance about the other L2 words in the test may cause learners to misrepresent their actual knowledge of the targeted words.

One common measure of breadth of vocabulary knowledge is the lexical sophistication free language production task. Milton describes the measure of lexical sophistication as a “calculation of the proportion of infrequent words in a text” (2009, p. 131). In other words, participants are asked to produce a lengthy stretch of written or spoken discourse, from which a specific measure of infrequent words can be pulled (see Coxhead, 1998; West & West, 1953 for example word lists based on frequency of use).

A common measure of depth of vocabulary knowledge is the Word Association Task (WAT) which measures learner’s knowledge of the collocations of a lexical item. WAT’s have been used for over a hundred years in the field of
psychology (e.g. Galton 1879; Jung 1910). These early studies used WAT’s to evaluate how individuals conceptualized via lexical associations. Responses to these tasks were then examined for possible signs psychological abnormality (e.g., Kent and Rosanoff 1910). Recently, WAT’s have been adopted by SLA researchers interested in better understanding the development and organization of the L2 mental lexicon and L2 proficiency (e.g. Meara, 2009; Wolter, 2002). Milton (2009, p.141) describes a typical WAT as getting a learner to focus on a specific stimulus word, such as white, and then asking them to produce a word in response. Learners are instructed to write down the first word that comes to mind.

**Measuring the Acquisition of L2 Polysemes**

Measuring the acquisition of polysemous L2 vocabulary poses unique challenges for researchers. Most L2 vocabulary testing methods are designed for measuring learner’s breadth of vocabulary knowledge rather than depth. Measurements of vocabulary breadth usually involve having learners demonstrate the total number of one-for-one form-meaning connections a learner can make within a given set of vocabulary. In other words, the learner is simply required to provide one meaning, usually a core or primary meaning, for each new form. However, when measuring L2 learner’s knowledge of polysemous L2 vocabulary it necessarily requires that they demonstrate a depth of knowledge of all the various meanings that these L2 forms can convey. This is much more demanding since in order to demonstrate a depth of vocabulary knowledge, learners must already have a very large breadth of vocabulary knowledge. In
other words, attempting to have a low novice learner demonstrate a depth of vocabulary knowledge of a polysemous word is virtually impossible because they simply do not possess the body of L2 vocabulary necessary to translate all of its interrelated meanings.

Vocabulary tests that measure breadth of vocabulary knowledge require that the researcher embed the L2 polysemous form into an L2 sentential context. Learners will be unable to render an accurate translation of the L2 polysemous form if they do not know the meaning of all the other words that surround it in the sentence. In other words, the meanings of polysemous words are dependent on the sense in which they are being used in a given context. Without knowledge of all the other words that provide the context, it is impossible to render the correct meaning.

2.10 RESEARCH QUESTIONS AND HYPOTHESES

The review of the literature provided above produces several lines of potential research into the acquisition of L2 vocabulary. As it is not possible to address all aspects of L2 vocabulary acquisition in a single study, so the present study narrowed the focus to three specific gaps in current research literature related to the acquisition of L2 polysemes. These three gaps produce three research questions as well as three related hypotheses all related to the acquisition of productive knowledge of L2 Spanish polysemous spatial prepositions by adult L1 English speakers. Taking into account current literature related to the acquisition of L2 polysemes, the present study’s main objectives motivate the following research questions:
Research Questions

1. Do *ab initio* learners acquire different levels of both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions under CLBI versus TBI?

2. Do individual differences in *ab initio* learner WMC differentially affect the acquisition of both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions?

3. Do individual differences in *ab initio* learner WMC interact with the instructional treatments CLBI and TBI to differentially affect both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions?

In light of the unique challenge of acquiring productive knowledge of L2 polysemous spatial prepositions, the completely different nature of the two instructional treatments CLBI and TBI as well as the evidence from previous research related to the predictive validity of WMC in SLA, the above research questions motivate the following three related hypotheses:

Related Hypotheses

1. *Ab initio* learners will acquire the same level of both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions under CLBI versus TBI.

2. Individual differences in *ab initio* learner WMC do not differentially affect the acquisition of both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions.
3. Individual differences in ab initio learner WMC do not interact with the instructional treatments CLBI and TBI to differentially affect both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions.

The first research question and hypothesis examine the general effectiveness of two techniques commonly used in the teaching of polysemous spatial prepositions. These two instructional techniques are (1) translation-based instruction (TBI) which treats the multiple meanings of a polyseme as an unrelated list of all the possible translations from the L2 to the L1, and (2) cognitive linguistics-based instruction (CLBI) which treats the multiple meanings of a polyseme as interrelated and connected to a common conceptual base.

The second research question and hypothesis examine the predictive nature of working memory capacity in the acquisition of L2 polysemous spatial prepositions. Working memory has been shown to be one of the most important learner variables in predicting success in L2 acquisition, (e.g. DeKeyser, 2000; Harley and Hart, 1997, 2002; Miyake & Friedman, 1998; Ross, Yoshinaga and Sasaki, 2002; Skehan, 2002), but there has yet to be a study of its specific effects on the acquisition of L2 polysemes.

The third research question and hypothesis examine the interaction of working memory capacity with two different instructional techniques commonly used for the teaching of L2 polysemous spatial prepositions. Therefore, the present study is not simply asking if WMC affects L2 vocabulary acquisition, but rather, it also explores the possible interactions that may occur between the type
of instructional technique and individual differences in the working memory capacity of the L2 learner.
CHAPTER 3: METHODOLOGY

3.1 EXPERIMENTAL DESIGN

The present study followed the classic design for studies involving the interaction of individual differences with differing instructional treatments for the same target structures. Corno, Cronbach, Kupermintz, Lohman, Mandinach, Porteus & Talbert (2002) state that the classic design for aptitude-treatment interaction research studies involves (1) alternative methods for teaching the same content, (2) random assignment of participants to treatments and (3) initial testing to measure abilities hypothesized to be more relevant to one treatment than another. The present study followed this design through the (1) use of CLBI and TBI as alternative methods for teaching polysemes, (2) random assignment of participants into two separate but equal participant groups, (3) and initial testing of working memory which is hypothesized to interact differently with each of the vocabulary teaching methods.

WORKING MEMORY PRE-TESTING

In order to answer the research questions given above, pre-qualified participants were subdivided into three separate groupings based on individual scores in working memory capacity (WMC). A detailed description of the requirements to be prequalified for the study is given in section 3.3. WMC was determined through a complex span task (CST). The common structure of a CST is the combining of a task followed by a recalled element (e.g., a letter,
word, or object). This is repeated for several trials resulting in subsequent tasks interfering with the required memory of the recalled item (Unsworth, Heitz, Schrock, & Engle, 2005).

The two most common tests for measuring working memory capacity are the operation span task (OST) and the reading span task (RST). Oswald, McAbee, Redick and Hambrick (2015) report that in a typical OST, examinees are given a series of very simple math problems to which they simply indicate whether or not the answer given is right or wrong. After the answer, they are shown a letter that they will have to recall later (e.g., $2 + 7 = ?, 9, Q$). After the series is completed, examinees are then prompted with a $4 \times 3$ matrix of letters and asked to click on the letters that were to be recalled in the order in which they were presented. The processing (arithmetic operation), decision (right or wrong), storage (letter), and recall (letter matrix) phases of the automated OST are each presented on separate computer screens to minimize rehearsal of the to-be-recalled items (e.g., Redick et al. 2012, Conway, Kane, Bunting, Hambrick, Wilhelm & Engle 2005; Unsworth et al., 2005).

The reading span task was first developed by Daneman and Carpenter (1980). In a typical RST, examinees are given a series of sentences to read to which they simply indicate true or false on whether the sentences made sense or not. After they answer true or false, they are shown a letter that they will have to recall later. After a series of three to seven sentences followed by a to-be-recalled letter is completed, examinees are then prompted with a matrix of letters and asked to click on the letters that were to be recalled in the order in which
they were presented. The processing (reading), decision (true or false), storage (letter), and recall (letter matrix) phases of the automated RST are each presented on separate computer screens to minimize rehearsal of the to-be recalled items. The RST is argued to more accurately measure WMC because of its combination of both a processing component (i.e., the sentence to be read) and a storage component (i.e., the recalled word), which is critical for assessing IDs in WMC (see also Engle, Tuholski, Laughlin, & Conway, 1999). For the present study, the RST was selected over the OST as the RST utilizes reading as the processing task rather than a math problem, and therefore more closely aligns itself with the linguistics aspects of the present research.

**EXPERIMENTAL GROUPINGS**

A total of 118 participants were pre-tested to determine their WMC using the RST developed by the Working Memory Lab at the Georgia Institute of Technology. After participants completed the RST, those who had WMC scores in the top one-third (61-75) were placed in a high scorers group. Those with WMC scores in the middle one-third (46-60) were placed together in a middle scorers group. And finally, those with WMC scores in the bottom one-third (31-45), were placed together in a low scorers group.

As soon as the high, middle and low WMC groups were finalized, the participants from the middle WMC group were removed from further participation in the study. The rationale for excluding the middle WMC group was to focus the study on two strongly contrasting groups of participants based on their working memory capacity. Focusing on the significant differences in WM capacity
between the two remaining groups provides the opportunity to obtain more
definitive results as to how WMC affects acquisition, as well as how it interacts
with the two vocabulary teaching methods, CLBI and TBI. There were also a
number of participants who scored well below the bottom one-third of scorers
and were therefore removed from the study as they could prove to be outliers
who could skew the results of the study. There were also a number of
participants who qualified to continue in the present study, but chose not to, and
therefore had to be replaced.

**Instructional Phase**

Once 30 high working memory participants and 30 low working memory
participants were established, these two sets of learners were divided into groups
of 15 participants, and then randomly assigned to two different instructional
treatments. These instructional treatments are commonly referred to as (1)
translation-based instruction (TBI) and (2) cognitive linguistics-based instruction
(CLBI). TBI is a traditional approach to vocabulary instruction in which the
various meanings or senses of a polyseme are taught one at a time as a distinct
and unrelated list of possible translated meanings. In contrast, Boers and
Lindstromberg (2008) state that CLBI consists in (1) making learners aware of a
polyseme’s central or prototypical sense and (2) showing how additional senses
of the polyseme are extended from this central sense. In other words, learners
are taught that the polyseme has just one central conceptual meaning and all
other meanings can be shown to link back to it through processes of metaphor
and metonymy. Arranging participants into two separate equal-sized groups
provided the opportunity to teach the target structures to four separate experimental groups. These experimental groups were high working memory CLBI (CB2), high working memory TBI (TB2), low working memory CLBI (CB1) and low working memory TBI (TB1). A more detailed explanation of each instructional treatment is provided in section 3.4.

**POST-TESTING**

Participants were tested twice, first with an immediate post-test followed one week later with a delayed post-test. Mackey and Gass (2016) suggest that delayed post-tests are most commonly given at intervals starting at one-week post treatment, followed possibly by another delayed post-test two weeks later. Since there was only one delayed post-test, a one-week interval was chosen in order to challenge the learner’s recall of the vocabulary but not so great a length of time that no participants would perform just above chance level. Both the immediate and delayed post-tests consisted of a written productive vocabulary task that used both sentences and pictures to elicit a specific response from the learner. The immediate and delayed post-tests were identical except for a change in the pictures used, and occasionally a small change in the specific words used. Having the two post-tests not match was done in order to prevent participants from simply recalling how they answered the immediate post-test. These post-tests were conducted using a Power point slide presentation of each of the picture and sentence combinations. Each of the post-tests ran for 12 minutes, with each slide presentation lasting for 15 seconds. The slides used in both post-tests may be seen in Appendix G and Appendix H.
In regards to pre-testing participant’s knowledge of the target structures, the current study followed what Mackey and Gass (2016) refer to as a post-test only design. In this design, participants are not given a pre-test to measure knowledge or skill prior to treatment. Pre-tests are only needed to ensure comparability of groups prior to treatment. For the present study, a pre-test was unnecessary as all study participants were pre-screened to ensure no prior knowledge of Spanish. As a result, all participant groups entered the study with the same initial level of knowledge. Table 3.2 below, shows the timeline used for the experimental portion of the study.

**Table 3.1 Experimental Timeline**

<table>
<thead>
<tr>
<th>Step 1</th>
<th>Day 1</th>
<th>Step 2</th>
<th>Day 2</th>
<th>Step 3</th>
<th>Day 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consent Form</td>
<td></td>
<td>Background Questionnaire</td>
<td>Instructional Treatment CLBI or TBI</td>
<td>Pre-Test of Working Memory</td>
<td>Immediate productive knowledge post-test</td>
</tr>
</tbody>
</table>

### 3.2 Target Prepositions

The field of L2 vocabulary acquisition has garnered substantial interest from researchers in SLA. This is due in large part to the numerous and wide-ranging possibilities for growth in our understanding of how second languages are learned. There are numerous categories of vocabulary items that can be researched, from the various parts of speech, degrees of abstractness to
concreteness, to whether or not a vocabulary item is categorized as lexical or functional/grammatical.

For the present study, the specific type of vocabulary chosen were all L2 Spanish polysemous prepositions. A unique property of words that are polysemous is that they are posited to profile multiple interrelated meanings or senses. As a result, polysemes allow a language user to express a much greater range of concepts with a single form. Taylor defines polysemy as the, “association of two or more related senses with a single linguistic form” (1995, p. 99). Although a lot of research has been conducted on vocabulary, there is still much that we don’t understand about the development and retention of new lexical items, especially polysemous vocabulary items which have garnered little attention in L2 acquisition studies. However, one thing SLA research has demonstrated is that vocabulary knowledge has arguably the greatest impact on overall L2 proficiency (e.g., Alderson, 2005; Hu & Nation, 2000; Laufer, 1989; Laufer & Sim 1985; Schmitt, Jiang & Grabe, 2011).

Gragg (1984) argues that polysemy reflects the possibility of language change, as well as the coexistence of meanings that range from more general to more specific, more literal to more figurative, to older as well as more innovative and new. In most cases, there is one sense of the word which appears to be dominant, and the other senses appear to be derived from this sense. Schmitt (2010) notes that knowledge of rarer meanings or senses can indicate a more comprehensive knowledge of a lexical item.
Nation (2001) notes that a feature of words that is especially striking is that they can have a lot of different meanings associated with them in the dictionary. He argues that this is particularly true of words that are used more frequently in a language. The relationship between the frequency of a word and the number of meanings or senses it may encode follows Zipf’s law of meaning distribution: words that are more frequent tend to have more meanings (Ferrer-i-Cancho, 2014; Zipf, 1945). Through his pioneering work, Zipf used varied predictor variables to show the qualitative tendency of the number of meanings of a word to increase as its frequency increases. Polysemes are therefore the most common word forms an L2 learner will encounter, making their importance in L2 acquisition even greater.

Empirical evidence demonstrates that polysemes demand a great depth of lexical understanding, and adult L2 learners rarely reach a native-like depth of knowledge of words which have multiple interrelated meanings (e.g., Bensoussan & Laufer, 1984; Pavlenko, 2009). Bensoussan and Laufer (1984) studied whether or not the meaning of some types of words are more easily understood and guessed than others. They attempted to answer this question by examining 60 first year EFL learners’ ability to accurately translate the meaning of five different lexically challenging structures. These five structures were (1) polysemes, (2) false cognates, (3) idioms, (4) synophone / homophone and (5)
morphological troublemakers\textsuperscript{1}. The results of this study show that from among these five different challenging lexical items, learners most frequently mistranslated polysemes. It also shows that learners vary greatly in how well they can recognize and use the multiple meanings of a polysemous word. Kantor (1978) shows that English-speaking learners of Hebrew are able to accurately use one meaning of a Hebrew polyseme, but struggle to accurately use all of its meanings.

**DISTINGUISHING HOMONYMS FROM POLYSEMES**

In order to conduct research into the acquisition of L2 polysemes, the present study had to first identify each of the target vocabulary items as being either a case of polysemy or homonymy. Words that share the same form but have completely unrelated meanings are called homonyms. Homonyms can be either homographs (words with identical written forms) or homophones (words with identical spoken forms). The various meanings of homonyms need to be counted and since their various meanings are unrelated, they must be learned individually. An example Nation (2001) gives of a homonym is the word \textit{bank}, which may refer to the bank of a river or a financial institution or even a small container with a slot in the top for inserting coins.

\textsuperscript{1} The term “morphological troublemakers” is used by Bensoussan and Laufer to refer to lexical items that are composed of multiple morphemes (i.e., unconditionally, underdeveloped, disestablishment, etc.)
By comparison, cases of polysemy involve words with multiple *interrelated* meanings. In other words, a lexical item is understood as polysemous if all its meanings or senses can be shown to be systematically related. One of the most important contributions of *cognitive linguistics* (CL) has been to show that the multiple senses or meanings of a polyseme are related not in an arbitrary way, but rather, in a systematic and natural way by means of several cognitive mechanisms such as image-schemas, metaphor and metonymy.

Several studies within the CL framework give strong support to this hypothesis (e.g., Beréndi, Csábi and Kovecses, 2008; Brugman, 1988; Lakoff, 1990). Working within the CL framework, Touplikioti (2009) cites *catch* as an example of polysemy in English. She argues that there are at least seven distinct, but interrelated concepts that can be expressed by *catch*. They are:

1. **getting hold of** as in *The boy caught the frizbee*.
2. **capturing** as in *The police caught the criminal*.
3. **perceiving** as in *I caught my son sneaking out after curfew*.
4. **becoming affected by** as in *The boy caught a bad cold*.
5. **overtaking** as in *The runner caught the leader at the finish line*.
6. **meeting** as in *I was able to catch my boss before he left his office*.
7. **becoming alight** as in *The forest caught fire*.

According to Touplikioti, all seven of these meanings can potentially be introduced in a single lesson via cognitive linguistic approaches to vocabulary instruction which utilize image-schemas, conceptual metaphor and metonymy. Touplikioti provides a schematic demonstration of how the various meanings or
senses associated with *catch* can all be related back to one core meaning which she argues is ‘get a hold of’. This one core meaning is in turn shown to interrelate to the other six core meanings through a visual schematic. The argument behind cognitive linguistic approaches to understanding polysemes is that through this visual presentation, learners more quickly learn and begin to understand all the various meanings of a polyseme by seeing that they are not an unrelated list of discrete meanings, rather, they are all conceptually linked to a common core meaning.

Numerous researchers have attempted to determine how language learners cope with the challenge of learning words with multiple meanings. Nagy (1997) argues that there are two main ways that language users can handle words with multiple meanings. First, they may have a permanent internal representation of each related meaning and simply select the appropriate sense of the word stored in the brain. He refers to this process as *sense selection*. Second, language users may have an underlying concept that is appropriate for the range of meanings with which the word is used. When language users encounter the word, they have to work out during the comprehension process what particular real-world items the word is referring to. He refers to this process as *referenced specification*.

Ruhl (1989) argues that rather than following dictionaries in seeing words as having multiple meanings, we should assume that each word has a single parent lexical meaning. He argues that there are two major sources of meaning when we comprehend a word in context: (1) its inherent lexical meaning which
refers to its meaning when used in isolation, and (2) inferential meaning which we infer when used alongside other words in the immediate context, and from our knowledge of the world.

The challenge for researchers is that lexical meaning is often very abstract. This is especially true of polysemes where a word has more than one sense, many of which developed over a long period of time through metonymy and metaphorical extensions. Ruhl (1989) argues that the senses are related to each other by general rules that apply to all words. These rules include the idea that words can have a range of senses concrete and abstract and these differences in concreteness and abstractness are inferred from the context. Ruhl produces evidence in support of his position by examining many examples of use and shows that apparent variations in meaning can be accounted for by inferential meaning that is stable and can be seen running through all senses of the word.

**POLYSEMY AND SEMANTIC OVERLAP**

Perhaps the most challenging cases for L2 learners in acquiring polysemes is when they are confronted with learning two different L2 polysemes that appear from their L1 perspective to overlap in their meanings, that is, they appear to be synonyms. This is often due to the fact that there is a single polyseme in the learners L1 that semantically overlaps with two different L2 polysemes. For example, the English preposition *for* overlaps with some of the meanings associated with both Spanish prepositions *por* and *para*, but not all of the meanings. The English preposition for essentially conflates a number of the
meanings associated with both of these Spanish prepositions which leads to a
great deal of confusion for the L2 learner in terms of distinguishing when to use
one rather than the other in a given productive language situation.

Another good example of this are the English words *make* and *do* which
Touplikioti’s (2007) study attempted to teach simultaneously to L1 Greek
speakers. Learners may have a single L1 word that covers much of the same
meanings that both *make* and *do* cover. English speakers say things like, “I did
the dishes and made my bed”, but they never say things like “I did a sandwich
and made my taxes.” The challenge that polysemes with overlapping meaning
present is that although on the surface they appear to share the same meaning
with another word, the underlying core concept of each of the words is actually
different. Touplikioti argues that the core or most salient meaning of *make* is
‘create’ and for *do* it’s ‘perform’. It is from these two central meanings that all
other senses of *make* and *do* branch.

Touplikioti’s research also demonstrates the importance in choosing
words that partially semantically overlap so that study participants are prevented
from correctly responding to study tasks through a process of elimination. In
other words, study participants are forced to make an informed decision between
more than one plausible L2 polyseme. The semantic overlap of words such as
*make* and *do* adds to the difficulty of learning these words, and also to the
difficulty in correctly answering the vocabulary knowledge tests in which they are
simultaneously presented.
For the present study, one of the most important features of polysemes that distinguishes them from other L2 vocabulary is that it appears that learners acquire the multiple meanings of a polyseme differently than they do learning individual meanings of multiple words. In other words, learning seven different words that each encode a single meaning is handled differently than learning a single word that encodes seven interrelated meanings or senses. This fact has been demonstrated in numerous L2 vocabulary acquisition studies. Boers, 2000; Boers & Demecheleer, 1998; Csábi, 2004; Kövecses, 2001; Kövecses and Szabó, 1996; Lam, 2009; Touplikioti, 2007; Verspoor and Lowie, 2003 all show that learners appear to handle the acquisition of several interrelated meanings differently than learning several unrelated meanings. This makes polysemes an especially attractive choice for L2 vocabulary acquisition studies involving individual differences (IDs), especially learner working memory since working memory is what language learners use to briefly store and process new linguistic input in order to analyze it for comprehension. The present study hypothesizes that learners may briefly store and process the meanings of new polysemes differently, not only because of a difference in the vocabulary teaching technique employed during the presentation of new vocabulary, but also as a result of individual differences in working memory capacity (WMC).

**Specific Target Prepositions**

The specific target prepositions selected for the present study were the four polysemous Spanish spatial prepositions 1) *por*, 2) *para*, 3) *en* and 4) *a*.

Table 3.2 below lists each of these spatial prepositions along with a list of English
prepositions that can potentially be used to render each of their meanings into English.

Table 3.2 shows that each of these four Spanish prepositions can be rendered in from between seven and eleven different English prepositions depending on the meaning that needs to be captured. In other words, English requires multiple prepositions to fully capture full lexical network of the spatial and/or temporal meanings captured in each of these Spanish prepositions. It serves to highlight the tremendous challenge that L1 English speakers have in learning these Spanish prepositions as no single English form equivalently matches any of these four prepositions.

**Table 3.2 Specific Target Prepositions**

<table>
<thead>
<tr>
<th>Target Prepositions</th>
<th>Possible English Renderings</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>por</em></td>
<td>by, to, at, for, per, through, as, about, because of, over, along</td>
</tr>
<tr>
<td><em>para</em></td>
<td>by, to, at, for, of, towards, as, for the purpose of, in order to, so as to</td>
</tr>
<tr>
<td><em>en</em></td>
<td>by, to, at, for, in, on, into, about</td>
</tr>
<tr>
<td><em>a</em></td>
<td>by, to, at, of, into, upon, per, next to</td>
</tr>
</tbody>
</table>

**Criteria for Selecting the Target Prepositions**

One of the most important parts of the present study was the selection of the ideal polysemes. There were four main criteria employed during the process of narrowing down the selection of target structures shown in Table 3.2. These criteria were 1) each word must be polysemous, 2) the words should overlap
semantically and grammatically with each other, 3) the words should be pedagogically significant in the acquisition of L2 Spanish, and finally 4) the total number of words should be set to optimally stress the average participant's ability to learn all of them, not too easy nor too hard. The following section contains a detailed explanation of the selection criteria for the set of the target prepositions.

Evidence for the Lexical Polysemy of Prepositions

The first criterion was that each word in the set must be polysemous. Polysemy is normally associated with lexical categories which traditionally includes verbs, nouns and adjectives. This is due to the fact that words that are classified as lexical have salient meanings, and the concept of polysemy is intricately connected to semantic theories of meaning. In contrast, words categorized as functional, such as determiners and conjunctions, tend to lack semantic content, and primarily express grammatical relationships between words in a sentence.

The subject of how to categorize prepositions has been a challenge to theories of linguistics and syntax. Within generative grammar (Chomsky, 1981; Jackendoff, 1973), prepositions are assumed to be one of the four main categories of lexical words along with nouns, verbs and adjectives. Prepositions project a prepositional phrase characterized by the features [-N,-V]. However, lexical words normally have either the feature [+N] or [+V]. There is a growing body of literature (Corver & van Riemsdijk, 2001; Littlefield, 2005; Mardale, 2011; Zwarts; 1997) that argues for a hybrid analysis of prepositions. Specifically, they argue that prepositions belong to a semi-lexical category. They have lexical and
functional properties at once. They also argue for a heterogenous view of prepositions with members that are more lexical and others that are more functional.

Since the present study was focused on the acquisition of word meanings, and not grammatical relationships, the chosen target prepositions had to possess salient meanings. Within the field of cognitive linguistics there have been numerous investigations into the polysemy of spatial prepositions. Beginning with Brugman's (1981) seminal analysis of the English preposition over, there is ample evidence that prepositions are highly polysemous lexical items (e.g., Brugman & Lakoff, 1988). And in general, spatial prepositions have been shown to be highly polysemous lexical items, and therefore, are ideally suited for an experiment involving the acquisition of polysemes (see Boers and Lindstromberg, 2008).

For example, Brugman (1981), Dewell (1994), Evans and Tyler (2004), Lindner (1981) and Tyler and Evans (2004), among others, have found that the peripheral senses or meanings of polysemous spatial prepositions are radially extended from central or prototypical meanings via general cognitive processes, such as image-schema transformations, metonymy and metaphor.

There is strong evidence that the Spanish prepositions a, en, por and para are polysemous (e.g., Delbecque, 1996; Huerta, 2009; Lafford & Ryan, 1995; Lam, 2009).
PREPOSITIONS AS A SEMI-LEXICAL HYBRID CATEGORY

There are debates within syntactic theory concerning how to classify prepositions, as lexical or functional. Within generative grammar, prepositions are categorized as lexical since they assign a theta role to their complements as well as project a prepositional phrase characterized by the binary features [-N,-V] (Jackendoff, 1973; Chomsky, 1981). Additional support for the lexical nature of prepositions is found primarily in spatial prepositions, such as the English prepositions over and through. Spatial prepositions add salient semantic content to a sentence, as demonstrated by their assignment of theta roles. However, a few prepositions, such as the genitive of and the dative to seem to be best categorized as functional in that they assign Case inherently to the structure, but do not add any semantic information or thematic properties.

Another distinction between lexical and functional categories is the ability to add new members. If a category of words allows new members to be added to it, it is considered to an open class. In contrast, categories that do not allow new members are considered to be a closed class. Prepositions are traditionally categorized as a closed class, but the categorization of prepositions as a closed class is tenuous. Their membership in English is taken to range from 50 – 60 as found in traditional grammars of English (Warriner & Griffith, 1977; Pollock, Sheridan, et.al., 1961). However, Fang (2000) found in a corpus study of prepositions that their membership was as high as 248. In addition, it is widely accepted that new prepositions can be added to the class (Kortmann & Konig, 1992; Vincent, 1999). The addition of new prepositions happens at a very slow
rate, but this may be due in large part to the limited number of spatial and
temporal relationships that need to be conveyed.

Although there are contradictions regarding the classification of
prepositions as lexical or functional, even within the category of prepositions
itself; there is a growing body of literature that argues that prepositions belong to
a hybrid classification (Corver & van Riemsdijk, 2001; Littlefield, 2005; Mardale,
2011; Zwarts; 1997). Littlefield (2005) argues that prepositions belong to a semi-
lexical category all on their own. Under this analysis, Littlefield argues that
prepositions can be characterized as having lexical and functional properties at
the same time.

For the present study, it was assumed that the category prepositions is a
very heterogeneous one, meaning that specific prepositions may vary in regards
to the level of lexical and functional characteristics that they possess. For
example, the preposition of may be more functional and the preposition over
more lexical. In other words, among words classified as prepositions, we may
distinguish between a minimum of two subclasses (1) more lexical ones and (2)
more functional ones.

Prepositions are among the most polysemous words in any language, and
are also extremely important to understanding the full meaning of a sentence.
However, for anyone who has studied a foreign language, the polysemy of
prepositions often times appears to be arbitrary and random. This view of
prepositional polysemy is further reinforced by the cross-language mismatches
that occur as the range of meanings associated with a given preposition in one
language almost never overlaps perfectly with the meanings of a preposition in another language. For example, Taylor (1995) notes that in English you put gloves on your hands and a ring on your finger; in Italian gloves go sulle mani, but a ring goes al dito. Italian uses two different prepositions to express these two spatial relationships, whereas English conflates these two semantic uses in a single preposition. Taylor also notes that in German, you go auf Urlaub, you live auf dem Lande, and you meet people auf einer Party, while in English you go on vacation, you live in the country, and you meet people at a party. In this example, the cross-language mismatch is even greater as German uses a single prepositional form to express three spatial meanings that English needs three different forms.

**THE SEMANTICS OF THE SPATIAL RELATIONSHIP PROFILED BY PREPOSITIONS**

Prepositions have proven to be especially intriguing and challenging to both those who work in theories of syntax and grammar as well as for those who work in the foreign language classroom. The challenge and frustration faced by both groups is determining a reasoned account that can explain and link all the various meanings or usages of any given preposition. Taylor argues that foreign language teachers have been relegated to a belief that prepositional usage is idiomatic, and the various meanings just have to be memorized one at a time. In other words, prepositional polysemy is reduced to homonymy. This has led to frustration for both student and teacher, as the acquisition of prepositions progresses extremely slowly. Taylor further argues that this belief is even held by many in mainstream linguistics, to the point that structuralist and generative
linguists have had very little to say about prepositions, and attention has been largely restricted to a small range of central senses. He claims that prepositional polysemy has been largely ignored due to the staggering complexity of prepositional polysemy not appearing to be subject to any obvious rule.

**COGNITIVE LINGUISTICS VIEW OF PREPOSITIONS**

In contrast, researchers working within the framework of cognitive linguistics have made developing a reasoned theoretical understanding of the polysemy of prepositions a priority for several decades. Taylor argues that the demonstration that prepositional polysemy is highly structured has probably been one of the major achievements of the cognitive paradigm.

Among the outstanding early contributions of cognitive linguistics is the dissertation by Brugman (1981). Brugman’s work clearly demonstrates how a cognitive linguistic approach to the understanding of prepositional polysemy can be used to explain the various usages of perhaps the most polysemous of the English prepositions, *over*. In order to provide the reader with a clearer understanding of prepositional polysemy from a cognitive linguistics framework, I have detailed the various meanings and usages of the English preposition *over* using Taylor (1995), Brugman (1981) monograph, Lakoff (1987).

Since all of the target prepositions for the present study are polysemous spatial prepositions, it was important to define the basic terms and concepts used in understanding the polysemy of prepositions. Prepositions, in their semantic usage, are used to spatially locate one entity in reference to another. This spatial locating of one item in reference to another is not limited to the domain of
physical reality, but may extend metaphorically to other domains such as temporal realms and other abstract non-physical realms.

Langacker (1987: 231) introduced two terms often used in literature associated with the study of the polysemy of prepositions. These terms are (1) trajector and (2) landmark. Langacker states that the entity which is located is referred to as the trajector (TR), while the entity which serves as a reference point is referred to as the landmark (LM). For the present study, I have chosen not to use this terminology. This is due to the confusion the term trajector produces in that it seems to imply that the TR is moving, however, the TR may be in a static location.

**THE SPATIAL RELATIONSHIPS PROFILED BY PREPOSITIONS**

Tyler, Mueller, and Ho (2011) use two different terms for the same concepts. These terms are (1) focus element (F) and (2) ground element (G). These terms were adopted in the present study as well as they better capture the semantics of the elements involved in the spatial relationship profiled by prepositions. As with the term trajector, the focus element (F) is the element that is located, and as with the term landmark, the ground element (G) is the reference point. So for example, in the sentence, *The bird flew over the house*, the *bird* is the focus element (F) since it is being located, and the *house* is the ground element (G) since it is the reference point.

Taylor (2003) argues that prepositions may simultaneously refer to more than one aspect of the F-G relationship, but the two most important distinctions are between prepositions which profile a static versus a dynamic F-G
relationship. Taylor states that in a static relationship, the preposition denotes the location or place of an F in reference to the G. Typically in a static relationship, the F is viewed as being not in motion and located at a specific position in relation to the G. Taylor goes on to state that a static F-G relationship may be further nuanced by the 1) size and/or shape of the F and/or G, 2) if there is contact between the F and the G, 3) if the G is perceived to provide support to the F (as with on), 4) containment (within / in), or 5) accompaniment (with). He states that in the case of a dynamic F-G relationship, the preposition may refer to three different types of dynamic relationships, 1) the goal or end-point of the F 's movement, 2) the source or starting-point of the F’s movement or 3) the path which may refer to some or all of the trajectory followed by the F.

With these general characteristics of prepositions in mind, Taylor examines the following sentences with over (2003:113):

(1)  
   a. The lamp hangs over the table.  
   b. The plane flew over the city.  
   c. He walked over the street.  
   d. He walked over the hill.  
   e. He jumped over the wall.  
   f. He turned over the page.  
   g. He turned over the stone.  
   h. He fell over a stone.  
   i. He pushed her over the balcony.  
   j. The water flowed over the rim of the bathtub.
k. He lives over the hill.

It's apparent from these few sample sentences that the preposition over has a great diversity of meanings associated with it. At first glance, it appears that these diverse meanings are unrelated, however, Taylor does an excellent job of explaining how OVER actually constitutes a complex family of related meanings. Taylor states that in the first sentence (The lamp hangs over the table), over denotes a static relationship of place. The F is located vertical to, but not in contact with the G. In The plane flew over the city (b), the F is again vertical to, and not in contact with the G. Taylor notes however that the relationship has changed from static to dynamic since the plane is moving. The expression over the city denotes (part of) the path followed by the F. Taylor argues He walked over the street (c) is similar, except that now there is contact between the F and the G. He walked over the hill (d) is closely related to (c), that is, the F traces a path vertical to, and in contact with, the G. Taylor notes however that a new element has been introduced, namely the shape of the path. In walking over a hill, a person first ascends, reaches the highest point, and then descends. He shows in the example, He jumped over the wall (e) that this curved, arc-like path of the F is again in evidence. He goes on to argue that a new element linked conceptually to the arc-like path makes an appearance, namely, the notion of the G as an obstacle that the F must surmount by first ascending, then descending. Taylor makes the argument that an arc-like path is strongly associated with the preposition over, and can be seen in examples F-K.
The next few examples exploit the idea of a curved path, introduced in (d). In (f), the page moves through 180° as it is turned. (Note that in this and the next few examples, *over* is more of an adverb than a preposition. As suggested earlier, polysemy need not require absolute identity of syntactic function.) In (g), the stone, in being turned over, likewise rotates on its axis. In *He fell over a stone* (h), the subject of the verb traces a more limited arc-like path (say, through 90°), while the unfortunate victim in (I) (He pushed her over the balcony) traces a curved, downward path. In (j), water, in flowing over the rim of a bathtub, traces a path of a similar shape.

Taylor does not restrict his study to the physical realm of the spatial uses of *over*. He argues that there are also a vast number of non-spatial, metaphorical uses. He notes a metaphorical use of *over* is exemplified in the sentence, (2) *He has no authority over me*, which is a metaphorical extension of *The lamp hangs over the table* (1a). However, in (2) the relationship between the F and the G is one of power, not of spatial orientation. Taylor argues that this is due to a transfer of the F-G relationship from the domain of vertical space to the domain of power relations. He further argues that “power relations (like social organization, mentioned earlier) are typically conceptualized in terms of vertical space. Someone with power is 'higher' than someone without power. Hence a preposition denoting a higher vertical location comes to be employed to encode a position of greater power” (2003:115). Taylor notes several other metaphorical extensions of the spatial meanings of *over*, including the following:

(3) He got over his parents' death.
(4) a. Our troubles are over.
   b. The lesson is over.
   c. It isn't over till it's over.

Taylor argues that the family resemblance structure of OVER is a motivated conventionalization of the English Language. Due to this fact, he argues that there is no reason to expect that any one preposition in another language will be structured in a similar way, and indeed, a preposition in one language virtually never has a single translation equivalent in another language. However, Taylor states that “the non-equivalence of prepositions across languages is no reason for accepting the view that prepositional usage is essentially arbitrary” (2003:117). Taylor shows that non-equivalence can be explained very simply in terms of different structurings of the categories, and he demonstrates this in cross-language data from Italian and German (see Taylor 2003: 117-118).

The second criterion was that each word should have at least one other word in the proposed set with which it overlaps both semantically and syntactically. The reason for these two criteria is to force learners to make form meaning connections, rather than form syntactic function connections. For example, if learners are taught a set of four L2 words from four different lexical categories (i.e., one noun, one verb, one preposition, one adjective), they can potentially produce correct responses based simply on recognizing which lexical category best fits the empty syntactic slot. However, if learners are taught a list of new words that could all potentially fill a given syntactic slot, they are forced to
use their knowledge of the meanings of the words to produce accurate responses. In the proposed list, all four words are spatial prepositions, therefore, they all overlap in terms of lexical category as well as syntactic function. Crystal (1985) defines a syntactic function as the grammatical relationship of one constituent to another within a syntactic construction. This prevents participants from producing correct responses based solely on the knowledge of which lexical category best fits the empty syntactic slot.

In addition, all of the target structures semantically overlap in providing salient semantic information about spatial and temporal relationships. Lam (2009) addresses the challenge inherent in learning the L2 Spanish prepositions _por_ and _para_ by L1 English speakers. Speakers of English commonly translate both of these prepositions as the English preposition ‘for’, but may also translate either as _by_, _to_ or _for_. However, Spanish makes a clear distinction between these two prepositions, and Lam argues that only through a clear cognitive linguistic framing can these two Spanish prepositions be made clear.

In addition, these four prepositions are often confused by L1 English speakers learning Spanish. This is due to cross-linguistic mismatch that occurs when L1 English speakers attempt to apply the meanings of the closest natural equivalent English prepositions for these Spanish prepositions. For example, the Spanish prepositions _por_ and _para_ can both be translated using _for_ in English. However, care must be taken when translating _for_ into Spanish, as _por_ and _para_ are used quite differently than the English preposition _for_.

99
The third criterion was the pedagogic significance of the selected words. Even though the total number of prepositions in any given language is far fewer than the total number of verbs, nouns or adjectives, they are still one of the most frequently used words in the corpus of language. Statistically, in a corpus of one million English words, one in ten words is a preposition (Fang, 2000). These specific Spanish words traditionally place a greater learning burden on learners of L2 Spanish. Therefore, any gains in knowledge in how to lessen that burden is of great pedagogical significance.

The fourth and final criterion employed in creating the list of target structures was determining the ideal number of polysemes necessary for optimally pushing the participants’ ability to learn all the meanings taught in the instructional phase of the study. In other words, how many polysemes should be included in the list? If the list is too short and as a result proves to be easy to learn for the majority of participants than scores may not accurately reflect the differences between the identified learners (i.e., all scores tightly grouped at the high end of proposed range). On the other hand, if the proposed list is too long and as a result proves to be unduly difficult to learn for the majority of participants than the scores would again not accurately reflect the differences between the identified learners (i.e., all scores tightly grouped at the low end of the proposed range). In between too easy and too difficult, there is a transition range where the learner has an above-chance rate of accurately responding, but does not always respond correctly. Determining the optimal level of difficulty for the learning task is an important part of the pre-testing phase of the experiment.
There were no additional words included as distractors during the instructional phase of the experiment. Mackey and Gass (2016) argue when assessing discrete language, it is important to embed the lexical items being targeted in a much larger test. They claim that this prevents participants from figuring out the scope of the research, and more importantly, the study instruments are more likely to produce a valid characterization of their L2 knowledge. However, for the present study this was unnecessary. This is due to the fact that all participants had no prior knowledge of Spanish therefore they could not apply previous knowledge in any way that would produce an advantage. In addition, participants were left unaware that other participants in the present study were being taught the same set of four words using a different teaching method.

3.3 Participants

There were two main stages in the present study that each required a different number of total participants. In the first stage, a 118 participants were recruited to complete the test of working memory capacity. This first stage of the experiment took approximately one year to finish due to the challenge in finding pre-qualified participants who had to prior knowledge of Spanish or any cognate language. Once a range of WMC scores were determined for each participant, those who scored in the top 1/3, and the bottom 1/3 continued on to the second stage of the experiment. Those who scored in the middle 1/3 were eliminated from further participation in the study. Also, there were a number of participants who scored very low in the WMC pre-test and were eliminated in order to prevent
the introduction of outsiders who might skew the results of the study. There were also a number of participants who did not complete all the stages of the experimental study, and had to be replaced. The second stage of the experiment involved the instructional treatment and post-testing of the target structures. Having 60 participants in the second stage ensured that there would be at least 15 participants in each of the four main experimental groups, CB2, CB1, TB2 and TB1.

Participants were recruited primarily from the faculty, staff and student population of Columbia International University. In order to qualify to participate, all potential candidates were required to complete a background questionnaire as well as sign a consent form. The required characteristics of qualified participants were that they (1) be 18 years of age or older, (2) are L1 English speakers, (3) have no prior formal or informal Spanish language education or exposure and (4) do not speak or have formal education in a related Romance language (i.e., Portuguese, Italian, French, etc.).

In regards to the age requirement, the minimum age of 18 years was selected because it has been shown that maturational constraints on L2 vocabulary acquisition are present at this age (Long, 2013; Spadaro, 2013). Although research (Chiappe, Siegel, & Hasher, 2000) shows that WM begins to decline after age 29, no upper age limit was established because the present study was only interested in how a learner’s working memory score at a specific time in their adult life affects L2 vocabulary acquisition, as well as how it interacts with differing instructional methods.
In addition to the study participants, eight L1 Spanish speakers were recruited to provide native speaker feedback on the instructional materials as well as the post-testing instruments. These L1 Spanish speakers came from varied Spanish speaking countries, including Spain, Mexico, Costa Rica, Colombia and Ecuador. This ensured that the content of the instructional materials as well as the answers expected in the post-tests accurately reflected the native Spanish speakers’ productive knowledge of the target prepositions.

3.4 MATERIALS

BACKGROUND QUESTIONNAIRE AND CONSENT FORM

Prior to WM testing, all potential study participants were required to sign a consent form (Appendix A) that detailed the nature of the study, ensured participants of complete anonymity and emphasized that taking part in the research was completely voluntary. In addition, all potential study participants were required to complete a background questionnaire (Appendix B) in order to ensure that they met the basic qualifications for participation in the present study (i.e., age, Spanish language knowledge, L1 English speaker, no related Romance language L2s).

COMPLEX SPAN TASK FOR WORKING MEMORY

If a participant was deemed eligible to participate in the study, and had signed the consent form, he or she then moved on to complete the reading span task (RST) to measure his or her working memory capacity. This RST is computer-based, and takes an average of about 25 minutes to complete (Unsworth et al., 2005). The RST that was used in the present study was
developed by researchers with the Georgia Institute of Technology’s Attention and Memory lab. The reading span task is just one example of several possible complex span tasks for measuring working memory capacity. This RST has been thoroughly tested, and has proven to be a valid and reliable source of an individual learner’s WMC. The RST was conducted in a controlled testing room using E-prime software.

**INSTRUCTIONAL TREATMENTS**

Numerous researchers have attempted to determine how language learners cope with the challenge of learning words with multiple meanings. Nagy (1997) argues that there are two main ways that language users can handle words with multiple meanings. First, they may have a permanent internal representation of each related meaning and simply select the appropriate sense of the word stored in the brain. He refers to this process as *sense selection*. Second, language users may have an underlying concept that is appropriate for the range of meanings with which the word is used. When language users encounter the word, they have to work out during the comprehension process what particular real-world items the word is referring to. He refers to this process as *referenced specification*. Table 3.3 below contains a brief summary of the core teaching principles of translation-based and cognitive linguistics based instruction. This summary is based on similar instructional methods employed in Touplikioti (2007).
**Table 3.3 Core Teaching Principles of TBI and CLBI**

<table>
<thead>
<tr>
<th>Translation-based Instruction</th>
<th>Cognitive Linguistics-Based Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruction based on the technique of</td>
<td>Instruction based on the technique of</td>
</tr>
<tr>
<td>memorization and repetition</td>
<td>metaphor awareness</td>
</tr>
<tr>
<td>Focus on language</td>
<td>Focus on cognitive processes involved in relating meanings to a common core</td>
</tr>
<tr>
<td>Learning and automatic reproduction of the formal systems of</td>
<td>Constant interaction between the learner and the materials they are exposed to</td>
</tr>
<tr>
<td>language</td>
<td></td>
</tr>
<tr>
<td>Aim is to have students produce formally correct sentences</td>
<td>Aim is to develop an insight in the learner in order to make his or her own</td>
</tr>
<tr>
<td>Practice drills, translation, repetition and recycling of</td>
<td>Mastery of strategies instead of skill and drill: image schemata, strategy of</td>
</tr>
<tr>
<td>activities</td>
<td>etymological elaboration, conceptual metaphors, body language, mental images,</td>
</tr>
<tr>
<td></td>
<td>collaborative activities</td>
</tr>
<tr>
<td>Very little critical thinking, little focus on speaking and</td>
<td>Critical thinking activities, awareness raising of learners personal learning</td>
</tr>
<tr>
<td>listening skills</td>
<td>style</td>
</tr>
<tr>
<td>Focus is on the meanings-form accuracy of a polysemous word</td>
<td>Focus is on the various senses of a polysemous word rather than its form</td>
</tr>
<tr>
<td>Tends to be teacher centered, passive students instructor is seen as expert</td>
<td>Tends to be student center, active participation of students, instructor is</td>
</tr>
<tr>
<td></td>
<td>seen as facilitator</td>
</tr>
</tbody>
</table>

Participants were provided with an instructional handout that detailed the meanings of each preposition, either under CLBI or TBI (see Appendices D and E). All instructional treatments were taught by the researcher of this present study. With the CLBI treatment, learners were first taught a single conceptual core meaning for each preposition. This conceptual core meaning included a diagram of a proto-scene that visually represented the core spatial relationship that each preposition profiles. This was followed by taking each participant through a series of sentence and picture presentations, and showing how all the other apparently separate meanings in the spatial, temporal and metaphorical
domains can be linked back to a single core meaning. This method employed much of the same strategies from Taylor’s (2003) explanation of the polysemy of the English preposition *over* provided in Section 3.2 of the present study.

Under TBI, learners were introduced to the meanings of each preposition through a list of their common uses, along with the most common English translations. This included taking each participant through a series of sentence and picture presentations, and showing how to accurately translate each spatial situation in order to select the correct preposition. Both instructional treatments used the exact same set of sample sentences and visuals. The variable that separated the treatments was in how learners were instructed to negotiate the meanings of the prepositions. Under CLBI, the learner was shown how to negotiate the meaning by using cognitive linguistics based approaches which center on identifying the focus element (F) and the ground element (G) of each sentence, and then deciding which preposition best encodes the meaning being spatially profiled. In contrast, learners under TBI were instructed to attempt to memorize numerous uses and English translations for each preposition, and from this knowledge they had try to make an informed decision about the correct preposition for each of the post-test questions.

**PRODUCTIVE VOCABULARY TESTS**

There was both an immediate and delayed productive knowledge vocabulary test used in the present study. Both tests employed elicitation tasks using sentences combined with a picture. This combination of a sentence with a picture works to clarify the specific meaning of the preposition that is being
profiled at that moment. Both tests were identical in terms of the specific meanings being tested. However, the visuals used in the delayed post-tests were changed from the immediate post-test in order to avoid any advantage a participant might have if they have photographic memory. In addition to the change in visual, there was also occasionally a need for a slight change in the sentence. For example, in the immediate post-test the sentence may have read, *The leaf is on the end of the branch*, but in the delayed post-test it read, *The orange is on the end of the branch*. Since participants were only being tested for their knowledge of the meanings of the four target prepositions, all other words in the sentences were translated into English.

There was no receptive knowledge testing of the prepositions as this was found to be useless in better understanding the acquisition of L2 prepositional polysemes as pilot-testing showed that participants simply made receptive knowledge judgments based on the visual provided, not on their knowledge of the L2 prepositions. Both post-tests can be found in the Appendices G and H.

### 3.5 Pilot Testing

The Georgia Institute of Technology’s Attention and Memory lab conducted extensive pilot testing on the RST that was used to measure WMC, and has demonstrated that it is a reliable instrument for measuring WMC (Foster, Shipstead, Harrison, Hicks, Redick, Engle, 2014). Pilot testing of the post-tests was conducted with eight L1 Spanish speakers in order to get feedback from their native intuitions about correct responses, and to ensure that each test question had only one possible correct answer from among the four target
prepositions. Following native Spanish speaker testing, the post-tests were pilot tested with five L1 English speakers with some Spanish education in their background in order to see if there was any confusion with test questions, and to set the time-limit on the post-tests.

3.6 DATA CODING AND ANALYSIS

All statistical reports were produced using SPSS version 24. For all independent samples t-test reports and MANOVAs, statistical significance was set at $p \leq .05$. For the interpretation of effect sizes, (i.e., mean differences ($d$), correlation coefficients ($r$)) the present study used an L2 field-specific and empirically-based scale established by Plonsky and Oswald (2014) for the general description and interpretation of effect sizes. Their scale established the values for $d$ at small ($d = 0.40$), medium ($d = 0.70$), and large ($d = 1.00$), and the values for $r$ at small ($r = 0.25$), medium ($r = 0.40$) and large ($r = 0.60$). Each test question was worth one point, with each of the four prepositions tested 12 times, for a total possible score of 48 on both the immediate and delayed post-tests.
CHAPTER 4: RESULTS

This chapter provides a detailed report of the results of the present experimental study. The chapter begins with a general introduction followed by three main sections that cover results directly related to answering each of the three research questions. All statistical reports were produced using SPSS version 24. For all independent samples t-test reports, statistical significance was set at \( p \leq .05 \). For the interpretation of effect sizes, (i.e., mean differences \((d)\), correlation coefficients \((r)\)) the present study used an L2 field-specific and empirically-based scale established by Plonsky and Oswald (2014) for the general description and interpretation of effect sizes. Their scale established the values for \( d \) at small \((d = 0.40)\), medium \((d = 0.70)\), and large \((d = 1.00)\), and the values for \( r \) at small \((r = 0.25)\), medium \((r = 0.40)\) and large \((r = 0.60)\). The four experimental groups are High WMC CLBI (CB2), Low WMC CLBI (CB1), High WMC TBI (TB2) and Low WMC TBI (TB1).

GENERAL INTRODUCTION

The following introduction gives a basic overview of the contents of each section’s experimental results. Section 4.1 corresponds to experimental results related to answering the first research question (i.e., \( \text{Do ab initio learners acquire different levels of both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions under CLBI versus TBI?} \)). Section 4.1.1 reports a comparison of the statistical effects of each instructional treatment
on short-term productive knowledge. It begins with a report of the immediate post-test results in terms of combined mean scores for all four target prepositions from participants under each of the two instructional treatments. This is followed by an independent samples $t$-test report to determine whether there is a statistically significant difference between the immediate post-test mean scores. Along with the mean scores, there is a report of the effect size of the mean differences ($d$) and correlation coefficients ($r$). It also includes independent samples $t$-test reports to determine the statistical significance of the difference in mean scores for each of the separate prepositions. This section concludes with a descriptive report of each instructional treatment group’s immediate post-test prepositional selection pattern including an examination of the individual scores for each of the four separate target prepositions. The immediate post-test results are used to understand the differential effects of instructional treatment on the acquisition of short-term productive knowledge.

Section 4.1.2 reports a comparison of the statistical effects of each instructional treatment on long-term productive knowledge. It begins with a report of the delayed post-test results in terms of combined mean scores for all four target prepositions. This is followed by an independent samples $t$-test report to determine whether there is a statistically significant difference between the delayed post-test mean scores. It also includes independent samples $t$-test reports to determine the statistical significance of the difference in mean scores for each of the separate prepositions. This section concludes with a descriptive
Section 4.2 corresponds to experimental results related to answering the second research question (i.e., Do individual differences in ab initio learner WMC differentially affect the acquisition of both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions?). This first section reports statistical effects of WMC on the productive knowledge of the four target prepositions for each of four experimental groups. Sub-section 4.2.1 reports study findings for the effects of WMC on short-term productive knowledge for CB2 versus CB1, and sub-section 4.2.2 reports findings for TB2 versus TB1. Sub-section 4.2.3 reports study findings for the effects of WMC on long-term productive knowledge for the same two pairs of experimental groups as in sub-section 4.2.3, and subsection 4.2.4 reports study findings for the effects of WMC on long-term productive knowledge for the same two pairs of experimental groups as in sub-section 4.2.2. All four of these sub-sections begin with an examination of each experimental group’s productive knowledge through a comparison of the relevant post-test results in terms of combined mean scores for all four target prepositions. This is followed by an independent samples $t$-test report to determine whether there is a statistically significant difference between the relevant post-test mean scores. Along with the mean scores, there is a calculation of the mean differences ($d$) and correlation coefficients ($r$) to clearly show the effect size difference between the two means. They also include separate independent samples $t$-test reports to determine whether there is a
statistically significant difference between the relevant post-test mean scores of
the individual prepositions. All four of these sub-sections conclude with a
comparison of each experimental group’s specific prepositional selection pattern,
as well as individual mean scores for each of the four separate target
prepositions. The comparison of all of the results in these sub-sections is
between the two experimental groups that each received the same instructional
treatment, but were separated by scores in working memory capacity (i.e., CB2
vs. CB1; TB2 vs. TB1).

Section 4.3 corresponds to experimental results related to answering the
third research question (i.e., *Do individual differences in ab initio learner WMC
interact with the instructional treatments CLBI and TBI to differentially affect both
short-term and long-term productive knowledge of L2 Spanish polysemous
spatial prepositions?*). Section 4.3.1 reports findings from a comparison of the
immediate post-test results for CL1 versus TB2, and Section 4.3.2 reports
findings from a comparison of the immediate post-test results for CB1 versus
TB1. Section 4.3.3 reports findings from a comparison of the delayed post-test
results for the two high WMC experimental groups, and Section 4.3.4 reports
findings from a comparison of the delayed post-test results for the two low WMC
experimental groups. All of these sub-sections begin with a comparison of the
relevant post-test results in terms of combined mean scores for all four target
prepositions. This is followed by an independent samples *t*-test report to
determine whether there is a statistically significant difference between the
relevant post-test mean scores. Along with the mean scores, there is a
calculation of the mean differences (d) and correlation coefficients (r) to clearly show the effect size difference between the two means. Each sub-section concludes with a report comparing individual mean scores for each of the separate prepositions, and independent samples t-test reports to determine the statistical significance of the difference in mean scores for each. The comparison of all the results in section 4.3 is between the two experimental groups that each scored in the same range on the pre-test of WMC, but received different instructional treatments (i.e., TB2 vs. TB1; CB2 vs. CB1). That is, the results relate to the interaction between WMC and instructional treatment, and the resultant effects on both short-term and long-term productive knowledge.

4.1 Effects of Instructional Treatment on Short and Long-term Productive Knowledge

The presentation of the results of the effects of instructional treatment on both short and long-term productive knowledge is detailed below. This section compares post-test results from participants under CLBI versus participants under TBI for both short and long-term productive knowledge.

4.1.1 Effects of Instructional Treatment on Short-term Productive Knowledge

Table 4.1 below reports the immediate post-test results in terms of combined mean scores for all four target prepositions from participants under CLBI and TBI. These mean scores were determined using the individual scores from each participant on the immediate post-test. There were a total of 48
questions on the immediate post-test (Appendix F), each worth one point, for the highest possible mean score of 48.

**TABLE 4.1 CLBI VS. TBI IMMEDIATE POST-TEST RESULTS**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>d</th>
<th>r</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLBI</td>
<td>30.50</td>
<td>5.90</td>
<td>30</td>
<td>0.07</td>
<td>0.04</td>
<td>0.846</td>
</tr>
<tr>
<td>TBI</td>
<td>30.07</td>
<td>6.09</td>
<td>30</td>
<td>0.07</td>
<td>0.04</td>
<td>0.846</td>
</tr>
</tbody>
</table>

The results in Table 4.1 show that the effect size of the mean difference is small ($d = 0.07$) and of the correlation coefficient is small ($r = 0.04$). An independent samples $t$-test was conducted to compare mean differences between all participants under each of instructional treatments, the immediate post-test results did not yield a significant difference in scores for participants under CLBI ($M = 30.05$, $SD = 5.90$) and participants under TBI ($M = 30.07$, $SD = 6.09$) conditions; ($t(58) = 0.1964$, $p = 0.846$).

In addition to the combined mean scores of all four target prepositions shown in Table 4.1, each instructional group’s results were calculated to determine individual mean scores for each preposition separately, along with standard deviations and a $p$-value for each (Table 4.2). These results show that there are differential results for each of the four target prepositions.

Table 4.3 below breaks down the individual prepositional selection pattern for both instructional treatment groups. These descriptive results report both the raw number of times the *correct* prepositions were chosen, as well as the raw number of times each of the *incorrect* prepositions were chosen. By showing what prepositions learners are selecting in place of the correct preposition, these
data provide insights into the possible causes of each experimental group’s error patterns. The table is organized with one of the four target prepositions heading each of the rows as well as each of the columns for both of the instructional group’s results. The preposition that heads each row represents the correct answer. The preposition heading each column represents which preposition was actually chosen. The highest score that any instructional treatment group could receive on any individual preposition was 360. Each test question was worth one point, and the coding protocol involved multiplying the total number of participants who received each instructional treatment (30), by the total number of test questions for each preposition (12), to get to the 360 total. The results in each block show the total number of times each preposition was selected, along with percentage out of the possible 360 that this number represents. The correct number of answers along with the corresponding percentage correct is shaded in gray for each preposition.

**Table 4.2 CLBI vs. TBI Immediate Post-test Preposition Results**

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Treatment</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>para</td>
<td>CLBI</td>
<td>7.07</td>
<td>2.92</td>
<td>30</td>
<td>0.962</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>7.03</td>
<td>2.46</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>en</td>
<td>CLBI</td>
<td>8.77</td>
<td>1.70</td>
<td>30</td>
<td>0.713</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>8.97</td>
<td>2.43</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>por</td>
<td>CLBI</td>
<td>7.03</td>
<td>1.97</td>
<td>30</td>
<td>0.134</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>6.17</td>
<td>2.42</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>CLBI</td>
<td>7.63</td>
<td>2.08</td>
<td>30</td>
<td>0.653</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>7.90</td>
<td>2.53</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>
The results (Table 4.3) show that the greatest percentage difference in immediate post-test scores between the CLBI and TBI experimental groups occurred within the preposition *por* with CLBI at 59% correct versus TBI at 51% correct. The preposition *por* is also the only preposition in which CLBI outscored TBI. With the prepositions *en* and *a*, the TBI experimental group outscored the CLBI experimental group by 2% in both. Finally, with the preposition *para*, both groups scored 59% correct out of 360, with CLBI getting just one more problem correct (212 vs. 211).

In terms of the error patterns, both CLBI and TBI instructional groups have similar selection patterns between the four prepositions. However, the results do show a modestly higher tendency from the CLBI experimental group to incorrectly select the preposition *para* when the correct answer should have been *a* (19% incorrect vs 10% incorrect). The results also show a modestly higher tendency from the TBI experimental group to incorrectly select the preposition *en* when the correct should have been *a* (13% incorrect vs 4% incorrect).

**4.1.2 Effects of Instructional Treatment on Long-term Productive Knowledge**

Table 4.4 below reports the delayed post-test results in terms of combined mean scores for all four target prepositions from participants under CLBI and TBI. These mean scores were determined using the individual scores from each participant on the delayed post-test. This test followed the same coding protocol given in Section 4.1.1.
### Table 4.3 CLBI vs. TBI Immediate Post-test Selection Patterns

<table>
<thead>
<tr>
<th></th>
<th>para</th>
<th>en</th>
<th>por</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CLBI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate Post-test Selection Patterns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>para</td>
<td>212 / 59%</td>
<td>17 / 5%</td>
<td>104 / 29%</td>
<td>27 / 7%</td>
</tr>
<tr>
<td>en</td>
<td>25 / 7%</td>
<td>263 / 73%</td>
<td>18 / 5%</td>
<td>54 / 15%</td>
</tr>
<tr>
<td>por</td>
<td>75 / 21%</td>
<td>50 / 14%</td>
<td>211 / 59%</td>
<td>24 / 6%</td>
</tr>
<tr>
<td>a</td>
<td>68 / 19%</td>
<td>14 / 4%</td>
<td>49 / 13%</td>
<td>229 / 64%</td>
</tr>
<tr>
<td><strong>TBI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate Post-test Selection Patterns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>para</td>
<td>211 / 59%</td>
<td>7 / 2%</td>
<td>121 / 34%</td>
<td>21 / 5%</td>
</tr>
<tr>
<td>en</td>
<td>9 / 2%</td>
<td>269 / 75%</td>
<td>28 / 8%</td>
<td>54 / 15%</td>
</tr>
<tr>
<td>por</td>
<td>86 / 24%</td>
<td>56 / 16%</td>
<td>185 / 51%</td>
<td>33 / 9%</td>
</tr>
<tr>
<td>a</td>
<td>35 / 10%</td>
<td>49 / 13%</td>
<td>39 / 11%</td>
<td>237 / 66%</td>
</tr>
</tbody>
</table>

### Table 4.4 CLBI vs. TBI Delayed Post-test Results

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>d</th>
<th>r</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLBI</td>
<td>29.23</td>
<td>6.73</td>
<td>30</td>
<td>0.48</td>
<td>0.23</td>
<td>0.069</td>
</tr>
<tr>
<td>TBI</td>
<td>26.17</td>
<td>6.02</td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The results in Table 4.4 show that the effect size of the mean difference is small ($d = 0.48$) and of the correlation coefficient is small ($r = 0.23$). This is an increase in effect size in comparison to the immediate post-test (Table 4.1). An independent samples $t$-test showed that the delayed post-test results did not yield a significant difference in scores for participants under CLBI ($M = 29.23$, $SD = 6.73$) and TBI ($M = 26.17$, $SD = 6.02$) conditions; ($t(58) = 1.8562$, $p = 0.069$).
The effect size of the mean difference (Table 4.4) is small ($d = 0.48$) and of the correlation coefficient is small ($r = 0.23$).

In addition to the combined mean scores of all four target prepositions shown in Table 4.4, each instructional treatment group’s results were calculated to determine individual mean scores for each preposition separately, along with standard deviations and $p$-values for each. The combined delayed post-test results for CLBI and TBI (Table 4.4) did not yield a statistically significant difference in mean scores ($p = 0.069$). However, an independent samples $t$-test yielded a statistically significant difference in scores for the preposition *por* for participants under CLBI *por* ($M = 6.33$, $SD = 1.97$) and TBI *por* ($M = 4.70$, $SD = 1.96$) conditions; ($t(58) = 3.2127$, $p = 0.002$). The effect size of the mean difference of the preposition *por* (Table 4.5) is medium ($d = 0.83$) and of the correlation coefficient is medium ($r = 0.38$).

**Table 4.5 CLBI vs. TBI Delayed Post-test Preposition Results**

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Treatment</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>$t$-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>para</em></td>
<td>CLBI</td>
<td>6.53</td>
<td>3.39</td>
<td>30</td>
<td>0.218</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>5.60</td>
<td>2.29</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td><em>en</em></td>
<td>CLBI</td>
<td>9.00</td>
<td>1.36</td>
<td>30</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>9.00</td>
<td>2.73</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td><em>por</em></td>
<td>CLBI</td>
<td>6.33</td>
<td>1.97</td>
<td>30</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>4.70</td>
<td>1.96</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td><em>a</em></td>
<td>CLBI</td>
<td>7.37</td>
<td>2.81</td>
<td>30</td>
<td>0.472</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>6.87</td>
<td>2.53</td>
<td>30</td>
<td></td>
</tr>
</tbody>
</table>
Table 4.6 below breaks down the individual prepositional selection pattern for both instructional treatment groups from the delayed post-test. These descriptive results report both the raw number of times the correct prepositions were chosen, as well as the raw number of times each of the incorrect prepositions were chosen. By showing what prepositions learners are selecting in place of the correct preposition, these data provide insights into the possible causes of each experimental group’s error patterns. The table is organized with one of the four target prepositions heading each of the rows as well as each of the columns. The preposition that heads each row represents the correct answer. The preposition heading each column represents which preposition was actually chosen. The highest score that any instructional treatment group could receive on any individual preposition was 360. This test followed the same coding protocol given in Section 4.1.1. The results in each block show the total number of times each prepositions was selected, along with percentage out of the possible 360 that this number represents. The correct number of answers along with the corresponding percentage correct is shaded in gray for each preposition.

Like the immediate post-test results, the results show that the greatest percentage difference in delayed post-test scores between the CLBI and TBI instructional groups occurred within the preposition *por* with CLBI scoring with 53% accuracy versus TBI at 39% accuracy. With the preposition *para*, the CLBI experimental group outscored the TBI experimental group by 7% points. For the preposition *a*, the CLBI experimental group outscored the TBI experimental
group by 5% points. Finally, with the preposition *en*, both groups scored exactly the same with 270 correct answers or 75% accuracy.

**TABLE 4.6 CLBI VS. TBI DELAYED POST-TEST SELECTION PATTERNS**

| CLBI Delayed Post-test Selection Patterns |  |  |  
|------------------------------------------|---|---|---|
| *para*                                   | *en* | *por* | *a* |
| 196 / 54%                                | 16 / 4% | 121 / 34% | 27 / 8% |
| *en*                                     | 16 / 4% | 270 / 75% | 17 / 5% | 57 / 16% |
| *por*                                    | 70 / 19% | 66 / 18% | 190 / 53% | 34 / 9% |
| *a*                                      | 70 / 19% | 18 / 5% | 51 / 14% | 221 / 62% |

| TBI Delayed Post-test Selection Patterns |  |  |  
|------------------------------------------|---|---|---|
| *para*                                   | *en* | *por* | *a* |
| 168 / 47%                                | 12 / 3% | 142 / 39% | 38 / 11% |
| *en*                                     | 21 / 6% | 270 / 75% | 18 / 5% | 51 / 14% |
| *por*                                    | 94 / 26% | 81 / 23% | 141 / 39% | 44 / 12% |
| *a*                                      | 50 / 14% | 43 / 12% | 61 / 17% | 206 / 57% |

In terms of the error patterns, both CLBI and TBI instructional groups have similar selection patterns between the four prepositions. However, the results do show a modestly higher tendency from the CLBI experimental group to incorrectly select the preposition *para* when the correct should have been *a* (19% incorrect vs 14% incorrect). The results also show a modestly higher tendency from the TBI experimental group to incorrectly select the preposition *en* when the correct answer should have been *a* (12% incorrect vs 5% incorrect). Overall, the results show the greatest difference in scores occurred with the preposition *por*, as the TBI experimental group incorrectly selected *para* for *por* 26% of the time, and *en* for *por* 23% of the time.
4.2 Effects of WMC on Short and Long-term Productive Knowledge

The comparison of the results of the effects of WMC on both short and long-term productive knowledge is separated into pairs of experimental groups that each received the same instructional treatment (i.e., CB2 vs. CB1 and TB2 vs. TB1). This is done to isolate WMC as the main effect in focus under both pairs of experimental groups. General results show that for participants under both CLBI and TBI, higher scores in WMC positively correlate to higher scores in both the immediate and delayed post-tests. Results for participants under CLBI are presented first, followed by participants under TBI. There was a 6.83 point differential in the scores on the immediate post-tests between the two high WMC experimental groups ($M = 33.70$, $SD = 4.22$) and the two low WMC experimental groups ($M = 26.87$, $SD = 5.08$). There was a 6.27 point differential in the scores on the delayed post-tests between the high WMC experimental group ($M = 30.84$, $SD = 5.21$), and the low WMC experimental group ($M = 24.57$, $SD = 5.88$).

4.2.1 Effects of WMC on Short-term Productive Knowledge under CLBI

A total of 30 participants were taught the four target prepositions under CLBI. Prior to instruction, these 30 participants were separated into two equal sized experimental groups (CB2 and CB1) of 15 participants each based on individual scores in WMC. Table 4.7 below shows the mean score and standard deviation for each of the CLBI experimental groups on the immediate post-test. These mean scores were determined using the individual scores from each participant on the immediate post-test. This test followed the same coding protocol given in Section 4.1.1. The immediate post-test yields a significant
difference in scores (Table 4.7) for CB2 \((M = 32.73, \ SD = 4.81)\) and CB1 \((M = 28.27, \ SD = 6.03)\) conditions; \(t(58) = 3.1670, (p = 0.003)\). This works out to a difference in raw mean scores of 4.46 points or a 15.8% higher score for the CB2 experimental group versus the CB1 experimental group. The effect size of the mean difference (Table 4.7) is medium \((d = 0.82)\) and of the correlation coefficient is medium \((r = 0.38)\).

**TABLE 4.7 CB2 VS. CB1 IMMEDIATE POST-TEST RESULTS**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>d</th>
<th>r</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLBI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>32.73</td>
<td>4.81</td>
<td>15</td>
<td></td>
<td>0.82</td>
<td>0.38</td>
<td>0.003</td>
</tr>
<tr>
<td>low</td>
<td>28.27</td>
<td>6.03</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In addition to the combined mean scores (Table 4.7), each experimental group’s results were calculated to determine individual mean scores (Table 4.8) for each preposition separately, along with standard deviations and \(p\)-values for each. When running independent samples \(t\)-tests on the individual mean score differences for each preposition separately, the immediate post-test results yielded a significant difference in scores for CB2 preposition *por* \((M = 7.73, \ SD = 2.22)\) and TB1 preposition *por* \((M = 6.33, \ SD = 1.45)\) conditions; \(t(28) = 2.0466, (p = 0.050)\). The effect size of the mean difference for the preposition *por* (Table 4.9) is medium \((d = 0.75)\) and of the correlation coefficient is medium \((r = 0.35)\).

Table 4.9 below breaks down the individual prepositional selection pattern for both experimental groups under CLBI. These descriptive results report both the raw number of times the *correct* prepositions were chosen, as well as the raw number of times each of the *incorrect* prepositions were chosen. By showing
what prepositions learners are selecting in place of the correct preposition, these data provide insights into the possible causes of each experimental group’s error patterns. The table is organized with one of the four target prepositions heading each of the rows as well as each of the columns. The preposition that heads each row represents the correct answer. The preposition heading each column represents which preposition was actually chosen. The highest score that any experimental group could receive on any individual preposition was 180. Each test question was worth one point, and the coding protocol involved multiplying the total number of participants in each experimental group (15) by the total number of test questions for each preposition (12) to get to the 180 total. The results shown in each block represent the total number of times each preposition was selected by that group, as well as the percentage out of a 180 that this number represents. The results shaded in gray blocks represent the total number of times each preposition was correctly selected by that experimental group, as well as the percentage out of a 180 that this number represents.

**Table 4.8 CB2 vs. CB1 Immediate Post-test Preposition Results**

<table>
<thead>
<tr>
<th>Preposition</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>para</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>7.93</td>
<td>2.46</td>
<td>15</td>
<td></td>
<td>0.105</td>
</tr>
<tr>
<td>low</td>
<td>6.20</td>
<td>3.17</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>en</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>9.13</td>
<td>1.36</td>
<td>15</td>
<td></td>
<td>0.242</td>
</tr>
<tr>
<td>low</td>
<td>8.40</td>
<td>1.96</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>por</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>7.73</td>
<td>2.22</td>
<td>15</td>
<td></td>
<td><strong>0.050</strong></td>
</tr>
<tr>
<td>low</td>
<td>6.33</td>
<td>1.45</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>a</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>7.93</td>
<td>1.53</td>
<td>15</td>
<td></td>
<td>0.300</td>
</tr>
<tr>
<td>low</td>
<td>7.33</td>
<td>2.50</td>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The results (Table 4.9) show that the greatest percentage difference in immediate post-test scores between CB2 and CB1 experimental groups occurred within the preposition *para* (66% vs. 52%), followed by *por* (64% vs. 53%), *en* (76% vs. 70%) and finally *a* (66% vs. 61%). In terms of the error patterns, both the CB2 and CB1 experimental groups have similar selection patterns between the four prepositions. However, the results do show a higher tendency from the CB1 experimental group to select the preposition *por* when the correct answer should have been the preposition *para*. The CB1 experimental group incorrectly selected *por* for *para* 35% of the time versus the CB2 experimental group selecting *por* for *para* 23% of the time.
**4.2.2 Effects of WMC on Short-term Productive Knowledge under TBI**

A total of 30 participants were taught the four target prepositions using TBI. Prior to instruction, these 30 participants were divided into two equal sized experimental groups (TB2 and TB1) of 15 participants each based on their individual scores in WMC. Table 4.10 below shows the mean score and standard deviation for each of the TBI experimental groups on the immediate post-test. These mean scores were determined by combining the individual scores from each participant on the immediate post-test. This test followed the same coding protocol given in Section 4.1.1. The results show that participants in the TB2 experimental group earned a mean score of 34.67 versus a score of 25.27 for participants in the TB1 experimental group. This works out to a difference in mean scores of 9.40 points or a 37.2% higher score for the TB2 group versus the TB1 group.

**Table 4.10 TB2 vs. TB1 Immediate Post-test Results**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>d</th>
<th>r</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBI</td>
<td>high</td>
<td>34.67</td>
<td>3.29</td>
<td>15</td>
<td>2.66</td>
<td>0.88</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>25.47</td>
<td>3.63</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An independent samples t-test was conducted to compare mean differences between TB2 and TB1 (Table 4.10), and the immediate post-test results yielded a significant difference in scores for TB2 ($M = 34.67$, $SD = 3.29$) and TB1 ($M = 25.47$, $SD = 3.63$) conditions; ($t(28) = 7.2731$, $p < 0.001$). The effect size of the mean difference is large ($d = 2.66$) and of the correlation coefficient is large ($r = 0.88$).
In addition to the combined mean scores (Table 4.10), each experimental
group’s results were calculated to determine individual mean scores for each
preposition separately, along with standard deviations and p-values for each
(Table 4.11).

**TABLE 4.11 TB2 vs. TB1 IMMEDIATE POST-TEST PREPOSITION RESULTS**

<table>
<thead>
<tr>
<th>Preposition</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>para</td>
<td>high</td>
<td>8.27</td>
<td>2.12</td>
<td>15</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>5.80</td>
<td>2.18</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>en</td>
<td>high</td>
<td>10.27</td>
<td>1.36</td>
<td>15</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>7.67</td>
<td>1.96</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>por</td>
<td>high</td>
<td>7.53</td>
<td>2.36</td>
<td>15</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>4.80</td>
<td>1.61</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>high</td>
<td>8.60</td>
<td>2.53</td>
<td>15</td>
<td>0.114</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>7.20</td>
<td>2.40</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

When running separate independent samples t-test on the individual mean
score differences for each preposition under TBI, three out of the four
prepositions met the criteria for statistical significance. For the data in Table
4.11, the difference in the immediate post-test mean scores for the preposition
*para* yielded a significant difference in scores for TB2 *para* (*M* = 8.27, *SD* = 2.12)
and TB1 *para* (*M* = 5.80, *SD* = 2.18) conditions; (*t*(28) = -3.143, *p* = 0.004). The
preposition *en* yielded a significant difference in scores for TB2 *en* (*M* = 10.27,
*SD* = 1.36) and TB1 *en* (*M* = 7.67, *SD* = 1.96) conditions; (*t*(28) = 4.2210, *p* <
0.001), and the preposition *por* yielded a significant difference in scores for TB2
*por* (*M* = 7.53, *SD* = 2.36) and TB1 *por* (*M* = 4.80, *SD* = 1.61) conditions; (*t*(28) =
-3.708, $p < 0.001$). The effect size of the mean difference for the preposition *para* is large ($d = 1.15$) and of the correlation coefficient is medium-large ($r = 0.50$). The effect size of the mean difference for the preposition *en* is large ($d = 1.54$) and of the correlation coefficient is large ($r = 0.61$). The effect size of the mean difference for the preposition *por* is large ($d = 1.35$) and of the correlation coefficient is large ($r = 0.56$).

Table 4.12 below breaks down the individual prepositional selection pattern for both experimental groups under TBI. These descriptive results report both the raw number of times the *correct* prepositions were chosen, as well as the raw number of times each of the *incorrect* prepositions were chosen. By showing what prepositions learners are selecting in place of the correct preposition, these data provide insights into the possible causes of each experimental group’s error patterns. The table is organized with one of the four target prepositions heading each of the rows as well as each of the columns. The preposition that heads each row represents the correct answer. The preposition heading each column represents which preposition was actually chosen. The highest score that any experimental group could receive on any individual preposition was 180. This test followed the same coding protocol given in Section 4.2.1. The results shown in each block represent the total number of times each preposition was selected by that group, as well as the percentage out of a 180 that this number represents.

The correct number of answers along with the corresponding percentage is shaded in gray for each preposition. The results show that the greatest
percentage difference in immediate post-test scores between the high and TB1 experimental groups occurred within the preposition *por* (63% vs. 40%) followed by *en* (86% vs. 64%), *para* (69% vs. 48%) and finally *a* (72% vs. 60%).

In terms of the error patterns, both the TB2 and TB1 experimental groups have similar selection patterns between the four prepositions. However, the results do show a much higher tendency from the TB1 experimental group to select the preposition *por* when the correct answer should have been *para*. The TB1 experimental group incorrectly selected *por* for *para* 41% of the time versus the TB2 experimental group selecting *por* for *para* 26% of the time. The TB1 experimental group also had a much higher tendency to incorrectly select *a* for *en* (21% vs 9%) as well as to select *a* for *por* (14% vs. 4%).

**TABLE 4.12 TB2 vs. TB1 Immediate Post-test Selection Patterns**

<table>
<thead>
<tr>
<th></th>
<th>TB2 Immediate Post-test Selection Patterns</th>
<th></th>
<th>TB1 Immediate Post-test Selection Patterns</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>para</td>
<td>en</td>
<td>por</td>
<td>a</td>
</tr>
<tr>
<td><em>para</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>en</em></td>
<td>3 / 2%</td>
<td></td>
<td><strong>154 / 86%</strong></td>
<td>6 / 3%</td>
</tr>
<tr>
<td><em>por</em></td>
<td>35 / 20%</td>
<td>24 / 13%</td>
<td><strong>113 / 63%</strong></td>
<td>8 / 4%</td>
</tr>
<tr>
<td><em>a</em></td>
<td>14 / 8%</td>
<td>21 / 11%</td>
<td>16 / 9%</td>
<td><strong>129 / 72%</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>en</em></td>
<td>6 / 3%</td>
<td></td>
<td><strong>115 / 64%</strong></td>
<td>22 / 12%</td>
</tr>
<tr>
<td><em>por</em></td>
<td>51 / 28%</td>
<td>32 / 18%</td>
<td><strong>72 / 40.0%</strong></td>
<td>25 / 14%</td>
</tr>
<tr>
<td><em>a</em></td>
<td>21 / 12%</td>
<td>28 / 15%</td>
<td>23 / 13%</td>
<td><strong>108 / 60.0%</strong></td>
</tr>
</tbody>
</table>
4.2.3 Effects of WMC on Long-term Productive Knowledge under CLBI

A total of 30 participants were taught the four target prepositions under CLBI. Prior to instruction, these 30 participants were divided into two equal sized experimental groups (CB2 and CB1) of 15 participants each based on their individual scores in WMC. Table 4.13 below shows the mean score and standard deviation for each of the CLBI experimental groups on the delayed post-test. These mean scores were determined by combining the individual scores from each participant on the delayed post-test. This test followed the same coding protocol given in Section 4.1.1. The results show that participants in the CB2 experimental group earned a mean score of 31.87 versus a score of 26.60 for participants in the CB1 experimental group. This works out to a difference in mean scores of 5.27 points or a 19.8% higher score for CB2 versus CB1.

An independent samples t-test was conducted to determine the statistical significance of the difference in mean scores shown in Table 4.13. The delayed post-test results yielded a significant difference in scores for CB2 (\( M = 31.87, \ SD = 5.67 \)) and CB1 (\( M = 26.60, \ SD = 6.70 \)) conditions; (\( t(28) = 2.3254, \ p = 0.028 \)). The effect size of the mean difference is medium (\( d = 0.85 \)) and of the correlation coefficient is medium (\( r = 0.39 \)).

In addition to the combined mean scores (Table 4.13), each experimental group’s results were calculated to determine individual mean scores for each preposition separately, along with standard deviations and \( p \)-values for each (Table 4.14)
TABLE 4.13 CB2 VS. CB1 DELAYED POST-TEST RESULTS

<table>
<thead>
<tr>
<th>Treatment</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>d</th>
<th>r</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLBI</td>
<td>high</td>
<td>31.87</td>
<td>5.67</td>
<td>15</td>
<td>0.85</td>
<td>0.39</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>26.60</td>
<td>6.70</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When running an independent samples \( t \)-test on the individual mean score differences (Table 4.14) for each preposition separately, the preposition *para* yielded a statistically significant difference in scores for CB2 *para* (\( M = 7.73, SD = 2.71 \)) and CB1 *para* (\( M = 5.33, SD = 3.66 \)) conditions; (\( t(28) = -2.041, p = 0.051 \)). The effect size of the mean difference for the preposition *para* (Table 4.15) is medium (\( d = 0.75 \)) and of the correlation coefficient is medium (\( r = 0.35 \)).

TABLE 4.14 CB2 VS. CB1 DELAYED POST-TEST PREPOSITION RESULTS

<table>
<thead>
<tr>
<th>Preposition</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>para</em></td>
<td>high</td>
<td>7.73</td>
<td>2.71</td>
<td>15</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>5.33</td>
<td>3.66</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td><em>en</em></td>
<td>high</td>
<td>9.27</td>
<td>1.10</td>
<td>15</td>
<td>0.292</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>8.73</td>
<td>1.58</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td><em>por</em></td>
<td>high</td>
<td>6.73</td>
<td>2.31</td>
<td>15</td>
<td>0.274</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>5.93</td>
<td>1.53</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td><em>a</em></td>
<td>high</td>
<td>8.13</td>
<td>2.62</td>
<td>15</td>
<td>0.141</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>6.60</td>
<td>2.90</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.15 below breaks down the individual prepositional selection pattern for both experimental groups under CLBI. These descriptive results report both the raw number of times the *correct* prepositions were chosen, as
well as the raw number of times each of the *incorrect* prepositions were chosen. By showing what prepositions learners are selecting in place of the correct preposition, these data provide insights into the possible causes of each experimental group’s error patterns. The table is organized with one of the four target prepositions heading each of the rows as well as each of the columns. The preposition that heads each row represents the correct answer. The preposition heading each column represents which preposition was actually chosen. The highest score that any experimental group could receive on any individual preposition was 180. This test followed the same coding protocol given in Section 4.2.1. The results shown in each block represent the total number of times each preposition was selected by that group, as well as the percentage out of a 180 that this number represents.

The correct number of answers along with the corresponding percentage is shaded in gray for each preposition. The results show that the greatest percentage difference in delayed post-test scores between the CB2 and CB1 experimental groups occurred within the preposition *para* (64% vs. 45%), followed by *a* (68% vs. 55%), *por* (56% vs. 49%) and finally *en* (77% vs. 73%).

In terms of the error patterns, both the CB2 and CB1 experimental groups have similar selection patterns between the four prepositions. However, the results do show a higher tendency from the CB1 experimental group to select the preposition *por* when the correct answer should have been *para*. The CB1 experimental group incorrectly selected *por* for *para* 42% of the time versus the CB2 experimental group selecting *por* for *para* just 26% of the time. The CB1
experimental group also selected *en* for *por* at a moderately higher rate (22% vs. 14%).

### Table 4.15 CB2 vs. CB1 Delayed Post-test Selection Patterns

<table>
<thead>
<tr>
<th></th>
<th>CB2 Delayed Post-test Selection Patterns</th>
<th>CB1 Delayed Post-test Selection Patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>para</strong></td>
<td>116 / 64%</td>
<td>80 / 45%</td>
</tr>
<tr>
<td><strong>en</strong></td>
<td>6 / 3%</td>
<td>10 / 5%</td>
</tr>
<tr>
<td><strong>por</strong></td>
<td>46 / 26%</td>
<td>75 / 42%</td>
</tr>
<tr>
<td><strong>a</strong></td>
<td>12 / 7%</td>
<td>15 / 8%</td>
</tr>
</tbody>
</table>

### 4.2.4 Effects of WMC on Long-term Productive Knowledge under TBI

A total of 30 participants were taught the four target prepositions using TBI. Prior to instruction, these 30 participants were divided into two equal sized experimental groups (TB2 and TB1) of 15 participants each based on their individual scores in WMC. Table 4.16 below shows the mean score and standard deviation for each of the TBI experimental groups on the delayed post-test. These mean scores were determined by combining the individual scores from each participant on the delayed post-test. This test followed the same coding protocol given in Section 4.1.1. The results show that participants in the TB2 experimental group earned a mean score of 29.80 versus a score of 22.53
for participants in the TB1 experimental group. This works out to a difference in mean scores of 7.27 or a 32.3% higher score for the high WMC group versus the low WMC group.

**Table 4.16 TB2 vs. TB1 Delayed Post-test Results**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>d</th>
<th>r</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TBI</td>
<td>high</td>
<td>29.80</td>
<td>4.77</td>
<td>15</td>
<td>1.59</td>
<td>0.62</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>22.53</td>
<td>4.37</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

An independent samples t-test was conducted to determine the statistical significance of the difference in mean scores (Table 4.16). The delayed post-test results yielded a significant difference in scores for TB2 ($M = 29.80, SD = 4.77$) and TB1 ($M = 22.53, SD = 4.37$) conditions; ($t(28) = 4.3524, p < 0.001$). The effect size of the mean difference is large ($d = 1.59$) and of the correlation coefficient is large ($r = 0.62$).

In addition to the raw scores for each preposition, each experimental group’s results were calculated to determine individual mean scores for each preposition separately, along with standard deviations and p-values for each (Table 4.17). When running separate independent samples t-tests on the individual delayed post-test mean score differences for each preposition under TBI, the results (Table 4.17) show that the difference in the mean scores for the preposition *para* yielded TB2 *para* ($M = 6.73, SD = 1.84$) and TB1 *para* ($M = 4.47, SD = 2.23$) conditions; ($t(28) = 3.0275, p = 0.005$), and the preposition *a* yielded TB2 *a* ($M = 8.40, SD = 1.84$) and TB1 *a* ($M = 5.33, SD = 2.26$) conditions; ($t(28) = 4.0799, p < 0.001$). The effect size of the mean difference for the
preposition *para* (Table 4.18) is large ($d = 1.09$) and of the correlation coefficient is medium ($r = 0.48$). The effect size of the mean difference for the preposition *para* (Table 4.18) is large ($d = 1.49$) and of the correlation coefficient is large ($r = 0.60$).

**TABLE 4.17 TB2 vs. TB1 DELAYED POST-TEST PREPOSITION RESULTS**

<table>
<thead>
<tr>
<th>Preposition</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>$t$-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>para</em></td>
<td>high</td>
<td>6.73</td>
<td>1.84</td>
<td>15</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>4.47</td>
<td>2.23</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td><em>en</em></td>
<td>high</td>
<td>9.60</td>
<td>2.76</td>
<td>15</td>
<td>0.231</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>8.40</td>
<td>2.61</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td><em>por</em></td>
<td>high</td>
<td>5.07</td>
<td>2.31</td>
<td>15</td>
<td>0.283</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>4.33</td>
<td>1.23</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td><em>a</em></td>
<td>high</td>
<td>8.40</td>
<td>1.84</td>
<td>15</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>low</td>
<td>5.33</td>
<td>2.26</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.18 below breaks down the individual prepositional selection pattern on the delayed post-test for both experimental groups under TBI. These descriptive results report both the raw number of times the correct prepositions were chosen, as well as the raw number of times each of the incorrect prepositions were chosen. By showing what prepositions learners are selecting in place of the correct preposition, these data provide insights into the possible causes of each experimental group’s error patterns. The table is organized with one of the four target prepositions heading each of the rows as well as each of the columns. The preposition that heads each row represents the correct answer. The preposition heading each column represents which preposition was
actually chosen. The highest score that any experimental group could receive on any individual preposition was 180. This test followed the same coding protocol given in Section 4.2.1. The results shown in each block represent the total number of times each preposition was selected by that group, as well as the percentage out of a 180 that this number represents.

The correct number of answers along with the corresponding percentage is shaded in gray for each preposition. The results show that the greatest percentage difference in delayed post-test scores between the high and TB1 experimental groups occurred within the preposition *para* (56% vs. 37%) followed by *a* (70% vs. 44%), *por* (42% vs. 36%) and finally *en* (80% vs. 70%).

In terms of the error patterns, both the TB2 and TB1 experimental groups have similar selection patterns between the four prepositions. Both experimental groups under TBI scored the worst on *por* followed by *para*, *a* and finally they both scored the best with *en*. However, the results do show a moderately higher tendency from the TB1 experimental group to select the preposition *por* when the correct answer should have been *para*. The TB1 incorrectly selected *por* for *para* 44% of the time versus the TB2 experimental group selecting *por* for *para* just 35% of the time. The TB1 experimental group also had a moderately higher tendency to incorrectly select *en* for *por* (27% vs 18%) as well as to select *en* for *a* (17% vs. 7%).
4.3 Effects of the Interaction between WMC and Instructional Treatment on Both Short and Long-term Productive Knowledge

The presentation of the results of the effects resulting from the interaction between instructional treatment and IDs in WMC on both short and long-term productive knowledge is separated into pairs of experimental groups that belong to the same WMC group, but received different instructional treatments. This is done to isolate the interactional effects of each instructional treatment with each level of WMC (low and high). The presentation of the data begins with results from the immediate post-test of CB2 versus the TB2 experimental groups followed by the CB1 versus the TB1 experimental groups. This is followed by the results of the delayed post-test for the same pairs of experimental groups.
In summary, the results of a two-way MANOVA (Table 4.19) suggest that there is not a statistically significant difference in post-test scores based on the interaction between WMC and instructional treatment, \( F(12, 3) = 1.90, p = .328; \) Wilk's \( \Lambda = 0.116, \) partial \( \eta^2 = .88 \) with an observed power of .20. This was in relation to the main effects of the interaction between WMC and instructional treatment. However, there are two key findings revealed from the MANOVA. First, the observed power (.20) is well below the recommended level of .80 for quantitative experiments (see Cohen, 1988). These results mean that the present study lacked the power to detect statistical significance, if statistical significance actually exists. Technically, the power of the present study only has a 20% chance of detecting significance, and therefore, may be failing to reject the third null hypothesis, when it should be rejected (Type II error). The second key finding is the score of .88 under Partial Eta Squared (Partial \( \eta^2 \)) which reveals the proportion of variance in mean scores accounted for by the main effect of the interactions between WMC and instructional treatment. In other words, 88% of the variance in mean scores from the immediate and delayed post-tests is due to interactions between learner WMC in the two instructional techniques.

### 4.3.1 Effects of the interaction between high WMC and Instructional Treatment on Short-term Productive Knowledge

A total of 30 high WMC participants were taught the four polysemous prepositions that are the target of the present study. Prior to instruction, these 30 participants were randomly assigned to one of two different instructional techniques.
treatments, CLBI or TBI. Table 4.20 below shows the mean score and standard deviation of both the CB2 experimental group and the TB2 experimental group on the immediate post-test. These mean scores were determined by combining the individual scores from each participant on the immediate post-test. This test followed the same coding protocol given in Section 4.1.1. The results show that participants in the TB2 experimental group earned a higher mean score of 34.67 versus a mean score of 32.73 for participants in the CB2 experimental group. This works out to a difference in raw mean scores of 1.94 points or a 5.9% higher score for the TB2 experimental group versus the CB2 experimental group.

**Table 4.19 MANOVA Results for Interactions**

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
<th>Partial Eta Squared</th>
<th>Observed Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immediate score</td>
<td>treatment</td>
<td>5.155</td>
<td>22</td>
<td>.234</td>
<td>1.40</td>
<td>.444</td>
<td>.91</td>
<td>.16</td>
</tr>
<tr>
<td></td>
<td>memory</td>
<td>3.155</td>
<td>22</td>
<td>.143</td>
<td>.</td>
<td>.</td>
<td>1.00</td>
<td>.</td>
</tr>
<tr>
<td>Delayed score</td>
<td>treatment</td>
<td>4.210</td>
<td>21</td>
<td>.200</td>
<td>1.20</td>
<td>.508</td>
<td>.89</td>
<td>.14</td>
</tr>
<tr>
<td></td>
<td>memory</td>
<td>2.871</td>
<td>21</td>
<td>.137</td>
<td>.</td>
<td>.</td>
<td>1.00</td>
<td>.</td>
</tr>
<tr>
<td>Immediate Score x Delayed Score</td>
<td>treatment</td>
<td>3.795</td>
<td>12</td>
<td>.316</td>
<td>1.90</td>
<td>.328</td>
<td>.88</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>memory</td>
<td>3.006</td>
<td>12</td>
<td>.250</td>
<td>.</td>
<td>.</td>
<td>1.00</td>
<td>.</td>
</tr>
</tbody>
</table>

The immediate post-test results (Table 4.20) did not yield a significant difference in scores for CB2 ($M = 32.73$, $SD = 4.81$) and TB2 ($M = 34.67$, $SD = 3.29$) conditions; ($t(28) = 1.2893$, $p = 0.133$). The effect size of the mean difference is small ($d = 0.47$) and of the correlation coefficient is small ($r = 0.23$).
Table 4.20 CB2 vs. TB2 Immediate Post-test Results

<table>
<thead>
<tr>
<th>Treatment</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>d</th>
<th>r</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLBI</td>
<td>high</td>
<td>32.73</td>
<td>4.81</td>
<td>15</td>
<td>0.47</td>
<td>0.23</td>
<td>0.201</td>
</tr>
<tr>
<td>TBI</td>
<td>high</td>
<td>34.67</td>
<td>3.29</td>
<td>15</td>
<td>0.47</td>
<td>0.23</td>
<td>0.201</td>
</tr>
</tbody>
</table>

Table 4.21 contains a comparison of the separate preposition results from the immediate post-test for the two high WMC experimental groups. These results include mean scores, the percentage difference of the instructional treatment group that scored better as well as an independent samples t-test produced p-value for the differences for each pair of prepositional means. The difference in mean scores from the combined data set (Table 4.20) was not statistically significant (p = 0.133). When running an independent samples t-test on the individual mean score differences for each preposition separately, only the preposition *en* produces a statistically significant score for CB2 *en* (M = 9.13, SD = 1.36) and TB2 *en* (M = 10.27, SD = 1.67) conditions; (t(28) = -2.042, p = 0.051). The effect size of the mean difference for the preposition *en* (Table 4.22) is medium (d = 0.75) and of the correlation coefficient is medium (r = 0.35).

Table 4.22 below breaks down the individual prepositional selection pattern for both of the high WMC experimental groups. These descriptive results report both the raw number of times the *correct* prepositions were chosen, as well as the raw number of times each of the *incorrect* prepositions were chosen. By showing what prepositions learners are selecting in place of the correct preposition, these data provide insights into the possible causes of each experimental group’s error patterns. The table is organized with one of the four
target prepositions heading each of the rows as well as each of the columns. The preposition that heads each row represents the correct answer. The preposition heading each column represents which preposition was actually chosen. The highest score that any experimental group could receive on any individual preposition was 180. This test followed the same coding protocol given in Section 4.2.1. The results in each block show the total number of times each preposition was selected along with a percentage out of the possible 180 that this number represents. The number of correct answers along with the corresponding percentage correct is shaded in gray for each preposition.

**Table 4.21 CB2 vs. TB2 Immediate Post-test Preposition Results**

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>para</em></td>
<td>CLBI</td>
<td>15</td>
<td>7.93</td>
<td>2.46</td>
<td>0.694</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>8.27</td>
<td>2.12</td>
<td></td>
</tr>
<tr>
<td><em>en</em></td>
<td>CLBI</td>
<td>15</td>
<td>9.13</td>
<td>1.36</td>
<td>0.051</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>10.27</td>
<td>1.67</td>
<td></td>
</tr>
<tr>
<td><em>por</em></td>
<td>CLBI</td>
<td>15</td>
<td>7.73</td>
<td>2.22</td>
<td>0.813</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>7.53</td>
<td>2.36</td>
<td></td>
</tr>
<tr>
<td><em>a</em></td>
<td>CLBI</td>
<td>15</td>
<td>7.93</td>
<td>1.53</td>
<td>0.389</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>8.60</td>
<td>2.53</td>
<td></td>
</tr>
</tbody>
</table>

The results show that the greatest percentage difference in correct answers on the immediate post-test scores between the two high WMC experimental groups occurred within the preposition *en* (76% vs. 86%) followed by *a* (66% vs. 72%), *para* (66% vs. 69%) and finally *por* (64% vs. 63%). The TB2
experimental group scored better on three of out of the four target prepositions (i.e., *en*, *a*, *para*) with CB2 performing slightly better on the preposition *por*.

In terms of the error patterns, both the CB2 and TB2 experimental groups have similar selection patterns between the four prepositions. However, the results do show a moderately higher tendency from the TB2 experimental group to select the preposition *en* when the correct answer should have been *a* (11% vs. 2%). There is also a moderately higher tendency for the CB2 experimental group to select *para* when the correct answer should have been *a* (16% vs. 8%), as well as to select *por* when the correct answer should have been *a* (16% vs. 9%).

**Table 4.22 CB2 vs. TB2 Immediate Post-test Selection Patterns**

<table>
<thead>
<tr>
<th>CB2 Immediate Post-test Selection Patterns</th>
<th>para</th>
<th>en</th>
<th>por</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>para</em></td>
<td>119</td>
<td>7</td>
<td>42</td>
<td>12</td>
</tr>
<tr>
<td><em>en</em></td>
<td>9</td>
<td>137</td>
<td>8</td>
<td>26</td>
</tr>
<tr>
<td><em>por</em></td>
<td>32</td>
<td>21</td>
<td>116</td>
<td>11</td>
</tr>
<tr>
<td><em>a</em></td>
<td>28</td>
<td>5</td>
<td>28</td>
<td>119</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TB2 Immediate Post-test Selection Patterns</th>
<th>para</th>
<th>en</th>
<th>por</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>para</em></td>
<td>124</td>
<td>1</td>
<td>47</td>
<td>8</td>
</tr>
<tr>
<td><em>en</em></td>
<td>3</td>
<td>154</td>
<td>6</td>
<td>17</td>
</tr>
<tr>
<td><em>por</em></td>
<td>35</td>
<td>24</td>
<td>113</td>
<td>8</td>
</tr>
<tr>
<td><em>a</em></td>
<td>14</td>
<td>21</td>
<td>16</td>
<td>129</td>
</tr>
</tbody>
</table>

141
4.3.2 Effects of the Interaction between Low WMC and Instructional Treatment on Short-term Productive Knowledge

A total of 30 low WMC participants were taught the four polysemous prepositions that are the target of the present study. Prior to instruction, these 30 participants were randomly assigned to one of two different instructional treatments, CLBI or TBI. Table 4.23 below shows the mean score and standard deviation of both of the low WMC experimental groups on the immediate post-test. These mean scores were determined by combining the individual scores from each participant on the immediate post-test. This test followed the same coding protocol given in Section 4.1.1. The results show that participants in the CB1 experimental group earned a higher mean score of 28.27 versus a mean score of 25.47 for participants in the TB1 experimental group. This works out to a difference in raw mean scores of 2.80 or an 11.0% higher score for the CB1 experimental group versus the TB1 experimental group. The comparison of the results of the two low WMC experimental groups differs significantly from the comparison of the results of the two high WMC experimental groups. In the comparison of the immediate post-test results of the two high WMC experimental groups, the TB2 experimental group outscored the CB2 experimental group, but this time, the CB1 experimental group outscored the TB1 experimental group.

Table 4.23 CB1 vs. TB1 Immediate Post-test Results

<table>
<thead>
<tr>
<th>Treatment</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>d</th>
<th>r</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLBI</td>
<td>low</td>
<td>28.27</td>
<td>6.03</td>
<td>15</td>
<td>0.56</td>
<td>0.27</td>
<td>0.135</td>
</tr>
<tr>
<td>TBI</td>
<td>low</td>
<td>25.47</td>
<td>3.63</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
An independent samples *t*-test was conducted to compare mean differences between CB1 and TB1 (Table 4.23), and the immediate post-test results did not yield a significant difference in scores for CB1 (*M* = 28.27, *SD* = 6.03) and TB1 (*M* = 25.47, *SD* = 3.63) conditions; (*t*(28) = 1.5408, *p* = 0.135). The effect size of the mean difference is medium (*d* = 0.56) and of the correlation coefficient is small (*r* = 0.27).

Table 4.24 contains a comparison of the separate preposition results from the immediate post-test for the two low WMC experimental groups. These results include mean scores, the percentage difference of the instructional treatment group that scored better as well as an independent samples *t*-test produced *p*-value for the differences for each pair of prepositional means. The difference in mean scores from the combined data set (Table 4.23) were not statistically significant (*p* = 0.135). When running an independent samples *t*-test on the individual mean score differences for each preposition separately, the preposition *por* was statistically significant for CB1 *por* (*M* = 6.33, *SD* = 1.45) and TB1 *por* (*M* = 4.80, *SD* = 1.61) conditions; (*t*(28) = 2.741, *p* = 0.011). The effect size of the mean difference for the preposition *por* (Table 4.24) is large (*d* = 1.00) and of the correlation coefficient is medium (*r* = 0.45).

Table 4.25 below breaks down the individual prepositional selection pattern for both of the low WMC experimental groups. These descriptive results report both the raw number of times the *correct* prepositions were chosen, as well as the raw number of times each of the *incorrect* prepositions were chosen. By showing what prepositions learners are selecting in place of the correct
preposition, these data provide insights into the possible causes of each experimental group’s error patterns. The table is organized with one of the four target prepositions heading each of the rows as well as each of the columns. The preposition that heads each row represents the correct answer. The preposition heading each column represents which preposition was actually chosen. The highest score that any experimental group could receive on any individual preposition was 180. This test followed the same coding protocol given in Section 4.2.1. The results in each block show the total number of times each preposition was chosen. Along with these numbers, the percentage of the times that preposition was selected out of the possible 180 is provided.

**Table 4.24 CB1 vs. TB1 Immediate Post-test Preposition Results**

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>para</strong></td>
<td>CLBI</td>
<td>15</td>
<td>6.20</td>
<td>3.17</td>
<td>0.690</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>5.80</td>
<td>2.18</td>
<td></td>
</tr>
<tr>
<td><strong>en</strong></td>
<td>CLBI</td>
<td>15</td>
<td>8.40</td>
<td>1.96</td>
<td>0.368</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>7.67</td>
<td>2.41</td>
<td></td>
</tr>
<tr>
<td><strong>por</strong></td>
<td>CLBI</td>
<td>15</td>
<td>6.33</td>
<td>1.45</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>4.80</td>
<td>1.61</td>
<td></td>
</tr>
<tr>
<td><strong>a</strong></td>
<td>CLBI</td>
<td>15</td>
<td>7.33</td>
<td>2.50</td>
<td>0.886</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>7.20</td>
<td>2.40</td>
<td></td>
</tr>
</tbody>
</table>

The correct number of answers along with the corresponding percentage correct is shaded in gray for each preposition. The results show that the greatest percentage difference in correct answers on the immediate post-test scores between the CB1 and TB1 experimental groups occurred within the preposition
por (53% vs. 40%) followed by en (70% vs. 64%), para (52% vs. 48%) and finally a (61% vs. 60%). The CB1 experimental group scored better on all four of the target prepositions with the greatest difference in performance occurring with the preposition por.

In terms of the error patterns, both the CB1 and TB1 experimental groups have similar selection patterns between the four prepositions. However, the results do show a moderately higher tendency for the TB1 experimental group to select the preposition en when the correct answer should have been a (15% vs. 5%). There is also a moderately higher tendency for the CB1 experimental group to select para when the correct answer should have been a (22% vs. 12%).

**Table 4.25 CB1 vs. TB1 Immediate Post-test Selection Patterns**

<table>
<thead>
<tr>
<th></th>
<th>para</th>
<th>en</th>
<th>por</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>CB1 Immediate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patterns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>para</td>
<td>93 / 52%</td>
<td>10 / 5%</td>
<td>62 / 35%</td>
<td>15 / 8%</td>
</tr>
<tr>
<td>en</td>
<td>10 / 9%</td>
<td>126 / 70%</td>
<td>10 / 5%</td>
<td>28 / 16%</td>
</tr>
<tr>
<td>por</td>
<td>43 / 24%</td>
<td>29 / 16%</td>
<td>95 / 53%</td>
<td>13 / 7%</td>
</tr>
<tr>
<td>a</td>
<td>40 / 22%</td>
<td>9 / 5%</td>
<td>21 / 12%</td>
<td>110 / 61%</td>
</tr>
<tr>
<td>TB1 Immediate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selection</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patterns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>para</td>
<td>87 / 48%</td>
<td>6 / 3%</td>
<td>74 / 41%</td>
<td>13 / 7%</td>
</tr>
<tr>
<td>en</td>
<td>6 / 3%</td>
<td>115 / 64%</td>
<td>22 / 12%</td>
<td>37 / 21%</td>
</tr>
<tr>
<td>por</td>
<td>51 / 28%</td>
<td>32 / 18%</td>
<td>72 / 40.0%</td>
<td>25 / 14%</td>
</tr>
<tr>
<td>a</td>
<td>21 / 12%</td>
<td>28 / 15%</td>
<td>23 / 13%</td>
<td>108 / 60.0%</td>
</tr>
</tbody>
</table>
4.3.3 Effects of the Interaction Between High WMC and Instructional Treatment on Long-term Productive Knowledge

A total of 30 high WMC participants were taught the four polysemous prepositions that are the target of the present study. Prior to instruction, these 30 participants were randomly assigned to one of two different instructional treatments, CLBI or TBI. Table 4.26 below shows the mean score and standard deviation of both of the high WMC experimental groups on the delayed post-test. These mean scores were determined by combining the individual scores from each participant on the delayed post-test. This test followed the same coding protocol given in Section 4.1.1. The results show that participants in the CB2 experimental group earned a higher mean score of 31.87 versus a mean score of 29.80 for participants in the TB2 experimental group. This works out to a difference in raw mean scores of 2.07 points or a 6.9% higher score for the CB2 experimental group versus the TB2 experimental group. This is a change from the results of the immediate post-test where the TB2 experimental group outscored the CB2 experimental group by 5.9%. This is due to the fact that the TB2 experimental group’s mean score dropped by 4.87 points or 14.0% between the immediate and delayed post-tests, while the CB2 experimental group’s mean score dropped by 0.86 points or 2.6%.

Table 4.26 CB2 vs. TB2 Delayed Post-test Results

<table>
<thead>
<tr>
<th>Treatment</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>d</th>
<th>r</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLBI</td>
<td>high</td>
<td>31.87</td>
<td>5.67</td>
<td>15</td>
<td>0.40</td>
<td>0.19</td>
<td>0.289</td>
</tr>
<tr>
<td>TBI</td>
<td>high</td>
<td>29.80</td>
<td>4.77</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
An independent samples $t$-test was conducted to compare mean differences between CB2 and TB2 (Table 4.26), and the delayed post-test results did not yield a significant difference in scores for CB2 ($M = 31.87$, $SD = 5.67$) and TB2 ($M = 29.80$, $SD = 4.77$) conditions; ($t(28) = 1.0820$, $p = 0.289$). The effect size of the mean difference is small ($d = 0.40$) and of the correlation coefficient is small ($r = 0.19$).

Table 4.27 below contains a comparison of the separate preposition results from the delayed post-test for the two high WMC experimental groups. These results include mean scores, the percentage difference of the instructional treatment group that scored better as well as an independent samples $t$-test produced $p$-value for the differences for each pair of prepositional means. The difference in mean scores from the combined data set (Table 4.26) were not statistically significant ($p = 0.289$). When running an independent samples $t$-test on the individual mean score differences for each preposition separately, none of the mean differences is statistically significant. However, the effect size of the mean difference for the preposition *por* (Table 4.27) is medium ($d = 0.70$) and of the correlation coefficient is medium ($r = 0.33$).

Table 4.28 below breaks down the individual prepositional selection pattern for both high WMC experimental groups. These descriptive results report both the raw number of times the *correct* prepositions were chosen, as well as the raw number of times each of the *incorrect* prepositions were chosen. By showing what prepositions learners are selecting in place of the correct preposition, these data provide insights into the possible causes of each
experimental group’s error patterns. The table is organized with one of the four target prepositions heading each of the rows as well as each of the columns. The preposition that heads each row represents the correct answer. The preposition heading each column represents which preposition was actually chosen. The highest score that any experimental group could receive on any individual preposition was 180. This test followed the same coding protocol given in Section 4.2.1. The results in each block show the total number of times each preposition was selected along with the corresponding percentage that this number represents out of the possible 180.

**Table 4.27 CB2 vs. TB2 Delayed Post-test Preposition Results**

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>para</strong></td>
<td>CLBI</td>
<td>15</td>
<td>7.73</td>
<td>2.71</td>
<td>0.247</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>6.73</td>
<td>1.84</td>
<td></td>
</tr>
<tr>
<td><strong>en</strong></td>
<td>CLBI</td>
<td>15</td>
<td>9.27</td>
<td>1.10</td>
<td>0.670</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>9.60</td>
<td>2.76</td>
<td></td>
</tr>
<tr>
<td><strong>por</strong></td>
<td>CLBI</td>
<td>15</td>
<td>6.73</td>
<td>2.31</td>
<td>0.067</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>5.07</td>
<td>2.46</td>
<td></td>
</tr>
<tr>
<td><strong>a</strong></td>
<td>CLBI</td>
<td>15</td>
<td>8.00</td>
<td>2.62</td>
<td>0.632</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>8.40</td>
<td>1.84</td>
<td></td>
</tr>
</tbody>
</table>

The correct number of answers along with the corresponding percentage is shaded in gray for each preposition. The results show that the greatest percentage difference in correct answers on the delayed post-test scores between the CB2 and TB2 experimental groups occurred within the preposition **por** (56% vs. 42%) followed by **para** (64% vs. 56%), **en** (77% vs. 80%) and finally
a (68% vs. 70%). The TB2 experimental group scored better on the prepositions en and a with CB2 performing better on the prepositions para and por.

In terms of the error patterns, both the CB2 and TB2 experimental groups have similar selection patterns between the four prepositions. However, the results do show a higher tendency from the TB2 experimental group to select the preposition por when the correct answer was should have been para (35% vs. 26%). There is also a moderately higher tendency for the TB2 experimental group to select para when the correct answer should have been por (28% vs. 20%). Finally, the CB2 experimental group showed a slightly higher tendency to select para when the correct answer should have been a (17% vs. 10%).

**Table 4.28 CB2 vs. TB2 Delayed Post-test Selection Patterns**

<table>
<thead>
<tr>
<th>CB2 Delayed Post-test Selection Patterns</th>
<th>para</th>
<th>en</th>
<th>por</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>para</td>
<td>116 / 64%</td>
<td>6 / 3%</td>
<td>46 / 26%</td>
<td>12 / 7%</td>
</tr>
<tr>
<td>en</td>
<td>6 / 3%</td>
<td>139 / 77%</td>
<td>7 / 4%</td>
<td>28 / 16%</td>
</tr>
<tr>
<td>por</td>
<td>35 / 20%</td>
<td>26 / 14%</td>
<td>101 / 56%</td>
<td>18 / 10%</td>
</tr>
<tr>
<td>a</td>
<td>31 / 17%</td>
<td>4 / 2%</td>
<td>23 / 13%</td>
<td>122 / 68%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TB2 Delayed Post-test Selection Patterns</th>
<th>para</th>
<th>en</th>
<th>por</th>
<th>a</th>
</tr>
</thead>
<tbody>
<tr>
<td>para</td>
<td>101 / 56%</td>
<td>2 / 1%</td>
<td>63 / 35%</td>
<td>14 / 8%</td>
</tr>
<tr>
<td>en</td>
<td>6 / 3%</td>
<td>144 / 80%</td>
<td>5 / 3%</td>
<td>25 / 14%</td>
</tr>
<tr>
<td>por</td>
<td>50 / 28%</td>
<td>33 / 18%</td>
<td>76 / 42%</td>
<td>21 / 12%</td>
</tr>
<tr>
<td>a</td>
<td>18 / 10%</td>
<td>12 / 7%</td>
<td>24 / 13%</td>
<td>126 / 70%</td>
</tr>
</tbody>
</table>
4.3.4 Effects of the Interaction Between Low WMC and Instructional Treatment on Long-term Productive Knowledge

A total of 30 low WMC participants were taught the four polysemous prepositions that are the target of the present study. Prior to instruction, these 30 participants were randomly assigned to one of two different instructional treatments, CLBI or TBI. Table 4.29 below shows the mean score and standard deviation of both of the low WMC experimental groups on the delayed post-test. These mean scores were determined by combining the individual scores from each participant on the delayed post-test. This test followed the same coding protocol given in Section 4.1.1. The results show that participants in the CB1 experimental group earned a higher mean score of 26.60 versus a mean score of 22.53 for participants in the TB1 experimental group. This works out to a difference in raw mean scores of 4.07 points or a 18.1% higher score for the CB1 experimental group versus the TB1 experimental group. Like the TB2 experimental group, the TB1 experimental group experienced a greater decline in score between the immediate and delayed post-test than their CB1 counterparts. The TB1 experimental group’s score dropped by 2.94 points or 11.5% versus the CB1 experimental group which experienced a drop of just 1.67 points or 5.9%.

Table 4.29 CB1 vs. TB1 Delayed Post-test Results

<table>
<thead>
<tr>
<th>Treatment</th>
<th>WMC</th>
<th>Mean</th>
<th>SD</th>
<th>N</th>
<th>d</th>
<th>r</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLBI</td>
<td>low</td>
<td>26.60</td>
<td>6.70</td>
<td>15</td>
<td>0.72</td>
<td>0.34</td>
<td>0.059</td>
</tr>
<tr>
<td>TBI</td>
<td>low</td>
<td>22.53</td>
<td>4.37</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
An independent samples $t$-test was conducted to compare mean differences between CB1 and TB1 (Table 4.29), and the delayed post-test results show a tendency towards statistical significance in the difference in scores for CB1 ($M = 26.60$, $SD = 6.70$) and TB1 ($M = 22.53$, $SD = 4.37$) conditions; ($t(28) = 1.9706$, $p = 0.059$). The effect size of the mean difference is medium ($d = 0.72$) and of the correlation coefficient is medium ($r = 0.34$).

Table 4.30 contains a comparison of the separate preposition results from the delayed post-test for the two high WMC experimental groups. These results include mean scores, the percentage difference of the instructional treatment group that scored better as well as an independent samples $t$-test produced $p$-value for the differences for each pair of prepositional means. When running an independent samples $t$-test on the separate mean score differences for each preposition, the preposition *por* yielded a statistically significant score for CB1 *por* ($M = 5.93$, $SD = 1.53$) and TB1 *por* ($M = 4.33$, $SD = 1.23$) conditions; ($t(28) = 3.147$, $p = 0.004$). The effect size of the mean difference for the preposition *por* (Table 4.31) is large ($d = 1.15$) and of the correlation coefficient is medium-large ($r = 0.50$).

Table 4.31 below breaks down the individual prepositional selection pattern for both of the low WMC experimental groups. These descriptive results report both the raw number of times the *correct* prepositions were chosen, as well as the raw number of times each of the *incorrect* prepositions were chosen. By showing what prepositions learners are selecting in place of the correct preposition, these data provide insights into the possible causes of each
experimental group’s error patterns. The table is organized with one of the four target prepositions heading each of the rows as well as each of the columns. The preposition that heads each row represents the correct answer. The preposition heading each column represents which preposition was actually selected. The highest score that any experimental group could receive on any individual preposition was 180. This test followed the same coding protocol given in Section 4.2.1. The results in each block show the total number of times each preposition was selected along with the corresponding percentage that this number represents out of the possible 180.

**Table 4.30 CB1 vs. TB1 Delayed Post-test Preposition Results**

<table>
<thead>
<tr>
<th>Preposition</th>
<th>Treatment</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>t-test (two-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>para</strong></td>
<td>CLBI</td>
<td>15</td>
<td>5.33</td>
<td>3.66</td>
<td>0.440</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>4.47</td>
<td>2.23</td>
<td></td>
</tr>
<tr>
<td><strong>en</strong></td>
<td>CLBI</td>
<td>15</td>
<td>8.73</td>
<td>1.58</td>
<td>0.676</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>8.40</td>
<td>2.61</td>
<td></td>
</tr>
<tr>
<td><strong>por</strong></td>
<td>CLBI</td>
<td>15</td>
<td>5.93</td>
<td>1.53</td>
<td><strong>0.004</strong></td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>4.33</td>
<td>1.23</td>
<td></td>
</tr>
<tr>
<td><strong>a</strong></td>
<td>CLBI</td>
<td>15</td>
<td>6.60</td>
<td>2.90</td>
<td>0.193</td>
</tr>
<tr>
<td></td>
<td>TBI</td>
<td>15</td>
<td>5.33</td>
<td>2.26</td>
<td></td>
</tr>
</tbody>
</table>

The correct number of answers along with the corresponding percentage is shaded in gray for each preposition. The results show that the greatest percentage difference in correct answers on the delayed post-test scores between the CB1 and TB1 experimental groups occurred within the preposition **por** (49% vs. 36%) followed by **a** (55% vs. 44%), **para** (45% vs. 37%) and finally
en (73% vs. 70%). The CB1 experimental group scored better on all four of the target prepositions with the greatest difference in performance occurring with the preposition *por*.

In terms of the error rates, both the CB1 and TB1 experimental groups have similar selection patterns between the four prepositions. However, the results do show a moderately higher tendency for the TB1 experimental group to select the preposition *en* when the correct answer should have been *a* (17% vs. 8%).

**Table 4.31 CB1 vs. TB1 Delayed Post-test Selection Patterns**

<table>
<thead>
<tr>
<th></th>
<th>CB1 Delayed Post-test Selection Patterns</th>
<th>TB1 Delayed Post-test Selection Patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>para</em></td>
<td><em>en</em></td>
</tr>
<tr>
<td><em>para</em></td>
<td>80 / 45%</td>
<td>10 / 5%</td>
</tr>
<tr>
<td><em>en</em></td>
<td>10 / 5%</td>
<td>131 / 73%</td>
</tr>
<tr>
<td><em>por</em></td>
<td>35 / 20%</td>
<td>40 / 22%</td>
</tr>
<tr>
<td><em>a</em></td>
<td>39 / 21%</td>
<td>14 / 8%</td>
</tr>
</tbody>
</table>
CHAPTER 5: DISCUSSION

The following chapter contains a detailed discussion of the experimental results provided in CHAPTER 4, and how those results fit within current literature. It is divided into three major sections, each one organized around the results directly pertaining to one of the three research questions and related hypotheses. The following chapter overview summarizes the contents of each of these three major sections:

Section 5.1 contains a discussion of experimental results related to the statistical effects that the two instructional treatment had on productive knowledge. These two instructional methods are (1) translation-based instruction (TBI), which treats the multiple meanings of polysemes as arbitrary, discreet and unrelated; (2) cognitive linguistics-based instruction (CLBI), which treats the multiple meanings of polysemes as interrelated and motivated by an association to a common conceptual base via the processes of metaphor and metonymy. This major section is divided into two subsections, one concerned with short-term productive knowledge (Section 5.1.1) as observed in the experimental results from the immediate post-test, and the other (Section 5.1.2) which is concerned with long-term productive knowledge as observed in the experimental results from the delayed post-test. The focus of the discussion in both of these subsections is on the observed effects that each instructional treatment had on
productive knowledge of the target prepositions (i.e., CLBI participants vs. TBI participants).

Section 5.2 contains a discussion of experimental results related to the effects of WMC on the acquisition of productive knowledge for both pairs of experimental groups that received the same instructional treatment, but were separated by scores in working memory capacity (i.e., CB2 vs. CB1; TB2 vs. TB1). Section 5.2.1 is focused on short-term productive knowledge as observed in the immediate post-test results and section 5.2.2 is focused on long-term productive knowledge as observed in the delayed post-test results. The focus of the discussion in both of these subsections on understanding how individual differences in WMC affect the acquisition of productive knowledge of the target prepositions.

Section 5.3 contains a discussion of the experimental results related to the observed effects resulting from the interaction of individual differences in WMC with two different instructional treatments (CLBI and TBI). The main focus of the discussion in on understanding how productive knowledge of the target prepositions is affected by the interaction between individual differences in WMC and the type of instructional treatment that learners received.

5.1 Discussion of the Effects of Instructional Treatment on Productive Knowledge

The following section contains a discussion of experimental results related to the observed effects of instructional treatment on productive knowledge of the four target prepositions. This major section is divided into two subsections. The
first subsection contains a discussion concerning experimental results related to short-term productive knowledge as observed in the immediate post-test results, and the second subsection contains a discussion concerning experimental results related to long-term productive knowledge as observed in the delayed post-test results. The focus of the discussion in both of these subsections is on answering the second research question (i.e., Do ab initio learners under CLBI versus TBI perform differently in the acquisition of both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions?)

In summary, experimental results suggest that for ab initio learners, CLBI and TBI do not differentially affect short-term productive knowledge of the target prepositions, but these instructional treatments do have a marginally differential effect on long-term productive knowledge. In addition, the observed results from the short-term productive test of the present study do not mirror the results from a number of previous studies that show beneficial effects of CLBI over TBI in the acquisition of polysemous words (cf. Khodadady & Khaghaninizhad, 2011; Lam, 2009; Makni, 2013; Morimoto & Loewen, 2007; Touplikioti, 2007). The primary reason for the results of the present study differing from previous studies is due to a variable directly associated with the participants involved. The present study had the requirement that participants must have no prior knowledge of the target language (i.e., Spanish) or any related language (e.g. French, Portuguese, Italian, etc.). This was done to isolate each instructional treatment as the only source of instruction for the learners. However, using ab initio learners was not a
requirement of participants in any previous studies, and in fact, they used participants who already had substantial knowledge of the target language.

5.1.1 Discussion of the Effects of Instructional Treatment on Short-term Productive Knowledge

This section discusses a comparison of the immediate post-test results between the participants under CLBI versus the participants under TBI. A $t$-test was conducted to compare mean differences between all participants under each of instructional treatments, and the immediate post-test results did not yield a significant difference in scores for participants under CLBI ($M = 30.50$, $SD = 5.90$) and participants under TBI ($M = 30.07$, $SD = 6.09$) conditions; ($t(58) = 0.129$, $p = 0.846$). The effect size of the mean difference is also small ($d = 0.07$) as well as the correlation coefficient ($r = 0.04$). Since the only variable that distinguishes the two groups of participants is the instructional treatment that they received, the observed results suggest that in terms of short-term productive knowledge, there is no observable difference in the rate of acquisition under one instructional method over the other for ab initio learners. This is only in terms of the combined mean scores from all four target prepositions, not the separate mean scores for the individual target prepositions, which will be discussed later.

The lack of any statistically significant difference in the combined mean scores of these two different instructional treatment groups is significant for two reason, (1) the two instructional treatments are significantly different from each other, and (2) current literature contains a number of diverse studies that report beneficial effects of CLBI over TBI. To address the first reason, it must first be
clearly understood how CLBI and TBI differ from each other. CLBI is designed to start the process of learning a polysemous word by focusing on a single common conceptual core that runs through all its other meanings, whereas TBI involves having learners memorize a list of the various ways in which a polysemous word can be translated into their L1 without any attempt to find a unifying conceptual core for all those translations. Touplikioti (2007) describes CLBI as a motivated approach to the teaching of polysemes that provides a systematic model for linking the multiple meanings of these highly complex words back to one common proto-meaning. This is in sharp contrast to translation-based instruction (TBI) that she defines as instruction based on the technique of memorization and repetition where students are simply guided to learn list of translated L2 to L1 words without any guidance as to how these various translated meanings interrelate and behave.

These differences between CLBI and TBI lead to differing advantages and disadvantages that offset each other when two independent sets of ab initio learners are separately instructed under only one of these instructional treatments. In terms of the advantages of CLBI, it can be argued from previous studies that it lessens the learning burden of polysemes by reducing the number of discrete pieces of semantic information that learners must maintain in working memory. This has led to greater gains in proficiency for learners taught under CLBI over TBI in a number of diverse studies. For example, Khodadady and Khaghaninizhad (2011) reported greater gains under CLBI versus TBI in the teaching of the polysemous French verb arriver and the polysemous French
preposition *sur* to 49 L1 Farsi speakers who had already had six semesters of French. Touplikioti (2007) reported greater gains under CLBI versus TBI in the teaching of the polysemous English verbs *make* and *do* to L1 Greek speakers who were at the low intermediate level in English at the start of the study. Makni (2013) reported greater gains under CLBI over TBI in the acquisition of the English polysemes *hand, break, head, over, burn, push, beyond* and *root* to 40 L1 Arab speakers who were at a low-intermediate level in English. Morimoto and Loewen (2007) reported greater gains under CLBI versus TBI in the acquisition of two L2 English polysemes *break* and *over* to 58 L1 Japanese speakers. Finally, Lam (2009) compared the effectiveness of teaching the Spanish prepositions *por* and *para* according to translation-based techniques versus cognitive linguistics-based techniques. In her study, participants experienced greater gains under CLBI, especially on a delayed post-test.

All of the studies cited above report greater gains for participants under cognitive linguistics-based techniques over translation-based techniques, however; they involved participants with at least intermediate level (e.g. B1-B2 CEFR Scale) proficiency in the language of the target prepositions, whereas the present study involved participants with no prior knowledge of the target language or of any related language (i.e., *ab initio* learners). In addition, the target vocabulary used in each of these studies are all high frequency words that would have been somewhat familiar to the participants prior to the beginning of the studies. This fact is most evident from the pre-tests in these studies that showed that participants already had measurable knowledge of the target
structures. Because the studies report that CLBI is a novel technique, it can be presumed that any exposure that participants would have had prior to the experiment would be from more translation-based techniques that require repetition and memorization. This means that the participants in all of the previous studies cited above were not separated by experimental groups in which their only exposure to the target prepositions came under either CLBI or TBI, but rather, these experimental groups would have actually been made up of participants who either continue to be exposed to the target prepositions under TBI, or participants who have some previous TBI exposure plus the advantage of a new exposure under CLBI. The lack of control over the learner’s previous knowledge of the target vocabulary weakens the internal validity of these previous studies. For example, Lam’s study on the acquisition of por and para under CLBI versus TBI involved intermediate level Spanish students in their fourth semester of university studies. It is impossible that after that much formal instruction to Spanish that they have not already gained a good bit of translation-based knowledge of por and para. Therefore, participants under CLBI in her experiment would have had an advantage in that they would be receiving an additional method to negotiate meaning of these polysemous spatial prepositions, whereas the comparison group just got more of the same translation and memorization methods.

In terms of the pedagogical disadvantages of CLBI, Beréndi, Csábi, and Kovecses (2008) reported an especially high standard deviation (SD 6.90) in the scores of the experimental group taught according to cognitive linguistic
principles. A similarly higher standard deviation for CLBI versus TBI experimental groups was observed in the present study. Table 5.1 shows that there is a large differential in SD across experimental groups with differing WMC, and across experimental groups who received different instruction treatments. In general, the results suggest that experimental groups of ab initio learners under CLBI will produce larger SDs than those under TBI, in addition, experimental groups of ab initio learners with lower WMC will produce larger SDs than their high WMC counterparts.

**Table 5.1 CLBI versus TBI Immediate Post-test Standard Deviations**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>WMC</th>
<th>SD</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLBI</td>
<td>low</td>
<td>6.03</td>
<td>15</td>
</tr>
<tr>
<td>CLBI</td>
<td>high</td>
<td>4.81</td>
<td>15</td>
</tr>
<tr>
<td>TBI</td>
<td>low</td>
<td>3.63</td>
<td>15</td>
</tr>
<tr>
<td>TBI</td>
<td>high</td>
<td>3.29</td>
<td>15</td>
</tr>
</tbody>
</table>

Beréndi, Csábi, and Kovecses argued that the presence of the much higher SD under CLBI suggests that there may be considerable individual differences when it comes to successfully applying cognitive-linguistic techniques autonomously. Boers (2004: 223) suggested that not all IDs in cognitive styles, (i.e., preferred strategy to organize and process information), are equally receptive to cognitive linguistics based instruction. Boers suggested that analytic learners who prefer to process information in separate parts and those who prefer to use images were more likely to benefit from an explanation of metaphors based on a concrete scene than holistic learners who process
information as large integrated chunks or verbalizers who prefer using words. In the present study, the cognitive styles of the learners may also have influenced their receptiveness to the treatment. In other words, learners are not only being exposed to a new foreign language target structure, they are also being exposed to a whole new way of understanding meaning, and a single lesson may not provide adequate practice in how to apply CLBI techniques to the process of learning the complex network of meanings associated with each of the four target polysemes.

In contrast, TBI employs the very common pedagogical techniques of repetition and memorization. Language learners are highly accustomed to being simply guided to memorize a list of translated L2 to L1 words without the worry for how these various meanings may all interrelated back to a single proto-meaning. Therefore, participants in the present study who were instructed under TBI were able to immediately apply the technique to the task of learning the four target prepositions. This is also seen in their smaller standard deviations. However, this advantage comes with the disadvantage that learners under TBI have to memorize a long list of seemingly arbitrary and discrete uses for each polysemous preposition, and this taxes the resources of their working memory much more than CLBI.

All of the above background information helps to provide an explanation for why the observed results from the short-term productive test of the present study contradict the results from a number of previous studies that show beneficial effects of CLBI over TBI in the acquisition of polysemous words. The
most salient reason is that the present study was divided between two experimental groups made up of *ab initio* learners who received exposure to the target prepositions exclusively from just one of the instructional treatments. Previous studies were actually using two groups, (1) a control group made up of participants under continued translation-based instruction, and (2) an experimental group made up of participants with some prior translation-based instruction plus the advantage of cognitive linguistics-based instruction.

Even though the results of the present study do not reflect the findings of previous studies, this does not necessarily mean that CLBI is not a more effective method for teaching polysemes. The evidence from these studies for the beneficial effect of CLBI are still compelling, rather, the results of the present may suggest that the full benefits of CLBI cannot be realized in a single lesson. This is especially true when applying CLBI to the many uses of polysemous spatial prepositions.

The discussion thus far has been centered on the combined mean scores for all four target prepositions, however, in examining the results for each of the prepositions separately, a slightly different picture emerges. The immediate post-test mean scores for three out of the four target prepositions (i.e., *para, en, a*) are very close, just as the combined mean scores were. However, the preposition *por* stands out with a much wider mean score differential (Table 4.2). However, the effect size of the mean difference for the preposition *por* is small ($d = 0.39$) and the correlation coefficient is small ($r = 0.19$). This wider difference in mean score for the preposition *por* seems to suggest some pedagogical
advantage of CLBI over TBI for this specific preposition. Arguably, this is due to
the fact that *por* has the most complex network of meanings among the four
target prepositions (as seen in its lower scores), and therefore it would seem to
place more demand on participant’s WMC. As an instructional treatment, CLBI
seems to have attenuated that demand, and in turn produced higher immediate
post-test scores for participants who were taught under it.

In terms of the error patterns, both CLBI and TBI instructional groups have
similar error selection patterns between the four prepositions. However, the
descriptive data (Table 4.3) does show a modestly higher tendency for
participants under CLBI to incorrectly select the preposition *para* when the
correct answer should have been *a* (CLBI 19% incorrect vs. TBI 10% incorrect).
This is the result of confusing the cognitive linguistics-based proto-meaning of
*para* where the focus element (F) spatially relates to the ground element (G) as a
destination or end goal, with the cognitive linguistics-based proto-meaning of *a*
where the focus element (F) spatially relates to the ground element (G) as simply
the direction of movement on an axis line. However, from the standpoint of an *ab
initio* learner, it takes time to recognize and apply these distinctions in meanings
productively. Figures 5.1 and 5.2 are the original work of the present author, but
were informed by the works of Brugman (1981), Delbecque (1996), Huerta
(2009), Lam (2009), Langacker (1987) and Taylor (2003). These figures visually
illustrate the cognitive-linguistics based proto-scene for each of these
prepositions, and they help to explain how learners may confuse these two
visually similar image-schemas.
The immediate post-test data also shows a modestly higher tendency from the TBI experimental group to incorrectly select the preposition *en* when the correct answer should have been *a* (TBI 13% incorrect vs CLBI 4% incorrect). This is due to the fact that depending on the context, both *en* and *a* may be translated from Spanish into English as the prepositions *to, in* or *on*, and this overlap in the L2 to L1 translation under TBI causes confusion for L1 English speakers in knowing when *a* or *en* is appropriate for a given spatial scene.

5.1.2 Discussion of the Effects of Instructional Treatment on Long-term Productive Knowledge

This section discusses a comparison of the delayed post-test results between the participants under CLBI versus the participants under TBI. In a comparison of the delayed post-test scores from the participants under each of
these instructional treatments, the effect size of the mean difference is small ($d = 0.48$) and of the correlation coefficient is small ($r = 0.23$). This is an increase in effect size in comparison to the immediate post-test (Table 4.1). An independent samples $t$-test was conducted to compare mean differences between participants under each of instructional treatments, and the delayed post-test results did not yield a significant difference in scores for participants under CLBI ($M = 29.23$, $SD = 6.73$) and TBI ($M = 26.17$, $SD = 6.02$) conditions; ($t(58) = 1.8562$, $p = 0.069$). Since the only variable that distinguishes the two groups of participants is the instructional treatment that they received, the observed delayed post-test results suggest that in terms of long-term productive knowledge, there is only a small beneficial effect of CLBI over TBI for $ab initio$ learners.

However, in examining the changes in raw score, it can be seen that the decline in the score between the immediate and delayed post-test was observably less for participants under CLBI than for participants under TBI. Participants under CLBI experienced a decline of 1.27 points or 4.2% between the immediate and delayed post-test. By comparison, participants under TBI experienced a greater decline of 3.90 points or 13.0% between the immediate and delayed post-test scores. This roughly works out to a rate of decline from immediate to delayed post-test for participants under TBI that is three times as great as for participants under CLBI. Although the difference in the two groups post-test scores did not reach a $p$-value that was statistically significant, the change in raw scores does suggest some pedagogical advantage of CLBI over TBI in terms of long-term retention of the meanings of the target prepositions.
These results also mirror those in Lam (2009) which showed significant advantages on a delayed post-test for participants under CLBI versus participants under TBI in the acquisition of *por* and *para*. Lam does not explain the reason why participants under CLBI performed better on the delayed post-test, but the present author argues that the advantage observed in long-term productive knowledge under CLBI is likely due to the fact that this instructional method requires memory of fewer discrete details that can easily decay over time.

In examining the delayed post-test results for each of the prepositions separately, a slightly different picture emerges. The immediate post-test mean scores for three out of the four target prepositions (i.e., *para*, *en*, *a*) are still very close, just as the combined mean scores were. However, the preposition *por* stands out with a much wider mean score differential (Table 4.5). An independent samples *t*-test yielded a statistically significant difference in scores for participants under CLBI *por* (*M* = 6.33, *SD* = 1.97) and TBI *por* (*M* = 4.70, *SD* = 1.96) conditions; (*t*(58) = 3.2127, *p* = 0.002). The effect size of the mean difference of the preposition *por* is medium (*d* = 0.83) and of the correlation coefficient is medium (*r* = 0.38). Since the only variable that distinguishes the two groups of participants is the instructional treatment that they received, the observed delayed post-test results for the preposition *por* seem to suggest some pedagogical advantage of CLBI over TBI for this specific preposition. This may be due to the fact that *por* has the most complex network of meanings among the four target prepositions, as observed in the lower productive scores produced for
this preposition. These lower productive scores (Table 4.5) suggest that the preposition *por* places more demand on individual participant’s WMC. As an instructional treatment, CLBI seems to have attenuated that demand, and in turn produced higher delayed post-test scores for participants who were taught under it.

Like the immediate post-test results, the data shows that the greatest percentage difference in delayed post-test scores between the CLBI and TBI instructional groups occurred within the preposition *por* with CLBI scoring with 53% accuracy versus TBI at 39% accuracy. With the preposition *para*, the CLBI experimental group outscored the TBI experimental group by 7%. TBI participants’ lower score for *para* is likely due to confusion that L1 English learners have in distinguishing when to use *por* versus *para* as these learners tend to conflate the meaning of these two prepositions into the single English preposition *for*. The higher scores in the preposition *para* for participants under CLBI may be an indirect reflection of their greater acquired productive knowledge of the preposition *por* as L1 English speakers have a tendency to confuse productive usage of these two prepositions. However, the English preposition *for* does not semantically overlap with all of the use of the Spanish prepositions *por* and *para*. For the preposition *a*, the CLBI experimental group outscored the TBI experimental group by 5% points. Finally, with the preposition *en*, both groups scored exactly the same with 270 correct responses or 75% accuracy. The preposition *en* arguably has the least complex network of meanings as observed in the higher scores for this preposition (Table 4.5), and is also somewhat
cognate with the English prepositions in and on, and this appears to have helped all participants, regardless of instructional treatment, to yield high scores in this preposition.

In terms of the error patterns, both CLBI and TBI instructional groups have similar selection patterns between the four prepositions. However, the results do show a modestly higher tendency from the CLBI experimental group to incorrectly select the preposition *para* when the correct answer should have been *a* (19% incorrect vs 14% incorrect). This is a repeat of the same error selection problem discussed under the immediate post-test results, and is certainly due to the continued confusion between the proto-meanings of these two prepositions as scene in the proto-scene for both. The delayed post-test results also show a modestly higher tendency from the TBI instructional group to incorrectly select the preposition *en* when the correct answer should have been *a* (12% incorrect vs 5% incorrect). Again, this is a repeat of the same error selection problem discussed under the immediate post-test results, and is certainly due to the fact that depending on the context, both *en* and *a* may be translated from Spanish into English as the prepositions *to, in or on*, and this overlap in the L2 to L1 translation under TBI causes confusion for L1 English speakers in knowing how to distinguish *a* and *en* in a given spatial scene. Overall, the data shows the greatest difference in scores occurred with the preposition *por*, as the TBI experimental group incorrectly selected *para* for *por* 26% of the time, and *en* for *por* 23% of the time. In using L2 to L1 translations to explain the meanings of
these prepositions, TBI has caused the all too common confusion that L1 English speakers have between *por* and *para* and even *en*.

This section is concluded with an examination of the immediate and delayed post-test results for CLBI versus TBI through the filter of the first null hypothesis (H₀) (i.e., *Ab initio learners under CLBI versus TBI do not perform differently in the acquisition of both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions*). The observed results of the present study suggest that the second null hypothesis should be accepted in terms of short-term productive knowledge. There are results from a number of previous studies that show learners performing better in the acquisition of both short and long-term productive knowledge under CLBI over TBI, however, these studies did not use *ab initio* learners. The use of *ab initio* learners had a profound effect on the present study as compared to previous studies. The results from the delayed post-test provide enough evidence to not fully accept the first null hypothesis in regards to long-term productive knowledge, especially if observing the separate delayed post-test results for the preposition *por*. A possible alternative hypothesis in regards to long-term productive knowledge would read, “*Ab initio learners under CLBI will moderately outperform learners under TBI in the acquisition of long-term productive knowledge of L2 Spanish polysemous spatial prepositions, especially the preposition por.*”
5.2 Discussion of the Effects of WMC on Short and Long-term Productive Knowledge

The discussion of the results of the effects of WMC on both short and long-term productive knowledge is separated by instructional treatments. This is done to isolate WMC as the main effect that accounts for any differences in experimental results. In summary, experimental results provide a clear answer of yes to the second research question for both instructional treatments (i.e., Do individual differences in ab initio learner WMC differentially affect the acquisition of both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions?) Observed immediate and delayed post-test results show that for ab initio learners under both CLBI and TBI, higher scores in WMC correlate to increased acquisition of both short and long-term productive knowledge of the target prepositions. For the present study, participants who scored in the top one-third (i.e., 61-75) of all participants on a pre-test to measure their WMC outscored participants who scored in the bottom one-third (i.e., 31-45). Overall, these results fit into current literature, as there is a growing body of research that demonstrates that adult L2 learners rely on certain types of cognitive resources, especially WMC, to attain high levels of proficiency in an L2 (e.g. DeKeyser, 2000; Harley and Hart, 1997, 2002; Ross, Yoshinaga and Sasaki, 2002).
5.2.1 Discussion of the Effects of WMC on Short-term Productive Knowledge

This section discusses the results of an immediate post-test conducted to compare the acquisition of short-term productive knowledge of four Spanish polysemous spatial preposition under two instructional treatments and two levels of working memory capacity. The immediate post-test results (Tables 4.7 and 4.10) yielded a significant difference in scores for CB2 ($M = 32.73$, $SD = 4.81$) and CB1 ($M = 28.27$, $SD = 6.03$) conditions; ($t(58) = 3.1670$, ($p = 0.003$), as well as for TB2 ($M = 34.67$, $SD = 3.29$) and TB1 ($M = 25.47$, $SD = 3.63$) conditions; ($t(28) = 7.2731$, $p < 0.001$). The effect size of the mean difference under CLBI (Table 4.7) was medium ($d = 0.82$) and of the correlation coefficient was medium ($r = 0.38$). The effect size of the mean difference under TBI (Table 4.10) was large ($d = 2.66$) and of the correlation coefficient was large ($r = 0.88$). Since WMC is the only independent variable that distinguishes these two pairs of experimental groups from each other, the results of the immediate post-test suggest that possessing high WMC gives learners some cognitive advantage in acquiring short-term productive knowledge of the four polysemous Spanish spatial prepositions that are the target of the present study.

There are no studies with which to directly compare the immediate post-test results, since no previous study has specifically examined the effects of WMC on the acquisition of short-term productive knowledge of polysemes under CLBI or TBI. However, there are studies that had similar results as the present study that investigated the effect of WMC on the acquisition of productive speech
as well as vocabulary acquisition. For example, Mota (2003) investigated whether there was a relationship between WMC and L2 speech production. The participants were 13 advanced learners of English as a second language. The results of the study revealed that three aspects of speech production, fluency, accuracy and complexity all correlated positively with higher working memory. Four aspects of speech production were assessed: fluency, accuracy, complexity, and weighted lexical density. Statistical analyses showed that working memory correlated positively with fluency, accuracy, and complexity. Mendonça (2003) investigated the relationship between WMC and the retention of L2 English vocabulary by L1 Portuguese speakers. Statistical results revealed that learners with higher working memory capacity were able to retain more L2 English vocabulary than low working memory learners.

In addition to these specific studies, the results of the present study fit our general understanding of the role of WMC in SLA. Kormos (2013, p. 142) argues that cognitive IDs, especially WMC, have a demonstrated influence on the various processing stages involved in language acquisition. These stages can be found in Table 2.4. Because acquiring productive knowledge of polysemes involves all four of these processing stages, and WMC is implicated as an important cognitive factor in each, it is no surprise that the scores on the immediate post-test and even the delayed post-test showed higher levels of productive knowledge for learners with high WMC over those with low WMC.

The results of the present study strongly support the argument that IDs may cause great variability between learners. These results also fit current
literature, which shows that IDs, especially WMC, are one of the primary causes of observed variability in adult L2 proficiency. WMC is used here according to the definition by Gathercole & Baddeley (1993) and Baddeley (2003), which describe this component of aptitude as a system of “temporary storage and manipulation of information that is assumed to be necessary for a wide range of complex cognitive activities” (Baddeley, 2003, p. 189).

Canner (2013) revealed correlations between language learner aptitude, especially working memory and levels of proficiency in L2 Russian attained in naturalistic versus formal learning contexts. The focus of this study was the learner’s oral proficiency in three distinct areas related to both fluency and accuracy: (1) a measure of fluency in terms of the number of meaningful syllables uttered per minute, (2) an assessment of accuracy in morphology and syntax measured by eliciting specific and frequently-uttered constructions that contain the major elements necessary for effective, native-like speech, and (3) overall ability in terms of both command of vocabulary and effective use of major prepositions commonly used in Russian.

Abrahamsson and Hyltenstam (2008) investigated the L2 proficiency and language aptitude of 42 near-native L2 speakers of Swedish. These participants were judged by mother-tongue speakers of Swedish as also being native speakers. The results of their study suggest that a high degree of language aptitude is required if adult learners are to reach a L2 proficiency that is indistinguishable from that of native speakers.
Engle, Kane and Tuholsky (1999) argue for a model of WMC in which WMC is an individual’s capacity to maintain controlled attention in the face of distraction. The results of the present study suggest that the individual participant’s ability to maintain controlled attention in the face of being distracted by all the information related to the many uses of the polysemous target prepositions greatly affect their acquisition of productive knowledge. Participants are dependent on their individual WMC to briefly store and process the new linguistic input during the instructional treatment in order to have that knowledge available during both the immediate and delayed post-tests. As a group, participants who pre-tested with high WMC were able to develop a higher level of proficiency in both short-term and long-term productive knowledge under both CLBI and TBI than their low WMC counterparts.

We know from current literature that IDs in WMC affect a learner’s rate of acquisition as well as ultimate attainment (Bley-Vroman’s, 1990, Skehan, 1998). This is in regard to overall proficiency in the L2. However, the results of the present study suggest that WMC may have differential effects on the acquisition of short-term productive knowledge depending on the specific target structure in focus. It appears that the more complex the lexical network of a given word, the greater the main effects of WMC will be on the acquisition of that word. In taking a deeper look at the separate immediate post-test results for each of the four target prepositions, for both participants under CLBI and TBI, the lowest raw scores occurred within the preposition *por* for three out of the four experimental groups (CB2, TB2 and TB1). For the CB1 experimental group, *para* yielded the
lowest raw scores on the immediate post-test. Under CLBI, only the preposition *por* yielded a *p*-value that was statistically significant. The immediate post-test results (Table 4.8) yielded a significant difference in scores for CB2 preposition *por* (*M* = 7.73, *SD* = 2.22) and TB1 preposition *por* (*M* = 6.33, *SD* = 1.45) conditions; (*t*(28) = 2.0466, *p* = 0.050). The effect size of the immediate post-test mean difference for the preposition *por* is medium (*d* = 0.75) and of the correlation coefficient is medium (*r* = 0.35).

When running separate independent samples *t*-test on the individual mean score differences for each preposition under TBI, three out of the four prepositions met the criteria for statistical significance. This suggests that the main effects of WMC are more prominent for experimental groups under TBI. The difference in the immediate post-test mean scores (Table 4.12) for the preposition *para* yielded a significant difference in scores for TB2 *para* (*M* = 8.27, *SD* = 2.12) and TB1 *para* (*M* = 5.80, *SD* = 2.18) conditions; (*t*(28) = -3.143, *p* = 0.004). The preposition *en* yielded a significant difference in scores for TB2 *en* (*M* = 10.27, *SD* = 1.36) and TB1 *en* (*M* = 7.67, *SD* = 1.96) conditions; (*t*(28) = 4.2210, *p* < 0.001), and the preposition *por* yielded a significant difference in scores for TB2 *por* (*M* = 7.53, *SD* = 2.36) and TB1 *por* (*M* = 4.80, *SD* = 1.61) conditions; (*t*(28) = -3.708, *p* < 0.001). The effect size of the mean difference for the preposition *para* (Table 4.11) is large (*d* = 1.15) and of the correlation coefficient is medium-large (*r* = 0.50). The effect size of the mean difference for the preposition *en* (Table 4.11) is large (*d* = 1.54) and of the correlation coefficient is large (*r* = 0.61). The effect size of the mean difference for the
preposition *por* (Table 4.11) is large ($d = 1.35$) and of the correlation coefficient is large ($r = 0.56$).

Overall, the observed immediate post-test results suggest that the preposition *por* places the greatest strain on individual working memory, as the majority of participants scored lowest in it. *Por* also produced the most statistically significant difference in scores in the immediate post-test between the high and low WMC experimental groups for both participants taught under CLBI and TBI. The results of the present study suggest that the preposition *por* has the most complex polysemy network among the four target polysemous spatial prepositions, and that the main effects of WMC are more prominent under it as a result of its complexity. There is evidence that some types of lexical items place a greater learning burden on second language learners over others. For example, empirical evidence demonstrates that polysemes demand a great depth of lexical understanding, and adult L2 learners rarely reach a native-like depth of knowledge of words which have multiple interrelated meanings (e.g., Bensoussan & Laufer, 1984; Pavlenko, 2009). Bensoussan and Laufer (1984) studied whether or not the meaning of some types of words are more easily understood and guessed than others. They attempted to answer this question by examining 60 first year EFL learners’ ability to accurately translate the meaning of five different semantically challenging lexical items. These five types of lexical items were (1) polysemes, (2) morphological troublemakers (i.e., lexical items that are composed of multiple morphemes), (3) idioms, (4) synophone / homophone and (5) false cognates. The results of this study show that from
among these five different challenging lexical items, learners most frequently mistranslated polysemes. The study also shows that learners vary greatly in how well they can recognize and use the multiple meanings of a polysemous word.

Two important questions that this and other studies fail to answer are (1) why learners perform worse in the acquisition of certain lexical items over others, and (2) what accounts for the differential rates of acquisition of polysemes see between learners. In regards to the first question, evidence from the present study suggest that each lexical item will differentially tax the limited storage of the individual learner's WMC, and that this differential taxing of WMC is directly related to the degree of complexity of each lexical item’s semantic network. In other words, the greater the amount of discrete pieces of information that the lexical item demands be kept in working memory, the greater the challenge for the learner to fully to acquire and accurately use that lexical item in language production. Ultimately, this means that there will be different rates of acquisition for each lexical items in accordance with the complexity of its semantic network. In addition, it may be possible to place all the lexical items of a given language on a cline that gradates from those that place the highest demand on WMC down to those that place the lowest demand on WMC. Nation (2001) claims that each word in the lexicon of language bears a unique learning burden for the learner. He defines learning burden as “the amount of effort required to learn it” (p.23). Nation (1990) argues that the learning burden of a word decreases the more a word represents patterns and knowledge that the learner is already familiar with. Nation’s results directly apply to the results of the present study where the
preposition *por* bears a greater learning burden for L1 English speakers than
does the preposition *en*. This is revealed in the scores for the two prepositions,
and is possible due to the fact that the preposition *en* represents spatial scenes
already known with the English preposition *in*, while the preposition *por* does not
match up well with any single English preposition. Kwon (2005) advances a case
for two additional factors that affect the learnability of words. One is semantic
complexity and, in particular, a hierarchy of semantic complexity where words
with multiple meanings are acquired later than those with fewer meanings.
Kwon’s results support the findings of the present study where the preposition
*por* appears to have more semantic complexity and learners under both CLBI
and TBI acquired it at a slower rate. differential rates of acquisition from one
learner to the next in the case of each preposition.

This section is concluded with an examination of the immediate and
delayed post-test results for both those under CLBI and TBI through the filter of
the second null hypothesis (H₀) (i.e., *Individual differences in ab initio learner
WMC do not differentially affect the acquisition of short and long-term productive
knowledge of L2 Spanish polysemous spatial prepositions*). The observed
results of the present study suggest that the first H₀ should be rejected, and that
an alternative hypothesis should be accepted that reads *individual differences in
ab initio learner WMC do differentially affect the acquisition of short and long-
term productive knowledge of L2 Spanish polysemous spatial prepositions*.

Overall, these results fit well into our current understanding of the
importance of IDs in WMC in L2 acquisition. Van Patten and Benati (2010) state
that all theories related to WMC agree that it has a limited capacity. The acquisition of polysemes requires learners to maintain a fairly large amount of discrete pieces of information in working memory. Because there are individual differences between learners in terms of their capacity to maintain information in working memory, there are also going to be individual differences in their rates of acquisition of these complex lexical items.

5.2.2 Discussion of the Effects of WMC on Long-term Productive Knowledge

This section discusses the results of a delayed post-test conducted to compare the acquisition of long-term productive knowledge of four Spanish polysemous spatial preposition under two instructional treatments and two levels of working memory capacity. The delayed post-test results under CLBI yielded a significant difference in scores for CB2 ($M = 31.87$, $SD = 5.67$) and TB1 ($M = 26.60$, $SD = 6.70$) conditions; ($t(28) = 2.3254$, $p = 0.028$). The effect size of the mean difference is medium ($d = 0.85$) and of the correlation coefficient is medium ($r = 0.39$). The delayed post-test results under TBI yielded a significant difference in scores for TB2 ($M = 29.80$, $SD = 4.77$) and TB1 ($M = 22.53$, $SD = 4.37$) conditions; ($t(28) = 4.3524$, $p < 0.001$). The effect size of the mean difference is large ($d = 1.59$) and of the correlation coefficient is large ($r = 0.62$). Since WMC is the only independent variable that distinguishes the two experimental groups being compared under each of the instructional treatments, the high WMC experimental groups’ significantly better performance on the delayed post-test would seem to only be accounted for due to some cognitive
advantage that their high WMC gives them in maintaining long-term productive knowledge of the four polysemous Spanish spatial prepositions. However, since WMC is understood to be the temporary storage of information, it would seem that the acquisition of long-term productive knowledge would not be affected by the beneficial effects of WMC since it is not linked with long-term storage of information. Robinson (1995) argues that individual differences in working memory and attentional capacity both affect a learner’s ability to notice new L2 input which directly affects SLA. Robinson defined noticing as the “detection with awareness and rehearsal in short-term memory necessary for learning and the subsequent encoding in long-term memory.” In other words, each participant in the present study could only encode into long-term memory what they were able to detect and rehearse in short-term memory. Robinson’s argument best explains the higher scores by the high WMC participants, as it is necessary for a learner to notice and gives attention to input but all of this is impacted by their working memory. The fact that all experimental groups experienced a decline in scores between the immediate and delayed post-tests that was relatively similar between the pairs of experimental groups being compared under each instructional treatment meant that the differential in scores observed in the immediate post-test would continue to be present in the delayed post-test. Specifically, for the two experimental groups being compared under TBI, the TB2 experimental group declined 14.0% while their TB1 counterparts declined 11.5%. For the two experimental groups being compared under CLBI, the CB2 experimental group declined 2.6% between the immediate and the delayed post-
test, while their CB1 counterparts declined 5.9%. The results of the delayed post-test show that the two experimental groups taught under CLBI maintained roughly 95% of what they initially acquired from their instructional treatment, while the two TBI experimental groups maintained roughly 87% of what they acquired from their initial instructional treatment. It comes down to each pair of experimental groups having a similar rate of decline in their productive knowledge of the four target prepositions between the immediate and the delayed post-test. In terms of the individual target prepositions, three out of four of the experimental groups (CB2, TB2 and TB1) all experienced their lowest scores with the preposition *por*. This suggests that *por* is the most difficult of the four target prepositions for learners to maintain in long-term productive knowledge. Arguably, this is due to the complexity of its lexical network of meanings, and the difficulty for learners to maintain memory of all of the uses of *por* in long-term storage.

In summary, the results of the delayed post-test suggest that higher scores in long-term productive knowledge do not directly correlate with higher scores in WMC, but rather, it is simply a reflection of the higher level of acquired productive knowledge observed in the immediate post-test results. Arguably, the only effect that WMC has on long-term productive knowledge is in aiding the learner to acquire more short-term productive knowledge which then allows for the possibility of a greater level of long-term productive knowledge. The clearest evidence that WMC does not beneficially effect long-term productive knowledge is the fact that the differential in the combined mean scores between the two high
WMC experimental groups and the two low WMC experimental groups did not widen, and in fact, it actually narrowed as the combined mean scores of the two high WMC experimental groups actually declined faster than their two low WMC counterparts.

5.3 Discussion of the Effects Resulting from the Interaction Between WMC and Instructional Treatment on the Acquisition of Productive Knowledge

The discussion of the effects caused by the interaction between individual differences in working memory capacity and instructional treatments is focused primarily on the results of the immediate post-test as this is where the main effects of working memory are more directly evident. The focus of this section is on answering the third research question (i.e., Do individual differences in ab initio learner WMC interact with the instructional treatments CLBI and TBI to differentially affect both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions?). In summary, the preponderance of the descriptive data from the post-tests suggest that the answer to this third research question is yes, however; the results of a MANOVA (Table 4.19) suggest that there is not a statistically significant difference in immediate and delayed post-test scores based on the interaction between WMC and instructional treatment, \( F (12, 3) = 1.90, p = .328; \) Wilk’s \( \Lambda = 0.116 \), partial \( \eta^2 = .88 \) with an observed power of .20. There were two key findings from the MANOVA. First, the results revealed a low level of observed power (.20). This means that the present study has only a 20% probability of finding a statistically significant difference in mean scores that is due to the interaction between WMC
and instructional treatment, if there is one to be found. It is generally accepted
that observed power should be .80 or greater (see Cohen, 1988); that is, you
should have an 80% or greater chance of finding a statistically significant
difference if there is one to be found. The low level of power of the present study
may have led to a Type II error which means that it failed to reject a null
hypothesis that was actually false. There are a number of factors that could have
affected the observed power, such as sample size and the magnitude of the
effect of the variable. Arguably, the observed power of the present study was
lowered by the relatively small sample size of 15 participants per experimental
group. The second key finding is the score of .88 under Partial Eta Squared
(Partial $\eta^2$) which reveals the proportion of variance in mean scores accounted
for by the main effect of the interactions between WMC and instructional
treatment. In other words, 88% of the variance in mean scores from the
immediate and delayed post-tests is due to interactions between learner WMC in
the two instructional techniques. This strongly suggests that the primary cause
of variance between learners is the result of learner working memory interacting
with the instructional treatment they received. Statistical significance was
detected when the effect size of the mean difference was large, such as in the
case of the preposition *por* ($d = 1.15$), ($r = 0.50$) in the results of the delayed
post-test (Table 4.31) comparing CB1 and TB1, then statistical significance was
reached ($p = 0.004$). Since the observed power was low, more attention was
placed in this discussion on the effect size and descriptive results rather than the
significance level found with the MANOVA.
The results of the immediate post-test comparing CB2 with TB2 (Table 4.20) and the results of the immediate post-test comparing CB1 with TB1 (Table 4.23) differ greatly. The overall results of these two comparisons suggest that there is an interactional effect between WMC and instructional treatment whereby learners with high WMC are beneficially affected in the acquisition of short-term productive knowledge under TBI, but low WMC learners are beneficially affected in the acquisition of short-term productive knowledge under CLBI.

The immediate post-test results for the two high WMC experimental groups (Table 4.20) did not yield a significant difference in scores for CB2 (\(M = 32.73, SD = 4.81\)) and TB2 (\(M = 34.67, SD = 3.29\)) conditions; \((t(28) = 1.2893, p = 0.133)\). However, there was a small effect size of the mean difference \((d = 0.47)\) and of the correlation coefficient is \((r = 0.23)\). Although this effect size is small, it does show a quantifiably higher mean score for TB2 over CB2 in the acquisition of short-term productive knowledge. The TB2 experimental group had a higher mean score of 34.67 versus a mean score of 32.73 for the CB2 experimental group. This works out to a difference in raw mean scores of 1.94 points or a 5.9% higher score for the TB2 experimental group versus the CB2 experimental group. Since the only difference between these two experimental groups was the instructional treatment that each received, the results suggest that learners with high WMC will be beneficially affected by TBI over CLBI in the acquisition of short-term productive knowledge of the four target prepositions.
The comparison of the immediate post-test results of the two low WMC experimental groups (Table 4.23) differs with the comparison of the results of the two high WMC experimental groups. An independent samples $t$-test was conducted to compare mean differences between CB1 and TB1 (Table 4.23), and the immediate post-test results did not yield a significant difference in scores for CB1 ($M = 28.27, SD = 6.03$) and TB1 ($M = 25.47, SD = 3.63$) conditions; ($t(28) = 1.5408, p = 0.135$). However, the effect size of the mean difference is approaching medium ($d = 0.56$) and of the correlation coefficient is small ($r = 0.27$). The effect size resulting from a comparison of the immediate post-test mean scores of the two low WMC experimental groups is slightly greater than those of the two high WMC experimental groups. The CB1 experimental group had a higher mean score of 28.27 versus a mean score of 25.47 for participants in the TB1 experimental group (Table 4.23). This works out to a difference of 2.80 points or an 11.0% higher score for the CB1 experimental group versus the TB1 experimental group. Since the only difference between these two experimental groups was the instructional treatment that each received, the results suggest that learners with low WMC will be beneficially affected by CLBI over TBI in the acquisition of short-term productive knowledge of the four target prepositions.

There are no studies with which to directly compare these immediate post-test results, since no previous study has specifically examined the effects resulting from the interaction between WMC and instructional treatment in the acquisition of short-term productive knowledge of the four target prepositions.
However, there are numerous empirical studies demonstrating strong correlations between individual differences and overall achievement in L2 acquisition. For example, Wesche (1981) found strong connections between language learner aptitude profiles and instructional treatments. Wesche reported that not only did learners perform better when matched with methods that aligned with their aptitude profile, but they also reported greater satisfaction with instruction. Levine and Reves (1990) researched the extent to which differences in vocabulary retention were related to different methods of vocabulary presentation. They also explored how different methods of vocabulary presentation interact with different learner factors such as personality, L1 background, word-processing habits and language attitudes. The findings from this study showed that the method of presenting new vocabulary leads to varying degrees of vocabulary retention. They argue that the retention of vocabulary seems to be related to the learner’s general learning patterns and/or cognitive styles of visual, auditory and contextual associations. Robinson (2002) argues that there exists a set of cognitive abilities, or aptitude complexes, that relate differently to language learning under different psycholinguistic processing conditions. Robinson describes these conditions as the situational level of classroom instruction and the specific pedagogic tasks that learners perform in classrooms. Robinson argues that purposely matching learners’ strengths in particular aptitude complexes with specific learning conditions and instructional techniques is an important element in the delivery of optimally effective classroom exposure and practice for L2 learners. Canner (2013) revealed
correlations between language learner aptitude and levels of proficiency in L2 Russian attained in naturalistic versus formal learning contexts.

There is also growing evidence that individual learners will interact differently with various techniques employed by second language teachers (Lam, 2009; Touplikioti, 2007). Robert Sternberg (2002, p. 13) stated in relation to three foreign language learning experiences that “I was being taught in different ways and responding differently to each of these ways. My aptitude was not internal to me, but in the interaction between my abilities and the way I was being taught.” Perhaps the most salient point that comes out of the immediate post-test results of the present study is the fact that no one instructional treatment can be claimed as best for ab initio learners in the acquisition of short-term productive knowledge of the four polysemous target prepositions. Rather, the results of the present study suggest that the best instructional treatment depends on the working memory capacity of the learner. The fact that the TB2 experimental group outscored the CB2 experimental group suggests that high WMC learners can not only handle the cognitive demands placed on working memory by TBI, they can even thrive under it. Cognitive linguistics based instruction may hold advantages in the teaching of words with complex networks of meanings (see section 5.1 of the present study), however, those advantages do not appear to outweigh the benefits of high working memory capacity, at least for ab initio learners.

In examining the results of the immediate post-test for the low WMC learners, it seems that the exact opposite is true. The results suggest that low
WMC learners are not able to handle the cognitive demands placed on working memory by TBI (as seen in TB1 experimental group’s results, Table 4.23), and are therefore much more dependent on a cognitive linguistics based approach that presumably reduces the demands on working memory by requiring learners to maintain fewer. Therefore, it can be argued that each *ab initio* learner will acquire a greater or lesser level of productive knowledge of polysemous spatial prepositions depending on their individual working memory capacity and the instructional treatment that they receive. In other words, learner WMC and instructional treatments interact to affect the acquisition of productive knowledge of the target prepositions. Past studies have attempted to show how learner IDs interact with instructional treatment. Nation (2001) argues that the best way to explain the meaning of a polyseme is to define the word by looking for the concept that runs through all its senses or uses, thereby reducing the learning burden. Levine and Reves (1990) researched the extent to which differences in vocabulary retention were related to different methods of vocabulary presentation. The methods of vocabulary presentation used in their study were (1) written presentation of L2 word and its L1 translation, (2) written presentation of L2 word in sentential context (3) L2 word with a picture, (4) written presentation of L2 word with its meaning, (5) auditory presentation of L2 word and its L1 translation, (6) auditory presentation of L2 word in sentential context and (7) three-fold computer presentation (word and its definition, word presented in analogy, word in context). The findings from this study showed that various learner factors or IDs combined differently with various methods of vocabulary presentation.
presentation. As a result, they argue that the processing of learning new vocabulary is a multifarious challenge; therefore, no single method should be imposed on learners.

The immediate post-test results (Table 4.21) of the separate mean scores of the individual target prepositions of the CB2 and TB2 experimental groups also serve to support the idea that IDs in WMC interact with CLBI and TBI. When running an independent samples t-test on the individual mean score differences for each preposition separately, only the preposition *en* produces a statistically significant score for CB2 *en* \((M = 9.13, SD = 1.36)\) and TB2 *en* \((M = 10.27, SD = 1.67)\) conditions; \((t(28) = -2.042, p = 0.051)\). The effect size of the mean difference for the preposition *en* (Table 4.22) is medium \((d = 0.75)\) and of the correlation coefficient is medium \((r = 0.35)\). The preposition *en* consistently received the highest scores out of the four target prepositions, and therefore suggests that it has the least complex network of meanings. It would also therefore place the least strain on learner working memory capacity. Therefore, there was no practical advantage for the CB2 experimental group in receiving a cognitive linguistics based instructional treatment of this preposition. In fact, the TB2 experimental group outscored the CB2 experimental group in every target preposition except for *por* which has already been argued to have the most complex network of meanings as shown in the consistently lower scores in receives. This suggests that for high WMC learners, CLBI is only beneficial when the specific target structure is sufficiently complex as to overwhelm the cognitive advantages that their high WMC gives them.
In comparison, the CB1 experimental group outscored their TB1 counterparts in each one of the target prepositions (Table 4.24). This suggests that low WMC learners will interact negatively with TBI even when their WMC is under less demand, such as with the preposition *en*. When running an independent samples *t*-test on the individual mean score differences for the preposition *por*, it was statistically significant for CB1 *por* (*M* = 6.33, *SD* = 1.45) and TB1 *por* (*M* = 4.80, *SD* = 1.61) conditions; (*t*(28) = 2.741, *p* = 0.011). The effect size of the mean difference for the preposition *por* (Table 4.24) is large (*d* = 1.00) and of the correlation coefficient is medium (*r* = 0.45). This result suggests that the negative interaction between TBI and low WMC will increase as the complexity of the target structure increases.

Since this section is focused on the interaction between learner WMC and instructional treatment, most of the focus has been on the immediate post-test as this is where the effects of working memory can be directly observed. However, there are some notable results from the delayed post-test. First, the delayed post-test results of CB2 and TB2 are now the opposite of the immediate post-test for these same two experimental groups. An independent samples *t*-test was conducted to compare mean differences between CB2 and TB2 (Table 4.26), and the delayed post-test results did not yield a significant difference in scores for CB2 (*M* = 31.87, *SD* = 5.67) and TB2 (*M* = 29.80, *SD* = 4.77) conditions; (*t*(28) = 1.0820, *p* = 0.289). The effect size of the mean difference is small (*d* = 0.40) and of the correlation coefficient is small (*r* = 0.19). When running an independent samples *t*-test on the individual mean score differences for each
preposition separately, none of the mean differences is statistically significant. The effect size of the mean difference for the preposition *por* (Table 4.27) is medium \((d = 0.70)\) and of the correlation coefficient is medium \((r = 0.33)\).

However, the descriptive results show that participants in the CB2 outscored their TB2 counterparts by 2.07 points or 6.9%. This suggests that the beneficial interaction of high working memory with TBI is primarily in terms of short-term productive knowledge. This is probably due to the fact that working memory is the “temporary storage and manipulation of information that is assumed to be necessary for a wide range of complex cognitive activities” (Baddeley, 2003, p. 189). In regards to SLA, working memory is what language learners use to briefly store and process new linguistic input in order to analyze it for comprehension.

Perhaps the most notable result of the delayed post-test is the fact that the low working memory group under CLBI (CB1) outscored the high working memory group under TBI (TB2) in the delayed post-test results for the preposition *por*. An independent samples *t*-test compared mean differences between CB1 and TB2, and the delayed post-test results did not yield a significant difference in scores for CB1 \((M = 5.93, SD = 1.53)\) and TB2 \((M = 5.07, SD = 2.46)\) conditions; \((t(28) = 1.1497, p = 0.260)\). The effect size of the mean difference for these post-test scores for the preposition *por* is small \((d = 0.42)\) and of the correlation coefficient is small \((r = 0.21)\). However, it does show a practical effect of an interaction between WMC, instructional treatment and time in regards to this specific target preposition. The interaction of these three main
effects can also be observed in the comparison of the delayed post-test results of CB1 versus TB1. An independent samples $t$-test showed that there is a tendency towards significance for the delayed post-test (Table 4.29) mean differences for CB1 ($M = 26.60, SD = 6.70$) and TB1 ($M = 22.53, SD = 4.37$) conditions; ($t(28) = 1.9706, p = 0.059$). Also, the effect size of the mean difference is medium ($d = 0.72$) and of the correlation coefficient is medium ($r = 0.34$). This suggest that the interaction between WMC, instructional treatment and time can produce a significant effect in the acquisition of productive knowledge of complex polysemous spatial prepositions.

In conclusion, when comparing the immediate post-test results of the two high WMC experimental groups, the TB2 experimental group outscored the CB2 experimental group, however, the exact opposite result occurred when comparing the mean scores of two low WMC experimental groups where the CB1 experimental group outscored the TB1 experimental group on the same immediate post-test. This suggests that the effects of instructional treatments in the acquisition of polysemous spatial prepositions will vary dependent on IDs in learner working memory capacity. Overall, the results of the present study suggest that the third null hypothesis should be rejected (i.e., Individual differences in ab initio learner WMC do not interact with the instructional treatments CLBI and TBI to differentially affect both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions). An alternative hypothesis would read, individual differences in ab initio learner WMC do interact with the instructional treatments CLBI and TBI to differentially affect
both short-term and long-term productive knowledge of L2 Spanish polysemous spatial prepositions. In regards to short-term productive knowledge, the present study suggests an even more specific hypothesis that reads, *ab initio learners with low WMC will acquire more short-term productive knowledge of L2 Spanish polysemous spatial prepositions under CLBI, however; ab initio learners with high WMC will acquire more short-term productive knowledge of L2 Spanish polysemous spatial prepositions under TBI.*
CHAPTER 6: CONCLUSION AND FUTURE RESEARCH

The results of the present study produce four general conclusions related to the main effects of instructional treatment, working memory capacity, and the interaction between working memory capacity and instructional treatment on the acquisition of both short-term and long-term productive knowledge of polysemous L2 Spanish spatial prepositions by *ab initio* L1 English speakers.

The first conclusion was drawn from the combined mean score of the four target prepositions from the immediate post-test. This general conclusion is *that for ab initio learners, CLBI and TBI do not differentially affect the acquisition of short-term productive knowledge of the target prepositions*. This key finding did not mirror the results from a number of previous studies that showed learners under CLBI outscored learners under TBI in the acquisition of polysemous words (cf. Khodadady & Khaghaninizhad, 2011; Lam, 2009; Makni, 2013; Morimoto & Loewen, 2007; Touplikioti, 2007).

The specific reason the results of the present study differ from the results of previous studies may be the present study’s use of learners with no prior knowledge of the target language, (i.e., Spanish) or any related cognate language. The use of *ab initio* learners was not a requirement of previous studies that explored the effectiveness of CLBI versus TBI, and in fact, these studies used participants who had prior knowledge of the target language and the target polysemes. Therefore, it could be argued that the conclusions of these
previous studies should have been qualified to say that learners under CLBI will acquire a greater level of knowledge of the target polysemes than learners under TBI, if the learners already have some prior translation-based knowledge of the target polysemes to build upon. If this is the case, than CLBI cannot be solely credited with the differential effects. Therefore, the first conclusion of the present study should also be qualified to say that CLBI and TBI are equally effective at developing short-term productive knowledge of the target prepositions, if (1) the learners have no prior knowledge of the target prepositions or target language, (2) each instructional treatment is the only source of information that the learners have available to negotiate meaning, and (3) the two experimental groups being compared are equally split between high WMC and low WMC learners. (This third qualification will be discussed in a later conclusion related to the interactions between learner WMC and instructional treatment.)

The second conclusion was drawn from the combined mean score of the four target prepositions from the delayed post-test. This general conclusion is that for ab initio learners, CLBI and TBI do have a moderately differential effect on the acquisition of long-term productive knowledge of the target prepositions. This key finding mirrors the results of Lam (2009) which also showed greater delayed post-test gains under CLBI.

The third conclusion is drawn from the results of the individual prepositional mean scores where the target preposition *por* revealed differential results from the other three target prepositions (*para, en*, and *a*). The results of the post-tests showed the preposition *por* yielded statistically significant
differences in mean scores on both immediate and delayed post-tests (Tables 4.5, 4.24, 4.30) with learners under CLBI higher scoring higher than learners under CLBI over those in TBI. These results produce the conclusion that one instructional treatment may actually be more effective in the teaching of a specific polysemous spatial preposition, even when other polysemous spatial prepositions in the same language do not experience differential rates of acquisition under these two instructional methods. This conclusion may seem to contradict the first conclusion, but the first conclusion was drawn from the results of a combined mean score of the four target prepositions. This conclusion was made by observing the separate post-test results of each preposition. This conclusion lines up with previous research that suggests each word may bear a differing learning burden for L2 learners (Kwon, 2005; Nation, 2001).

A fourth conclusion is that learners with high working memory capacity will acquire productive knowledge of polysemous spatial prepositions at faster rates than low working memory learners under both CLBI and TBI. This conclusion is not a novel discovery as there is a substantial body of research that demonstrates that adult L2 learners rely on certain types of cognitive resources, especially working memory, to attain high levels of proficiency in an L2. This conclusion is supported by previous studies that also investigated the role of working memory in the acquisition of productive knowledge and vocabulary learning (Mendonça, 2003; Mota, 2003)

The fifth conclusion is drawn from post-test results focused on the interaction between WMC and the two instructional treatments, CLBI and TBI. In
comparing the immediate post-test results of the two high WMC experimental groups, the high working memory learners under TBI outscored their high working memory counterparts under CLBI, however; the exact opposite result occurred when comparing the mean scores of the two low WMC experimental groups, where the low working memory learners under CLBI outscored the low working memory learners under TBI. This suggests that the effects of instructional treatments in the acquisition of productive knowledge of polysemous spatial prepositions will vary in accordance with the working memory capacity of the learner. The post-test results suggest the specific conclusion that *ab initio learners with low WMC will acquire more short-term productive knowledge of L2 Spanish polysemous spatial prepositions under CLBI, however; ab initio learners with high WMC will acquire more short-term productive knowledge of L2 Spanish polysemous spatial prepositions under TBI*. This conclusion fits well with a number of previous studies (Levine & Reves, 1990; Robinson, 2002; Wesche, 1981).

**Limitations and Recommendations**

There are two primary limitations, and four key recommendations that come out of the present study. The first limitation was the study design appeared to lack the observable power to detect statistical significance, especially as seen in the MANOVA results that examined the interaction between learner working memory and each of the two instructional treatments. Although there were 60 participants total, there were only 15 participants in each of the four experimental groups. This sample size may have been too small to reach statistical
significance in a number of the results where the effect sizes were ranging between small to medium. Therefore, it is recommended that any future research that explores the interaction between learner IDs and instructional treatment involve a sample size of 25 or more participants in each experimental group.

The second limitation was that learners were only given one lesson in their assigned instructional treatment. This may have had a differential effect on the results as learners under TBI would have already been familiar with the pedagogical techniques of repetition and memorization commonly used under TBI. Whereas learners under CLBI were being exposed to both novel L2 words and a novel way of learning words in a second language. Participants under CLBI varied in how well they were able to apply their understanding of the single core meaning of each preposition to all of its metaphorical extensions. As a result, it is recommended that any future research projects that attempt to compare CLBI to TBI should spend more time in giving lessons in each treatment to fully develop each learners understanding of not only the target structures, but more importantly, the pedagogical techniques that are to employed in learning them.

Each of the conclusions reached in the present study only apply to adult ab initio L1 English speakers attempting to acquire polysemous L2 Spanish spatial prepositions. Therefore, it is recommended that future studies should include learners from different L1 backgrounds, different ages, different L2 target polysemes as well as differing levels of prior knowledge and proficiency levels in
the target language. It is also recommended that when comparing the
effectiveness of two instructional treatments, researchers should clearly qualify
the level of prior knowledge and the source of prior knowledge of the target
structures that participants had prior to the start of the study. This will improve
the internal validity of these studies.
REFERENCES


http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf


Stern, H. H. (1975). What can we learn from the good language learner? 


West, M., & West, M. P. (Eds.). (1953). *A general service list of English words: with semantic frequencies and a supplementary word-list for the writing of popular science and technology.* Addison-Wesley Longman Limited.


APPENDIX A: STUDY PARTICIPANT CONSENT FORM

Please be informed that all information obtained during the course of this study will be known only to this researcher. All personal information will be kept completely confidential. Taking part in this research study is completely voluntary, and there is no penalty or other adverse effect if you decide to not take part. In addition, you are free to excuse yourself from further participation at any point during the course of this study without any adverse consequences.

If you wish to be considered as a candidate in this study, please sign and date your name below. Please also print your name and provide an email address to which follow up information may be sent.

Candidate Signature: ________________________   Date: ____/____/____

Candidate Name (Print): _____________________________

Candidate Email: ___________________________________

________________________________________ Date: ____/____/____

Joseph LeTexier (Researcher)

Participant ID Number: __________
APPENDIX B: BACKGROUND QUESTIONNAIRE

Please answers to each of the following questions:

1. Name: _________________________________   Age: ________

2. Is English your first language?   Yes   /   No
   If no, what is your first language? ________________

3. Where did you grow up? _________________________________

4. What languages do you speak? _________________________________
   __________________________________________________________________

5. What languages have you studied? _________________________________
   __________________________________________________________________

6. If you have studied Spanish, when and for how long did you study it? ______
   __________________________________________________________________

Signature: _________________________________       Date: ____/____/____

Researcher only
Participant ID Number: ________
APPENDIX C: SCHEMATIC REPRESENTATION OF THE VERB CATCH
(taken from Touplikioti, 2009, p.11)
APPENDIX D: CLBI INSTRUCTIONAL MATERIALS

Guidelines: On the following pages, your instructor will take you through a one-hour lesson on the meanings of the Spanish prepositions *para*, *en*, *por* and *a*. The lesson begins with a 30-minute explanation of the basic spatial relationship that each preposition may profile. This is followed by a series of sample sentences that exemplify the range of meanings for each preposition. These sentences are sub-divided into three separate domains of use – spatial (S), temporal (T) and metaphorical (M). Each sentence is accompanied by a picture that visually specifies the domain of meaning in focus. On each page you will also find a diagram that visually illustrates the central meaning of each preposition. These diagrams show the spatial relationship between a focus element (F) and a ground element (G). These diagrams and terms will be explained to you in the lesson. After the sample sentences, your instructor will guide you for ten minutes through a practice set of 12 sentences. These practice sentences are designed to both help teach you the meanings of the prepositions as well as prepare you for the quiz you will take right after the lesson. After you have completed the practice set, your instructor will go over the correct answer with you. After you have gone over the correct answers, you will have exactly 20 more minutes to review all the material before taking a 48 question fill-in-the-blank quiz. Use your answer sheet to circle the correct preposition for each sentence-picture presentation. Thanks for your participation!
The preposition *para* profiles a dynamic relationship in which the ground element (G) is viewed spatially as the destination or end goal of the focus element (F).

The chef prepared a *meal para* his *guests*.

The preposition *en* profiles a static relationship in which the ground element (G) is viewed spatially as place of containment or support for the focus element (F).

**The Spanish Preposition - En**

There is *bread en* the *basket*. 
The preposition *por* profiles a dynamic relationship in which the ground element (G) is viewed spatially as the path of movement of the focus element (F).

The F *taxis* drove *por* the G *streets of the city.*

The preposition *a* profiles a static or dynamic relationship in which the focus element (F) spatially relates to the ground element (G) as a relative point or line of direction on an axis.

F *She sat a* G *the side of the tiger.*
1. The chef prepared a meal para his guests. (S)

2. The pirate found gold para his treasure chest. (S)
3. The architectural plans will be ready para Monday. (T)

4. They must finish the construction project para the next month. (T)
5. The plumber used two wrenches **para** connect the pipes. (M)

6. The musician wrote a new song **para** his album.
The Spanish Preposition – en

1. There is bread en the basket. (S)

2. The plane landed en the water. (S)
3. The train will leave from the station en ten minutes. (T)

4. Valentine’s Day is en February. (T)
5. The marching band participated en the parade. (M)

6. He has a lot of problems en his life. (M)
The Spanish Preposition – Por

1. The taxis drove por the streets of the city. (S)

2. They crossed por the Brooklyn Bridge. (S)
3. They have been married *por* fifty years. (T)

4. The students have class *por* the morning. (T)
5. He cried **por** the death of his loved one. (M)

6. He traded the fish **por** a bottle of water. (M)
THE SPANISH PREPOSITION – A

1. The elephants walked a the watering hole. (S)

2. She sat a the side of the tiger. (S)
3. The wedding started at 1:30 p.m. (T)

4. The rooster crows at dawn. (T)
5. He has arrived a the most successful stage of his career. (M)

6. These flowers have reached a the end of their lives. (M)
PRACTICE SENTENCES

1. The mechanic used the wrench ___ fix the engine.

2. He saved her number ___ his mobile phone.

3. The baby slept ___ three hours.
4. They were praised ___ their musical talent.

5. They are ___ the end of a day of fishing.

6. He bought a toy ___ his dog.
7. The bird flew ___ its house.

8. There is lemonade ___ the glass.

9. She arrived ___ the conclusion that reading books is fun.
10. They walked ___ the cliffs.

11. He ran five miles ___ 28 minutes.

12. He prepared a work schedule ___ his employees.
Guidelines: On the following pages, your instructor will take you through a one-hour lesson on the meanings of the Spanish prepositions *para*, *en*, *por* and *a*. You will begin with a 30-minute lesson on the translated meanings of the Spanish prepositions *para*, *en*, *por* and *a*. Pages 2-3 cover the basic uses for each preposition, the common English translations as well as one sample sentence for each use. All sample sentences have been fully translated into English except for the preposition being taught. Pages 4 through 15 contain a series of sample sentences that exemplify the range of meanings and uses for each preposition. Each sentence is accompanied by a picture that visually specifies the meaning in focus. There is an English translation and a label of the use below each picture that explicitly indicate which meaning and use is in focus.

After the sample sentences, your instructor will guide you for ten minutes through a practice set of 12 sentences. These practice sentences are designed to both help teach you the meanings of the prepositions as well as prepare you for the quiz you will take right after the lesson. After you have completed the practice set, spend *exactly* 20 more minutes reviewing all the material by yourself. Immediately after this 20-minute study session, you will take a 48 question multiple-choice quiz. Use your answer sheet to circle the correct preposition for each sentence-picture presentation. Thanks for your participation!
Explanation of the basic uses and English translations for each preposition.

<table>
<thead>
<tr>
<th>SPANISH PREPOSITION - PARA</th>
<th>Uses of para</th>
<th>Common English Translation</th>
<th>Sample sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deadline or specific time in the future</td>
<td>by; for</td>
<td>The architectural plans will be ready <strong>para</strong> Monday.</td>
<td></td>
</tr>
<tr>
<td>Goal</td>
<td>in order to for</td>
<td>The plumber used two wrenches <strong>para</strong> connect the pipes.</td>
<td></td>
</tr>
<tr>
<td>Purpose / Reason</td>
<td>for; used for</td>
<td>The musician wrote a new song <strong>para</strong> his album.</td>
<td></td>
</tr>
<tr>
<td>Recipient / End Location</td>
<td>for</td>
<td>The chef prepared a meal <strong>para</strong> his guests.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>THE SPANISH PREPOSITION - EN</th>
<th>Uses of en</th>
<th>Common English Translation</th>
<th>Sample sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>on; onto; in; into</td>
<td>There is bread <strong>en</strong> the basket.</td>
<td></td>
</tr>
<tr>
<td>Time period</td>
<td>in; within</td>
<td>The train will leave from the station <strong>en</strong> ten minutes.</td>
<td></td>
</tr>
<tr>
<td>SPANISH PREPOSITION - POR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Uses of por</strong></td>
<td><strong>Common English Translation</strong></td>
<td><strong>Sample sentences</strong></td>
<td></td>
</tr>
<tr>
<td>Motion / General location</td>
<td>along; through; around; by; over</td>
<td>The taxis drove <em>por</em> the streets of the city.</td>
<td></td>
</tr>
<tr>
<td>Duration of an action</td>
<td>for; during; in</td>
<td>They have been married <em>por</em> fifty years.</td>
<td></td>
</tr>
<tr>
<td>Reason or motive for an action</td>
<td>because of; on account of; on behalf of</td>
<td>He cried <em>por</em> the death of his loved one.</td>
<td></td>
</tr>
<tr>
<td>Exchange or substitution</td>
<td>for; in exchange for</td>
<td>He traded the fish <em>por</em> a bottle of water.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SPANISH PREPOSITION - A</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Uses of a</strong></td>
</tr>
<tr>
<td>Direction</td>
</tr>
<tr>
<td>Specific Time</td>
</tr>
<tr>
<td>Relative Location</td>
</tr>
</tbody>
</table>
THE SPANISH PREPOSITION – PARA

The chef prepared a meal *para* his guests.

![Image of a chef serving a meal](image1.jpg)

RECIPIENT (FOR)

The pirate found gold *para* his treasure chest.

![Image of a pirate's treasure chest](image2.jpg)

END LOCATION (FOR)
**The Spanish Preposition – Para**

The architectural plans will be ready *para* Monday.

![Image](image1.png)

**Specific Future Time (by)**

They must finish the construction project *para* the next month.

![Image](image2.png)

**Deadline (by)**
**THE SPANISH PREPOSITION – PARA**

The plumber used two wrenches *para* connect the pipes.

*Goal (in order to)*

The musician wrote a new song *para* his album.

*Purpose/Reason (for)*
THE SPANISH PREPOSITION – EN

There is bread *en* the basket.

The plane landed *en* the water.

LOCATION (IN)

LOCATION (ON, ONTO)
THE SPANISH PREPOSITION – EN

The train will leave from the station en ten minutes.

TIME PERIOD (IN, WITHIN)

Valentine’s Day is en February.

TIME PERIOD (IN)
THE SPANISH PREPOSITION – EN

The marching band participated **en** the parade.

**LOCATION OR TIME PERIOD (IN)**

He has a lot of problems **en** his life.

**LOCATION (IN, WITHIN)**
The taxis drove *por* the streets of the city.

They crossed *por* the Brooklyn Bridge. (S)
THE SPANISH PREPOSITION – POR

They have been married *por* fifty years.

(Duration of Action (for))

The students have class *por* the morning.

(Duration of Action (in, during))
He cried *por* the death of his loved one.

**Reason for Action (because of, on account of)**

He traded the fish *por* a bottle of water.

**Exchange (for, in exchange for)**
The Spanish Preposition – a

The elephants walked a the watering hole.

Direction (to, towards)

She sat a the side of the tiger.

Relative Location (to, by, at, next to)
The wedding started a 1:30 p.m.

The rooster crows a dawn.
He has arrived at the most successful stage of his career.

These flowers have reached to the end of their lives.
**PRACTICE SENTENCES**

1. The mechanic used the wrench ___ fix the engine.

2. He saved her number ___ his mobile phone.

3. The baby slept ___ three hours.
4. They were praised ___ their musical talent.

5. They are ___ the end of a day of fishing.

6. He bought a toy ___ his dog.
7. The bird flew ___ its house.

8. There is lemonade ___ the glass.

9. She arrived ___ the conclusion that reading books is fun.
10. They walked ___ the cliffs.

11. He ran five miles ___ 28 minutes.

12. He prepared a work schedule ___ his employees.
APPENDIX F: ANSWERS TO PRACTICE SENTENCES

1. para
2. en
3. por
4. por
5. a
6. para
7. a
8. en
9. a
10. por
11. en
12. para
APPENDIX G: IMMEDIATE POST-TEST

**Quiz Instructions:** You are about to take a 48 question fill-in-the-blank quiz on the four Spanish prepositions *para, en, por* and *a*. During the quiz, sentences will appear on the screen one at a time for 15 seconds. These sentences have all been translated from Spanish into English, and each one contains a blank line where one of the four Spanish prepositions has been deleted. You must decide which preposition best completes the sentence according to the specified context. In order to visually specify the context, each sentence also has a corresponding picture. First read the sentence, and then examine the picture. Once you have decided which preposition you think best completes the sentence, simply circle this preposition on your answer sheet. Be sure the number on your answer sheet matches the number of each slide. Work quickly, but please do your very best to select the correct preposition for each sentence.

Please click your mouse when you are ready to begin.
Thanks for participating!

There are carrots ___ the soup.

[1] para en por a
He filled out the forms ___ the federal government.

The runner reached ___ the finish line.
Peter made the cake ___ Julia.

Immediate Post Test

para en por a

The leaf is ___ the tip of the branch.

Immediate Post Test

para en por a
The space shuttle blasted off ___ space.

The sand has been falling ___ eight minutes.
The owl hunts ___ the night.

There are a lot of words ___ these pages.
The ambulance came ___ his rescue.

Amy is nervous because her homework is ___ tomorrow.
I am going to travel ___ all of Europe this summer.

The wind blew leaves ___ all the ground.
The cook used a knife ___ cut the pepper.

She left her desk ___ 12:04 p.m.
The bird is ___ the cage.

The flight just arrived ___ the boarding gate.
She is praying ___ a sick friend.

He is studying ___ an exam.
She paid in cash ____ a glass of beer.

She has come ____ the point where she thinks leisure is best.
He is walking his dog ___ the woods.

He solved the Rubik’s cube ___ three minutes.
The monk spends many hours ___ meditation.

The boy moved ___ the front of the line.
They are protesting ____ animal rights.

These glasses are ____ wine.
28
He escaped ____ a tunnel.

29
The woman examined the painting ____ the wall.
She set the spoon ___ the side of the knife.

The eagle landed ___ the branch.
She stopped ___ the end of the path.

The pizza is ___ eating during the party.
He is ___ the doorway.

They work ___ a construction company.

34  para  en  por  a

35  para  en  por  a
The scientists found a cure ___ the disease.

The river flows ___ the valley.
The kids hid ___ the side of a tree.

He has a lot of credit cards ___ his pocket.
The sun sets ___ midnight.

He has a jelly stain ___ his tie.
Her maternity leave is scheduled ___ the next week.

The explorer climbed ___ the top of the mountain.
The soldier died fighting ___ his countrymen.

He used the lawn mower ____ cut the grass.
She has too much stress ___ her life.

She wants to exchange a small blouse ___ a large one.
They hope to save one million dollars ____ retirement.

Immediate Post Test

para en por a

You are done! Please be sure your name is on your answer sheet before returning it to the researcher.

You may now exit the quiz.
APPENDIX H: DELAYED POST-TEST

Quiz Instructions: You are about to take a 48 question fill-in-the-blank quiz on the four Spanish prepositions para, en, por and a. During the quiz, sentences will appear on the screen one at a time for 15 seconds. These sentences have all been translated from Spanish into English, and each one contains a blank line where one of the four Spanish prepositions has been deleted. You must decide which preposition best completes the sentence according to the specified context. In order to visually specify the context, each sentence also has a corresponding picture. First read the sentence, and then examine the picture. Once you have decided which preposition you think best completes the sentence, simply circle this preposition on your answer sheet. Be sure the number on your answer sheet matches the number of each slide. Work quickly, but please do your very best to select the correct preposition for each sentence.

Please click your mouse when you are ready to begin.
Thanks for participating!

There are peas ___ the soup.
He filled out the forms ___ the federal government.

The runners reached ___ the finish line.
She wrapped a gift ___ her boyfriend.

The orange is ___ the tip of the branch.
The space shuttle blasted off ___ space.

The sand has been falling ___ eight minutes.
The wolf hunts ___ the night.

There are a lot of words ___ these pages.
The helicopter came ___ his rescue.

Paul is nervous because his homework is ___ tomorrow.
I am going to travel ___ all of the United States this summer.

The wind blew leaves ___ all the ground.
The cook used a knife ___ cut the lemon.

The bus arrived ___ 12:03 p.m.
16
The bunny is ___ the cage.

17
The flight just arrived ___ the boarding gate.
They are praying ___ a sick friend.

She is studying ___ an exam.
He paid in cash ___ a pastry.

20

She has come ___ the point where she thinks leisure is best.

21
He is walking ___ the woods.

He solved the Rubik’s cube ___ three minutes.
The monk spends many hours ___ meditation.

The bicyclist went ___ the end of the line.
They are protesting ___ natural conservation.

These glasses are ___ the wine.
He escaped ___ a tunnel.

She examined the painting ___ the wall.
She set the fork ___ the side of the knife.

The eagle landed ___ the branch.
He stopped ___ the edge of the cliff.

The pizza is ___ eat during the party.
She is sitting ___ the doorway.

They work ___ a construction company.
The scientists found a cure ___ the disease.

The river flows ___ the valley.
The cat hid ___ the side of the tree.

She has her cellphone ___ her pocket.
The sun sets ___ midnight.

There is a coffee stain ___ his notebook.
Her maternity leave is scheduled ___ the next week.

The explorer climbed ___ the top of the mountain.
The soldier died fighting ___ his countrymen.

He used the lawn mower ___ cut the grass.
46
She has too much stress ___ her life.

47
She wants to exchange a small blouse ___ a large one.
They hope to save one million dollars ___ retirement.

You are done! Please be sure your name is on your answer sheet before returning it to the researcher.

You may now exit the quiz.
## APPENDIX I: POST-TESTS ANSWER SHEET

Name: ___________________________  Participant No. : ____________

<table>
<thead>
<tr>
<th>No</th>
<th>Preposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>por a para en</td>
</tr>
<tr>
<td>2</td>
<td>por a para en</td>
</tr>
<tr>
<td>3</td>
<td>por a para en</td>
</tr>
<tr>
<td>4</td>
<td>por a para en</td>
</tr>
<tr>
<td>5</td>
<td>por a para en</td>
</tr>
<tr>
<td>6</td>
<td>por a para en</td>
</tr>
<tr>
<td>7</td>
<td>por a para en</td>
</tr>
<tr>
<td>8</td>
<td>por a para en</td>
</tr>
<tr>
<td>9</td>
<td>por a para en</td>
</tr>
<tr>
<td>10</td>
<td>por a para en</td>
</tr>
<tr>
<td>11</td>
<td>por a para en</td>
</tr>
<tr>
<td>12</td>
<td>por a para en</td>
</tr>
<tr>
<td>13</td>
<td>por a para en</td>
</tr>
<tr>
<td>14</td>
<td>por a para en</td>
</tr>
<tr>
<td>15</td>
<td>por a para en</td>
</tr>
<tr>
<td>16</td>
<td>por a para en</td>
</tr>
<tr>
<td>17</td>
<td>por a para en</td>
</tr>
<tr>
<td>18</td>
<td>por a para en</td>
</tr>
<tr>
<td>19</td>
<td>por a para en</td>
</tr>
<tr>
<td>20</td>
<td>por a para en</td>
</tr>
<tr>
<td>21</td>
<td>por a para en</td>
</tr>
<tr>
<td>22</td>
<td>por a para en</td>
</tr>
<tr>
<td>23</td>
<td>por a para en</td>
</tr>
<tr>
<td>24</td>
<td>por a para en</td>
</tr>
<tr>
<td>25</td>
<td>por a para en</td>
</tr>
<tr>
<td>26</td>
<td>por a para en</td>
</tr>
<tr>
<td>27</td>
<td>por a para en</td>
</tr>
<tr>
<td>28</td>
<td>por a para en</td>
</tr>
<tr>
<td>29</td>
<td>por a para en</td>
</tr>
<tr>
<td>30</td>
<td>por a para en</td>
</tr>
<tr>
<td>31</td>
<td>por a para en</td>
</tr>
<tr>
<td>32</td>
<td>por a para en</td>
</tr>
<tr>
<td>33</td>
<td>por a para en</td>
</tr>
<tr>
<td>34</td>
<td>por a para en</td>
</tr>
<tr>
<td>35</td>
<td>por a para en</td>
</tr>
<tr>
<td>36</td>
<td>por a para en</td>
</tr>
<tr>
<td>37</td>
<td>por a para en</td>
</tr>
<tr>
<td>38</td>
<td>por a para en</td>
</tr>
<tr>
<td>39</td>
<td>por a para en</td>
</tr>
<tr>
<td>40</td>
<td>por a para en</td>
</tr>
<tr>
<td>41</td>
<td>por a para en</td>
</tr>
<tr>
<td>42</td>
<td>por a para en</td>
</tr>
<tr>
<td>43</td>
<td>por a para en</td>
</tr>
<tr>
<td>44</td>
<td>por a para en</td>
</tr>
<tr>
<td>45</td>
<td>por a para en</td>
</tr>
<tr>
<td>46</td>
<td>por a para en</td>
</tr>
<tr>
<td>47</td>
<td>por a para en</td>
</tr>
<tr>
<td>48</td>
<td>por a para en</td>
</tr>
</tbody>
</table>