1-1-2013

Working Memory-Learning Condition Interactions: Proficiency In L2 Russian Under Naturalistic and Formal Learning Conditions

Marc Canner
University of South Carolina

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

Recommended Citation

This Open Access Dissertation is brought to you for free and open access 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 SCHOLARC@mailbox.sc.edu.
WORKING MEMORY-LEARNING CONDITION INTERACTIONS
PROFICIENCY IN L2 RUSSIAN UNDER NATURALISTIC AND FORMAL LEARNING CONDITIONS

by

Marc T. Canner

Bachelor of Arts
University of New Mexico, 1984

Master of Arts
University of South Carolina, 2011

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
2013

Accepted by:
Paul Malovrh, Director of Committee
Nina Moreno, Committee Member
Curtis Ford, Committee Member
Alexandra Rowe, Reader
Lacy Ford, Vice Provost and Dean of Graduate Studies
Dedication

I dedicate this work to my loving wife, Dale, and to my incredible children, Aleksandre, Wesley and Rachel. Apart from their longsuffering patience, kindness and support, this project would never have been possible.
Acknowledgements

As with any challenging endeavor, this dissertation simply could not have been completed without the aid and friendship of a great number of people. I would like to first thank my wife, whose great patience and encouragement have been instrumental in the pace that I kept toward its completion. I would also like to thank Paul Malovrh, who was there every step of the way with the kind of focused advice, mentoring, and planning that enabled me to complete the task. My first advisor, Barbara Schulz, though absent for the completion of this work due to illness, also deserves my heartfelt thanks for helping me to get the research process off the ground. I also extend gratitude to Nina Moreno, Curt Ford and Alexandra Rowe for their outstanding recommendations as part of my committee. I wish to thank the many participants and friends in Russia and Ukraine whose cheerful willingness to endure testing was actually one of the most inspiring aspects of the project. To John Musgrove, missionary extraordinaire and faithful friend to me and to my Ukrainian brothers and sisters, I extend my warmest appreciation. The many long hours he spent helping me to contact potential participants in Kiev and Moscow speak volumes as to his selfless friendship. I also acknowledge the help of my friend and partner in training cross-cultural workers, Dave Wheeler. Apart from his hard work at running our institute in my absence, this project would not have been possible. Last of all, I again want to thank my incredible wife, Dale, who, in addition to her help on the home front, has diligently filled in for me as a teacher at our institute. The good Lord does not make language teachers better than her.
Abstract

The purpose of this dissertation was to examine the relationships between working memory (WM) and three commonly used learning strategies or conditions in the nature of proficiency among adult L2 Russian learners. Based on the aptitude-learning condition interaction framework articulated by Robinson (2002b), the study identifies two types of relationships between the fixed variables of working memory and learning condition, and four variables related to second language proficiency (SLP) in Russian: accuracy, fluency, overall general proficiency or ‘native-likeness’, and the occurrence of potentially fossilized forms in speech. Three main learning contexts were identified, including strictly naturalistic conditions resulting from immersion in-country (NC), naturalistic learning followed by formal learning experiences (NF), and formal training followed by naturalistic learning through immersion (FN).

37 participants were recruited, including naturalistic learners (n = 12), naturalistic/formal (n = 12), and formal/naturalistic participants (n=13). A proficiency interview was used to test learners for their L2 Russian proficiency, including measures for fluency, overall proficiency or native-likeness, accuracy via the elicitation of three prevalent structures in native speech, and the presence of potentially fossilized interlanguage (IL) forms. Based on proficiency scores and answers provided on an experience-related questionnaire, three experiments were conducted. The first experiment tests for predictive relationships between WM and the four aspects of proficiency for each learning condition. The second experiment investigates potential
interactions between WM and the three learning conditions in relation to the main aspects of proficiency. A third experiment tests for correlations between the amount of formal training and accuracy-related measures for the two combined (NF and FN) learning contexts.

The results of a logistic regression indicate significant predictive correlations between working memory and accuracy, and a negative correlation with the rate of potential fossilization for the naturalistic condition. No predictive significance for WM alone and aspects of proficiency was found in either NF or FN conditions. The second experiment yielded results indicating a significant interaction between WM and naturalistic conditions for fluency, and significance for interactions between WM and the amount of formal training for both accuracy and general proficiency for the formal/naturalistic (FN) context. The third experiment found predictive correlations between the amount of formal training and fossilization for the FN condition. The study found no significant correlations for the NF context. Results are interpreted to signify differential effects in relationships between WM and the major aspects of proficiency among different conditions, with an evident impact of the order or sequence of naturalistic and formal experience for the combined contexts. It appears that working memory does not act alone, but interacts with the nature of learning conditions to affect proficiency. Significant correlations between the amount of formal instruction and accuracy-related measures including fossilization are indicated.
# TABLE OF CONTENTS

DEDICATION ....................................................................................................................... iii

ACKNOWLEDGEMENTS ....................................................................................................... iv

ABSTRACT .............................................................................................................................. v

LIST OF TABLES ................................................................................................................... ix

LIST OF FIGURES ................................................................................................................ xi

LIST OF ABBREVIATIONS .................................................................................................... xii

CHAPTER I: INTRODUCTION .............................................................................................. 1

1.1 THE NATURE OF ADULT L2 VARIABILITY ................................................................. 1

1.2 THE RESEARCH PROBLEM ......................................................................................... 6

1.3 THE RELEVANCE OF APTITUDE AND LEARNING CONDITIONS ......................... 14

1.4 OVERVIEW OF THE DISSERTATION ..................................................................... 18

CHAPTER II: LITERATURE REVIEW ................................................................................ 20

2.1 VARIABILITY AND FOSSILIZATION IN ADULT SLA: COGNITIVIST PERSPECTIVES ... 20

2.2 THE NATURE OF THE APTITUDE COMPLEX OF ABILITIES ................................... 30

2.3 MODELS OF APTITUDE-ORIENTED INTERACTIONIST RESEARCH ..................... 54

2.4 WORKING MEMORY AND ITS ROLE IN THE ATTAINMENT OF L2 PROFICIENCY ..... 57

2.5 CHALLENGES RELATED TO THE ACQUISITION OF L2 RUSSIAN .................... 69

2.6 CHALLENGES IN ACQUIRING L2 RUSSIAN PROFICIENCY .............................. 81

2.7 RESEARCH QUESTIONS AND HYPOTHESES .......................................................... 85

CHAPTER III: RESEARCH METHODS ............................................................................. 88
LIST OF TABLES

TABLE 3.1 STUDY PARTICIPANTS ........................................................................... 94
TABLE 3.2 PARTICIPANT FLUENCY AS A PERCENTAGE OF NATIVE-LEVEL FLUENCY .... 108
TABLE 3.3 INDEPENDENT AND DEPENDENT VARIABLES .................................................. 114
TABLE 4.1 ALL PARTICIPANTS: AVERAGE PROFICIENCY SCORES ........................................... 122
TABLE 4.2 ALL PARTICIPANTS: ACCURACY, FLUENCY AND NATIVE-LIKENESS BY WM AND EXPERIENCE .............................................................. 126
TABLE 4.3 AVERAGE NATIVE-LIKENESS BY LEVEL OF WORKING MEMORY/EXPERIENCE ... 130
TABLE 4.4 AVERAGE FLUENCY BY LEVEL OF WORKING MEMORY/EXPERIENCE ............ 135
TABLE 4.5 AVERAGE ACCURACY BY LEVEL OF WORKING MEMORY/EXPERIENCE ........... 140
TABLE 4.6 ALL PARTICIPANTS: FR BY LEVEL OF WM AND EXPERIENCE ............................ 142
TABLE 4.7 AVERAGE FOSSILIZATION BY LEVEL OF WORKING MEMORY/EXPERIENCE ....... 146
TABLE 4.8 LEARNING CONDITIONS: GENERAL PROFICIENCY AVERAGES (NL) ............... 153
TABLE 4.9 LEARNING CONDITIONS: AVERAGE FLUENCY BY AMOUNT OF EXPERIENCE ... 155
TABLE 4.10 LEARNING CONDITIONS: AVERAGE ACCURACY ................................................ 156
TABLE 4.11 LEARNING CONDITIONS: AVERAGE RATE OF POTENTIAL FOSSILIZATION ... 158
TABLE 4.12 HIGHEST INDIVIDUAL PROFICIENCY COMPONENT SCORES .............................. 160
TABLE 4.13 NF PROFICIENCY AVERAGES: FORMAL TRAINING AS A PERCENTAGE .......... 165
TABLE 4.14 FN PROFICIENCY AVERAGES: FORMAL TRAINING AS A PERCENTAGE .......... 165
TABLE 4.15 NF LEARNERS: ACCURACY-RELATED SCORES BY FORMAL TRAINING (FT) ... 166
TABLE 4.16 FN LEARNERS: ACCURACY-RELATED SCORES BY FORMAL TRAINING (FT) ...166
TABLE 4.17 NF LEARNERS: FR WITH FORMAL TRAINING AS A PERCENT OF EXPERIENCE..168
TABLE 4.18 FN LEARNERS: FR WITH FORMAL TRAINING AS A PERCENT OF EXPERIENCE..168
LIST OF FIGURES

FIGURE 4.1 WMC BY ACCURACY FOR NATURALISTIC LEARNERS...............................141
FIGURE 4.2 WMC BY FR FOR NATURALISTIC LEARNERS........................................147
FIGURE 4.3 OBSERVED FOSSILIZATION BY FORMAL TRAINING FOR FN LEARNERS.....169
FIGURE 5.1 AVERAGE NATIVELIKENESS BY AMOUNT OF EXPERIENCE.........................203
FIGURE 5.2 AVERAGE FLUENCY BY AMOUNT OF EXPERIENCE.....................................205
FIGURE 5.3 AVERAGE ACCURACY BY AMOUNT OF EXPERIENCE.................................209
FIGURE 5.4 AVERAGE FOSSILIZATION RATE BY AMOUNT OF EXPERIENCE....................212
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANOVA</td>
<td>Analysis of Variance</td>
</tr>
<tr>
<td>ALA</td>
<td>Analytic Language Ability</td>
</tr>
<tr>
<td>AM</td>
<td>Associative Memory</td>
</tr>
<tr>
<td>CE</td>
<td>Control Executive (a function of Working Memory)</td>
</tr>
<tr>
<td>CLT</td>
<td>Communicative Language Teaching</td>
</tr>
<tr>
<td>CPH</td>
<td>Critical Period Hypothesis</td>
</tr>
<tr>
<td>DLAB</td>
<td>Defense Language Aptitude Battery</td>
</tr>
<tr>
<td>EB</td>
<td>Episodic Buffer (a function of Working Memory)</td>
</tr>
<tr>
<td>FDH</td>
<td>Fundamental Difference Hypothesis</td>
</tr>
<tr>
<td>FL</td>
<td>Foreign Language</td>
</tr>
<tr>
<td>FN</td>
<td>Formal/Naturalistic learning condition</td>
</tr>
<tr>
<td>FSH</td>
<td>Fundamental Similarity Hypothesis</td>
</tr>
<tr>
<td>g</td>
<td>General Intelligence</td>
</tr>
<tr>
<td>Gf</td>
<td>Fluid Intelligence</td>
</tr>
<tr>
<td>GS</td>
<td>Grammatical Sensitivity</td>
</tr>
<tr>
<td>ID</td>
<td>Individual Difference</td>
</tr>
<tr>
<td>L1</td>
<td>First or Native language</td>
</tr>
<tr>
<td>L2</td>
<td>Second language</td>
</tr>
<tr>
<td>MLAT</td>
<td>Modern Language Aptitude Test</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MMR</td>
<td>Moderated Multiple Regression</td>
</tr>
<tr>
<td>MRA</td>
<td>Multivariate Regression Analysis</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>NC</td>
<td>Naturalistic learning Condition</td>
</tr>
<tr>
<td>NL</td>
<td>Native Language</td>
</tr>
<tr>
<td>NF</td>
<td>Naturalistic/Formal learning condition</td>
</tr>
<tr>
<td>PL</td>
<td>Phonological Loop (a function of working memory)</td>
</tr>
<tr>
<td>PLAB</td>
<td>Pimsleur Language Aptitude Battery</td>
</tr>
<tr>
<td>SLA</td>
<td>Second Language Acquisition</td>
</tr>
<tr>
<td>SLP</td>
<td>Second Language Proficiency</td>
</tr>
<tr>
<td>STM</td>
<td>Short Term Memory</td>
</tr>
<tr>
<td>UG</td>
<td>Universal Grammar</td>
</tr>
<tr>
<td>VS</td>
<td>Visuospatial Sketchpad (a function of working memory)</td>
</tr>
<tr>
<td>WM</td>
<td>Working Memory</td>
</tr>
<tr>
<td>WMC</td>
<td>Working Memory Capacity</td>
</tr>
</tbody>
</table>
CHAPTER I: INTRODUCTION

One of the perennial problems in SLA research is the existence of a high degree of inter-learner variability in terms of both overall success and the nature of an adult’s developing or mature inter language (IL) grammar. One of the most prevalent characteristics of L2 learners is their general failure to acquire native-like proficiency, especially in terms of accuracy or complexity. Foreign language teachers are usually able to identify such differences among learners as they develop, but are rarely able to prevent such failure: the goal of promoting native-like proficiency can be daunting for the learner and teacher alike. The adult L2 acquisition of Russian is certainly no exception to this observation. Regardless of the type of strategy utilized, the amount of time spent learning in a Russian-speaking country, or the number of teachers, tutors or conversation partners employed, adult English L1 learners of L2 Russian rarely approach native levels of proficiency, and differ widely in their attained language abilities.

1.1 The Characteristics of Adult L2 Variability

Of particular interest to the present study are three key aspects of adult L2 acquisition related to variability in proficiency: an overall lack of success, general failure, and variation in the degree of success (Bley-Vroman, 1989). The first of these, an overall lack of success, involves the failure to master complex skills in adult learning for those areas that do not have a related domain-specific cognitive capacity. The degree to which such skills are mastered varies among adults, and few ever approach target-like proficiency. Such an absence of adult success presents a great obstacle to any theory that
would state that a single process is responsible for both adult and child language learning (see for example, Bley-Vroman, 1989; Dörnyei, 2005; Robinson, 2002b). A second aspect is general failure or the rarity of complete success, which means that “a (second) language is not merely difficult to learn with only general cognitive strategies, it is virtually impossible” (Bley-Vroman, 1989, p. 44). The extent of such failure is also variable. Lastly, variation in the degree of success relates to the common observation that different degrees of success in attainment are experienced by adult learners who utilize the same strategies, experience similar learning conditions or are influenced by similar affective variables, such as motivation (see for example, Dornyei, 2005; Dörnyei and Skehan, 2003). These three main aspects of variation in adult attainment of L2 proficiency point out the importance to any SLA theory of explaining this phenomenon.

The above three major types of variation in proficiency can be characterized as inter-learner variation, since they describe differences between learners. Another major type of variability also exists, however: intra learner variability. Intra learner variability refers to differences within the proficiency acquired by individual learners, or variation in the success the learner achieves between the different major areas of proficiency. For example, the same learner may have achieved a high level of fluency (native-like rapidity of speech), but without a high level of morphological accuracy. Sentence structure (syntax) may be native-like, but without target-like inflectional morphology. Research in the area of intra learner variation typically relates to an examination of the causes of fossilization (Han, 2004).

**Fossilization in Adult L2 Attainment**

In terms of proficiency or the degree to which a learner’s speech is native-like, the
phenomenon of fossilization stands out as one of the most intriguing aspects of SLA research, and its explanation remains elusive at best. Although the phenomenon has been defined in a number of ways, one of the more succinct definitions is provided by Han (2004), who describes fossilization as the “stabilized IL [inter language] forms that remain in learner speech or writing over time, no matter what the input or what the learner does” (p. 20). The literature also describes fossilization as having two dimensions: Cognitively, it relates to the primary systems involved in the generation of fossilized structures; experimentally, it refers to fossilized IL structures that are retained in communication (whether spoken or written) regardless of the input or of any strategies the learner uses to eliminate them (Han, 2004). The parts of speech that tend to fossilize include linguistic structures, various subsystems, and rules that a learner has adopted into the IL associated with a specific target language (TL) grammar, regardless of the adult learner’s age, amount of instruction in the L2, or extent of exposure to TL forms; research shows that such fossilized forms appear predisposed to recur in production even if they had been previously ‘eliminated’ (e.g. Selinker, 1972).

The evidence for fossilization varies widely, with much debate over the exact nature of the phenomenon. A number of potential characteristics of the issue have been discussed in the literature. One question is whether fossilization should be viewed as being a global phenomenon, or something that is expressed more locally. The global view of fossilization states that it impacts the entire IL system (e.g. Selinker, 1992), while those who view it as a local phenomenon have concluded that its effects are only seen on specific aspects of the interlanguage grammar (e.g. VanPatten, 1988; Han, 2004). Han (2004) points out that the global view of fossilization relies strictly on mere assumptions.
and is not supported in the evidence; on the other hand, the conclusion that it is largely 
local is predicated upon the existing evidence that demonstrates that fossilization only 
affects specific linguistic properties that are part of various sub-components of the IL 
graham, while other parts of the system tend to remain unaffected (Han, 2004).

One general and fairly common view of fossilization describes two basic types of 
learners: those who have fossilized, who are called ‘type 1’ learners, and learners who 
have not, called ‘type 2’. Han (2004) states that “such a bifurcation of L2 learners is 
conceptually flawed” (p. 21) since it is dependent on conclusions related to whether 
learning is continuing to occur, which are based in turn solely upon the learner’s 
observed performance; according to Han (2004), such a classification therefore represents 
a largely behaviorist perspective. Researchers also hold to competing perspectives on 
whether fossilization is a product or a process, with some referring to it as both. The 
cognitivist perspective holds to the view that it is largely a product, while those who see 
fossilization as more of a process have adopted a predominantly phenomenological 
perspective of the issue (Han, 2004). Jung (2002) points out the challenges associated 
with drawing conclusions about the exact nature of fossilization, such as the need to 
conduct long-term research to track its development. Such difficulties underscore the 
problematic issues involved in discovering its causes. However, a number of key 
observations have been made that help to clarify our understanding of the phenomenon 
and consequently the factors that contribute to its occurrence.

Findings related to fossilization are typically found either in longitudinal research 
that focuses on individual learners and their L2 development over time, or in studies that 
examine the L2 grammars of learners who have spent many years in an L2 context, have
experienced continued exposure to the target language, and have acculturated or adjusted well (e.g. Long, 2003). Different degrees of observed fossilization are usually described in two main respects: inter-learner fossilization (some learners come closer to target-like (TL) structures than others), and intra-learner fossilization (the individual comes closer to the target language in some respects than in others) (Han, 2004). Researchers have explained the different degrees of fossilization and consequently success in the attainment of proficiency in various ways. The number of possible explanations that have been suggested is staggering, and at least 50 different potential explanations of its causes have been identified (Han, 2004). Within cognitivist research, the explanation of fossilization emphasizes both cognitive and external factors that contribute to the phenomenon (e.g. DeKeyser, 2000; Schmidt, 1983; Seliger, 1975; Skehan, 1998). The present study relies upon this perspective.

Past approaches to the issue of learner variability and fossilization have been largely one-dimensional, and the emphasis of research has wavered between individual difference (ID) variables in cognitive, affective, or learning style domains, the type of learning context or environment, or among generativist researchers a learner’s access to UG. A focus on learning context as the main cause of such variation motivated the move toward more communicative or implicit styles of teaching (e.g. Krashen, 1981), and other subsequent trends, such as the more recent reaction to the failures of communicative teaching that has revived an emphasis on the importance of focus-on-form instruction (e.g. Long and Robinson, 1998). In recent years, a more multi-dimensional approach to the problem of inter-learner and intra-learner variability in SLA research has surfaced. According to Robinson (2001, 2002b), the key to understanding this long-standing
problem perhaps lies not in one or another of the many variables, but in an interaction of at least two key aspects of L2 acquisition: IDs in cognitive abilities (i.e. components of ‘aptitude’) and specific learning conditions, a model that has been called ‘aptitude-learning condition interaction research’ (Robinson, 2002b). For some researchers the concept of aptitude is now encompassed by Baddeley’s (2003) construct of working memory (e.g. Bowden, Sanz, and Stafford, 2005; Miyake and Friedman, 1998).

1.2 The Research Problem

It is the aforementioned research model that provides the basis for this study of adult L2 Russian. Based on Robinson’s (2002b) framework, the goal of the present study was to examine specific characteristics of adult second language proficiency (SLP) in L2 Russian which result from an interaction of working memory and the types of learning conditions and strategies that are most commonly experienced by English-speaking learners of Russian. Four main aspects of adult proficiency in L2 Russian were targeted for this study: accuracy, fluency, general proficiency (excluding pronunciation), and the presence of potentially fossilized forms in speech. Working memory was chosen because of its recognition as one of the primary components of adult language aptitude used in the acquisition of a second language (e.g. Miyake and Friedman, 1998). As a result of scores demonstrated during a WM reading span task, three profiles of working memory capacity were identified among study participants, including Low, Average, and High abilities.

To examine participant proficiency, an oral proficiency test was conducted including a general oral interview, a fluency monologue, and an accuracy elicitation task that targets three morpho-syntactic structures that are ubiquitous in everyday native Russian speech. Potential fossilization was measured by recording consistently repeated errors in
morphosyntax during the interview process. These measures of proficiency were tested among post-pubescent learners who have used three main types of acquisition strategies.

37 adult Russian speakers were recruited for the experiment, including 35 who were residing in Russia or the Russian-speaking region of Ukraine at the time of testing. The three learning conditions that were targeted for this research represent the strategies that are most commonly utilized by adult L2 Russian learners who have relocated for work to Russian-speaking regions of the former Soviet Union. These learning conditions are based on two predominant types of learning experienced by adult English L1 learners of Russian as a second language. The first major type of learning is naturalistic: many adult learners have attempted to acquire the language via largely naturalistic strategies. Such strategies include certain ‘barefoot’ approaches that focus on incidental or intentional interaction, or the use of adult conversation partners or conversation ‘tutors’. Others have approached the problem entirely differently, and have utilized various formal approaches including grammar instructors, college classroom training, or for many, intensive programs prior to entry into a Russian-speaking country.

This second general strategy has included both formal training and naturalistic experience resulting from immersion. When both naturalistic and formal conditions are utilized, the order in which the two types of experience are used results in two types of combined learning conditions: One combined strategy used by many learners includes initial, often intensive, formal study in the US which is subsequently followed by various naturalistic approaches and immersion. For our purposes this approach is called the formal/naturalistic (FN) condition. Other learners were first immersed and learned in a largely naturalistic fashion. Due to perceived inadequacies in their language
development, these learners later enrolled in more formal learning programs, such as intensive Russian programs following immersion experiences. This second combined strategy is called the naturalistic/formal (NF) approach. The third strategy that was targeted in this study is the strictly naturalistic condition or approach (NC). All of the participants in this study have experienced large amounts of immersion within Russian-speaking countries; no participants have had less than 4.3 years of largely full-time L2 Russian experience, including both formal and naturalistic types of exposure.

Three experiments were devised for this study. Two of the experiments were designed to examine working memory and its relationship to these three learning conditions. These two experiments have two different overall objectives. The objective for the first experiment was to investigate the degree to which working memory predicts four important characteristics of proficiency: accuracy, fluency, general proficiency or ‘native-likeness’, and the occurrence of potentially fossilized forms. A logistic UNIANOVA regression was utilized to test for statistical significance of the data for this experiment. The second experiment’s objective was to search for and examine potential interactions between working memory and the three learning conditions in relation to these aspects of proficiency. For the interactive experiment a moderated UNIANOVA regression was used. The third and final experiment involved an investigation of the learning conditions themselves to determine whether the types (i.e. formal instruction or naturalistic immersion) and order (the two different sequences of naturalistic and formal experience for the two combined contexts) have a significant predictive effect on the acquisition of these major components of proficiency. This third experiment also used
logistic regression for analysis of significance. All three experiments required a careful classification of participant experiences, including their duration and type.

**Attaining Accuracy in L2 Russian: Structures Targeted in the Current Study**

The acquisition of L2 Russian presents the adult English learner of this language with a great number of difficulties on various levels. One requirement is the mastery of morpho-syntactic accuracy: Russian has highly developed inflections for case, gender and number. Native Russian speakers also utilize some highly variable and context-specific patterns of syntax. In terms of morphosyntax, many of the most challenging forms are also very common in everyday speech; the learner must acquire a number of different complex relative clause structures. Another key difficulty is the need to master the complexities of verbal aspect and mood. For example, use of the perfective rather than the imperfective can create unintended consequences, such as the impression that the speaker is impatient, rude, or has an overbearing expectation of the listener, while the reverse error may indicate indecisiveness or a flippant attitude in some contexts. These challenges are further multiplied when overlaid on top of important paradigms such as motion verbs. Russian has grammaticalized a number of dimensions of motion, with distinctions between verb forms in terms of mode (e.g. vehicular versus bipedal motion), and the type of motion, including specific verbs for unidirectional and multidirectional or indeterminate movement. As with motion verbs, the Russian lexicon in general is highly context-specific: the English learner of Russian must acquire a multitude of forms that are highly specialized for use in specific contexts, with many instances of one-form to one-function lexical mappings.
Many of the forms that need to be mastered are characterized by a high degree of morpho-syntactic or semantic complexity, including several important relative clauses, such as the relative pronouns *kotoryj* and the *to, chto* construction or the subjunctive use of the conjunction *chtoby*. Because of their highly specialized nature and the importance of their usage, such structures are also relatively straightforward to elicit in a production task and are ubiquitous in native speech. While evidence of the use of such structures is elicited in this study, it is important to point out that the learners’ use of the structures in question is relied upon as a measure of conformity to native-like speech; the targeted items were not used in an instructional type of treatment, as is often seen in SLA research.

The first item targeted was a relative clause construction using the relative pronoun *kotoryj* ‘which, that’, which is used in a subordinate clause structure to refer back to specific nouns in a main clause and which demands accurate control of both the inflectional morphology and syntax demanded by the context of its clause. The relative pronoun must be in the initial position of its clause except when a preposition is used to modify it, when the preposition appears in front of *kotoryj*. In addition, the pronoun must reflect the gender/number of the noun it refers to in the initial clause, while also reflecting the inflectional morphology demanded by elements in its clause. Examples of this structure’s use following specific questions are seen in (1):

(1) a. ‘Kakaya zhenshchina molozhe?’ ‘Zhenshchina, u kotoroj chyornje volosy molozhe.’
   ‘Which woman younger?’ ‘(The) woman, with who (GEN/F) black hair younger.’
   ‘Which woman is younger?’ ‘The woman who has the black hair is younger.’

   b. ‘Kakije lyudi molodyje?’ ‘Lyudi, s kotorymi govorit Svetlana.’
   ‘Which people young (PL)?’ ‘People, with who (INSTR/PL.) speak (1SG) Svetlana.’
   ‘Which people are younger?’ ‘The people with whom Svetlana is speaking.’
Another similar relative pronominal structure used in complex sentences was also elicited: the common relative pronoun conjunction to, chto ‘that, what/which’, which demands control of both inflectional morphology and syntax. As with the relative pronoun kotoryj, this construction is highly representative of the syntactic and morphological control that is typical among native speakers, and is necessary in certain contexts due to the constraints of Russian. Examples are provided in (2):

(2) a. ‘O chym govorit Sergej?’ On govorit o tom, chem zanimaetsa Ivan.’
   ‘About what speak (3SG) Sergej?’ ‘He speak about that (LOC/SG), what (INSTR) busy Ivan.’
   ‘What is Sergej speaking of?’ ‘He’s speaking about what Ivan is doing.’

b. ‘O chym govorit Andrej?’ ‘On govorit o tom, chto emu nado idti domoj.’
   ‘About what speak (3SG) Andrej?’ ‘He speak about that (LOC/SG), that (NOM/SG) he (DAT/SG) need go (INF) home (ACC(INSTR)/SG)’
   ‘What is Andrej talking about? He is talking about the fact that he needs to go home.’

The third structure that was elicited is the subjunctive use of the conjunction chtoby ‘so that, in order that’. This construction demands accurate control of past tense verbal morphology (the subjunctive mood in Russian is expressed with the verb in the past tense), the control of verbal aspect, and a specific syntactic structure, as seen in (3):

(3) a. ‘Chto khochet Ivan?’ ‘On khochet, shtoby Sveta poshla s nim na vecherinku.’
   ‘What want (3SG) Ivan?’ ‘He want, that-be Sveta go (FSG/PAST) with him to party’
   ‘What does Ivan want?’ ‘He wants Svyeta to go with him to the party.’

These three constructions are representative of the level of syntactic and morphological structure and complexity that is attributed to native-like speech, and due to their important status for effective communication, are targeted for acquisition during the first or second semester of intensive formal instruction (usually between the second year and third year for learners in a standard college curriculum). They are very common in everyday speech in all areas of the former Soviet space and naturalistic learners
experience great amounts of exposure to the structures in the course of interaction with native speakers.

A Word about Terminology

Since the term “language aptitude” can often create considerable confusion due to its commonly broad application, for the purpose of the present study specific terms for component types of learner aptitude are identified and used. While other component abilities will at times figure into discussions of learner ID variables, the specific component of aptitude that serves as the focus of this research is working memory (WM), with the learner’s ability in WM characterized as the individual’s “working memory capacity” (WMC) or ‘working memory span’. Working memory is used here according to the definition found in Gathercole and 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). The WM construct includes a verbal or phonological component called the “phonological loop”, a similar visual component called the “visuospatial sketchpad”, a vital capacity that controls a learner’s attention control abilities called the “central executive”, and a storage component called the “episodic buffer” that processes information from various sources into a multi-faceted code (episode) that underlies one’s awareness capacity. Working memory may also be generalized to a broader discussion of “learner aptitude”.

The term learner “proficiency” or “second language proficiency” (SLP) is also used. In SLA literature proficiency is sometimes referred to as “observed performance” or “communicative competence”. In this study “proficiency” is defined as a diagnostic
measure or level of ability in the foreign or second language related to two main areas of oral performance: Accuracy, which for our purposes is defined as the extent to which the language used in performance demonstrates “the ability to produce error-free speech” (Lennon, 1990, p. 390) that is “elaborate and varied” (Ellis, 2003, p. 340) (“Elaborate and varied” speech is defined as a native-like level of complexity in sentence structure and command of vocabulary); and Fluency, which is defined as language characterized by “native-like rapidity” or the extent to which language “manifests pausing, hesitation, or reformulation” (Ellis, 2003, p. 342).

For determining accuracy in the present study, the language was analyzed in respect to the degree to which it incorporates specific morpho-syntactic structures common to native-like patterns of speech, while also evaluating discourse content (as compared to native speakers). It is important to note that this study is not an attempt to analyze the acquisition of these targeted structures following an instructional or non-instructional ‘treatment’. Rather, the accuracy part of the proficiency interview was designed to determine if the targeted structures had been previously acquired. An additional measure of overall accuracy was also utilized: the occurrence of potentially fossilized forms or fossilization rate (FR), which is measured by the number of recurrent errors in morphosyntax that occur per 100 words in the discourse elicited as part of the proficiency interview. The term ‘fossilization’ or ‘fossilized forms’ is used to describe the observation of apparently fossilized IL structures that have remained in the learner’s spoken communication in spite of years of exposure to and practice with the accurate use of such forms (Han, 2004), not in terms of the cognitive systems that are involved in the generation of fossilized structures.
The current study also involves an analysis of two general learning conditions (alone and in combination), formal instruction and naturalistic learning. The formal instruction condition refers to learning that occurs in a more traditional classroom context, such as college courses or intensive programs similar to Middlebury’s intensive Summer Russian Institute. Formal conditions typically include so-called “focus-on-form” instruction, including classroom instruction or tutoring in which the learner is consciously aware of the target forms or structures (phonological, morphological, or syntactic) that are being learned. However, for the current study it is important to distinguish between what in SLA research is traditionally called “explicit” instruction, which refers to specific focus-on-form techniques in controlled experiments, and “formal learning conditions”. The term “formal” in the sense used here simply refers to traditional types of classroom training that are commonly utilized in US-based language programs. The “naturalistic” condition is defined as a context in which learning occurs in the course of everyday communication or with a conversation partner, with primary emphasis on the learning of meaning. Under this condition, the acquisition of grammatical forms tends to occur incidentally, not in a focused or instructed manner.

1.3 The Relevance of Aptitude and Learning Conditions

Early researchers did not have the benefit of the body of knowledge related to language acquisition that has been acquired since the early 1980s. As a result of more recent SLA research, especially a number of key studies that were conducted after 1990, a quickly growing body of data has accumulated related to SLA theory. Much of this information has had a direct impact on our understanding of the role of aptitude in L2 acquisition and its importance in predicting second language proficiency.
There are a number of key variables that relate to the importance of learner aptitude in L2 acquisition, including age and learning context. Research over the past 20 years has demonstrated, for example, that adult L2 acquisition differs significantly from child acquisition not only in terms of success, but also in relation to the types of cognitive resources adults rely upon during learning, with analytic aspects of aptitude and the attention control capacity of working memory being far more crucial to adult acquisition of the L2 (e.g. DeKeyser, 2000; Harley and Hart, 1997, 2002; Ross, Yoshinaga and Sasaki, 2002). The results of these studies lend credence to Bley-Vroman’s (1989) Fundamental Difference Hypothesis (FDH), which posits a fundamental difference between adult and child acquisition processes. As such, IDs in aptitude have a direct bearing on any discussion of adult acquisition, since such differences naturally cause variability in these abilities, leading to different levels of proficiency. The connection between aptitude and proficiency, however, was ignored for over 20 years due to the erroneous perception that such cognitive abilities, though predictive of learning in the classroom, are unrelated to true language acquisition.

Perhaps the greatest impetus for the rejection of aptitude stemmed from the perception that there is a dichotomy between language ‘learning’ and true language ‘acquisition’. The predominant perspective among researchers during the early years of SLA was that true acquisition of an L2 occurred unconsiously and therefore only under implicit or incidental learning conditions. The belief was that more explicit or formal types of instruction, in which learning occurs consciously, merely promote language ‘learning’, not true ‘acquisition’ (see, for example, Bialystok and Fröhlich, 1978; Gardner, 1985; Krashen, 1981). It is apparent that since aptitude had been used to predict
success in the formal language learning classroom, it was assumed that measures of aptitude only predict learning under formal or explicit learning conditions. For these researchers aptitude had become irrelevant to the new communicative paradigm. Such sentiment was further compounded by research that erroneously held to a monolithic view of aptitude as being the equivalent to general (g) intelligence: researchers often used measures of IQ-related intelligence to test for correlations with L2 acquisition, and when none were found the conclusion was that ‘aptitude’ was irrelevant to true acquisition under implicit or incidental conditions. However, research has since demonstrated a strong predictive validity of aptitude measures for acquisition that occurs under any learning condition.

It has been demonstrated that learner aptitude is a more potent predictor of language-learning success than any other individual difference variable (Dörnyei, 2005; Dörnyei and Skehan, 2003; Skehan, 1989, 2002). Quantitative data from a number of studies has demonstrated the high predictive validity of aptitude, with correlations between aptitude assessment scores and achievement as high as .70 (Skehan, 1989). In a study conducted by Ehrman and Oxford (1995), it was found that measure of aptitude was the ID variable that correlated most powerfully with proficiency in an L2, with aptitude scores explaining as much as a 25% variance in second language proficiency (SLP) (Dörnyei, 2005). Such correlations between aptitude and achievement or proficiency are considerably higher than the next most important ID variable – motivation. ID research has also consistently demonstrated that measured aptitude is a potent predictor of L2 success in virtually any learning condition (e.g. Ehrman and Oxford, 1995; Harley and Hart, 2002; Horwitz, 1987; Reves, 1983).
In the last decade a growing number of researchers who have analyzed the relationships between various aspects of learner aptitude and variability in L2 performance or proficiency have come to the understanding that although aspects of aptitude represent key predictive variables related to L2 proficiency, such variables potentially do not function independently as predictors, but are closely related to elements of the context of learning. Robinson (2002b) has developed an aptitude-learning condition interactionist framework which describes the effects of aptitude in terms of the nature of different types of learning contexts. He emphasizes the need to ‘profile’ learners in terms of major aptitude abilities and match such learner profiles to optimum learning conditions. For example, based on research that has found strong effects of working memory capacity on the processing of lexical items in terms of speed and accuracy (e.g. Kroll and De Groot, 1997; Potter et al., 1984), learner profiles of WM can be used to predict learner difficulties in lexical processing under various learning conditions. Likewise, learner profiles in analytic language abilities, which have been strongly correlated with the acquisition of morphosyntax (e.g. Ross, Yoshinaga and Sasaki, 2002), can be used to create the optimum conditions for the acquisition of accurate L2 morphosyntax. The interactionist framework therefore provides a platform for the study of the relationships between measures or profiles of key abilities, such as working memory, and acquisition or proficiency under different learning conditions.

Among the different aspects of learner aptitude, working memory stands out as perhaps the most important cognitive ID variable (Bowden, Sanz, and Stafford, 2005; Miyake and Friedman, 1998). WM has been implicated in a great number of language-related abilities, including the acquisition of lexical items (e.g. Daneman, 1992; Kroll and
De Groot, 1997; Potter et al., 1984), phonological aspects (e.g. O’Brien, Segalowitz, Collentine, and Freed, 2006), as well as fluency, accuracy and complexity in speech production (e.g. Mota, 2003; Mota and Bergsleithner, 2007). Although our understanding of the role of working memory in L2 acquisition and performance has grown significantly, research is needed that can elucidate or reveal the extent to which abilities like WM interact with the nature of learning contexts to produce the different elements of a learner’s IL grammar. Robinson’s (2002b) interactionist framework appears to be the best fit for this type of research.

Besides the lack of research designed to discover potential interactions between aspects of aptitude and learning conditions in producing major characteristics of learner proficiency, past research has provided, at best, an incomplete picture of the acquisition processes involved with the different learning conditions and their relationships with the characteristics of the IL grammar that demonstrate the greatest variability among learners. Much remains to be learned about the exact nature of learning conditions and how they impact learner proficiency. The current study therefore represents a step in the direction of addressing these twin issues related to the roles of aptitude and learning context in the nature of learner proficiency. The information provided by such a research framework can potentially become a vital part of determining the causes of inter-learner and intra learner variability in various aspects of proficiency, especially as it relates to some of the more elusive accuracy-related characteristics such as fossilization.

1.4 Overview of the Dissertation

The remaining body of this dissertation is divided into five additional chapters.
Chapter Two provides a review of the current literature related to the following topics: 1) cognitive perspectives on variability and fossilization in Adult SLA; 2) the nature of the aptitude complex of abilities; 3) models of aptitude-oriented interactionist research; 4) working memory and its role in the attainment of L2 Proficiency; 5) challenges related to the acquisition of L2 Russian; 6) the challenges related to the attainment of L2 Russian proficiency; and 7) the study’s research questions and hypotheses. Chapter Three provides a description of the research methods that were utilized in the study, including the challenges of its design, participants, testing materials used, and data analysis procedures. Chapter Four reports the results of the three main experiments as well as a general overview of relationships seen in the data. Chapter Five provides a discussion of the results for both the general patterns seen in the data and for the three experiments. 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 the following subsections:

1) The results for working memory and its predictive relationship with participant fluency and accuracy, related to the first research question.

2) Results for the experiment testing for predictive relationships between WM and observed fossilization. This section relates to the second research question.

3) Results for the interaction experiment testing for potential interactions between working memory and learning conditions.

4) A description of the relationships between different learning conditions and accuracy-related measures of proficiency, with an emphasis on the sequencing of the two combined learning conditions.

Chapter Six provides a conclusion with recommendations for future research.
CHAPTER II: LITERATURE REVIEW

The objective of the current study is to describe the characteristics of adult second language proficiency (SLP) in L2 Russian which result from an interaction of working memory and the types of learning conditions and strategies that are most commonly experienced by English-speaking learners of Russian. In order to provide the proper backdrop for this study, the summary of literature that follows examines what is currently known about the following key areas of SLA research: a) the cognitivist perspective on the nature of inter-learner and intra-learner variability and fossilization in the interlanguage (IL) grammar; b) the nature of the aptitude complex of abilities and the construct of working memory; c) models of aptitude-related interactionist research, including Robinson’s aptitude-learning condition interaction model; d) what is currently known about the role of working memory in the attainment of L2 proficiency; and e) the challenges English L1 learners of L2 Russian experience in attaining proficiency, with emphasis on both the overall difficulties in becoming proficient and the types of morpho-syntactic structures and grammatical paradigms necessary for native-like accuracy and complexity, including those forms targeted in the current study;

2.1 Variability and Fossilization in Adult SLA: Cognitivist Perspectives

The existence of a high degree of inter-learner variability in terms of both overall success and the presence of fossilized grammatical forms has been one of the most perplexing problems in SLA. The cognitivist view of variability and fossilization is best characterized as a continuation of the information processing approach to SLA that was
established in cognitive psychology. Within the cognitivist paradigm, the high degree of variation in adult L2 attainment has perhaps been the main driving force behind individual difference (ID) research. In response to the limitations of contrastive analysis, cognitivists explained variation in adult L2 success largely by relying on IDs in affective and cognitive variables (internal factors). However, due to the impact of Krashen’s early model of SLA, many researchers turned chiefly to affective variables such as motivation (e.g. Gardner, 1985), learning style (e.g. McDonough, 1981), or degree of acculturation (e.g. Hubbard, 1975), and the role of social and individual affective factors in producing variability is well documented in research (e.g. Gardner, 1985; Gardner and Lambert, 1972; Williams, 1994) (Dörnyei, 2005). The most influential model of affective variables, integrative motivation, was developed by Gardner (2001) as a synopsis of the relationships between motivation and other IDs and success in L2 acquisition (Dörnyei and Skehan, 2003): measures of motivation have been widely used for predicting overall L2 attainment. Yet among information-processing approaches to the phenomenon of adult variability in ultimate attainment, cognitive IDs tend to be seen as the most crucial factors in the processing of input (Bowden et. al, 2005; Sanz, 2005).

Of all ID variables, IDs in memory, attention and other cognitive abilities are relied upon as the most potent predictors of adult L2 success. Such IDs have been widely used to explain variability in attainment (see Dörnyei, 2005; Dörnyei and Skehan, 2003; Skehan, 1989). In examples of modern cognitivist research the analysis of learner variability or fossilization in L2 proficiency may draw conclusions based on all the above ID factors, including aptitude, and affective variables such as motivation, learning style, and field dependency. With their status as the most powerful predictors of L2 success,
differences in aspects of learner aptitude are seen by many modern researchers as a primary cause of such variability (see discussion under section 1.1). Although cognitive variables remain a primary focus in the cognitivist tradition, a great number of factors may act together to cause learner variability in ultimate attainment, and consequently fossilization. Research has therefore come to include a wide variety of variables that may contribute to these phenomena.

In light of evidence for a critical period, especially as it relates to variable success in attaining native-like proficiency, cognitivists have come to emphasize the interaction of a broad array of variables in L2 acquisition that may act together to produce the variation seen in adult acquisition; because of its emphasis on such interactions, cognitivist research is sometimes described as ‘interactionist’. Variables often analyzed include characteristics of the input, learner ID variables (e.g. cognitive resources, or affective factors like motivation), and strategies or conditions (for example, see Ellis, N., 2003; Robinson, 1997b, 2002b). Cognitivist research tends to emphasize both the how and why of acquisition. Their findings are therefore generally applicable to pedagogical practices.

A number of researchers have cited an interaction of different constraints, such as both biological and cognitive factors to cause adult acquirers to be “preconditioned to fossilize” (Han, 2004, p. 9), such as Birdsong and Molis (2001) and Dekeyser (2000). Other researchers argue that phenomena such as fossilization and its cross-learner variation is caused by a combination of L1 transfer and other variables, including linguistic, social, psychological, or age factors (e.g. Han, 2000; Han and Selinker, 1999; Selinker and Lakshmanan, 1992). Selinker and Lakshmanan (1992) argue for an
important role of L1 transfer in fossilization in what they call the “multiple effects principle” (MEP), or an interaction of key variables producing fossilization. They state that the MEP is strongest when language transfer combines with one or more factors which together produce a stabilization of the IL grammar.

In recent years aptitude-related research has been used to speak to the issues of variation and fossilization in adult L2 attainment. One example is provided by the aptitude-processing stage interaction model developed by Skehan (1998), in which the learner’s various cognitive abilities are matched to specific acquisition processes. Dörnyei and Skehan (2003) describe strengths in certain cognitive abilities used to process language input as enabling the learner to avoid failure. For example, in Skehan’s model (1998), the stage of acquisition called ‘pattern restructuring and manipulation’ relates to the learner’s apprehension and reformulation of inaccurate or undeveloped rules in the IL grammar so that they more accurately reflect the TL grammar; according to this view, if successful, it is this processing stage that acts as the “anti-fossilization stage of development” (p. 599). Specific abilities necessary for the task are identified, including analytic language ability (inductive language learning ability) and grammatical sensitivity. Research can then test for correlations between these two cognitive abilities and success during this stage of processing by measuring aspects of observed performance or attained level of proficiency. As such, Skehan’s (1998) aptitude-processing stage interaction model can be used to examine potential cognitive causes of variation in adult ultimate attainment. Robinson (2001, 2002b) has articulated a similar approach that involves other variables besides the processing of input, such as various learning conditions.

Under the model articulated by Robinson (2001, 2002b), component cognitive
abilities interact with various learning conditions to produce overall attainment in the L2. Called ‘aptitude-learning condition interactions’, this particular model of adult L2 acquisition looks at both internal IDs related to aptitude and external factors (learning conditions) to describe variation in L2 attainment. The model describes the relationships between certain cognitive abilities and the nature of learning conditions, citing research that shows significant correlations between specific abilities and the type of learning that occurs in various learning contexts, including explicit, implicit and incidental conditions. Such research indicates that some aptitude abilities are more strongly implicated in learning under certain conditions than others, such as studies conducted by Harley and Hart (2002) and Ross, Yoshinaga and Sasaki (2002), which provide evidence for the idea that different learners rely on different clusters of abilities in their development of proficiency under various learning conditions, resulting in variability in the degree of success achieved. Research under interactionist models such as Skehan’s aptitude-processing stage approach or Robinson’s model of aptitude-learning condition interactions can therefore provide insights into the causes of variability in adult L2 attainment, and possibly the issue of fossilization as well.

**Cognitivist Research Related to Variability and Fossilization**

Cognitivist research that speaks to the issue of variation and potentially fossilization appeals to a number of potential explanations for these phenomena. Research related to the issue of variation in adult L2 attainment is particularly well attested, and a number of conclusions have been drawn pertaining to the role of cognitive factors in causing such variability. DeKeyser (2000) reported results that indicate significant positive correlations between analytic aptitude abilities and proficiency related
to grammaticality judgments for learners who arrived in the US following the critical period, whereas such correlations were not seen for learners who arrived prior to the CP. Harley and Hart (1997) specifically examined the interaction of aptitude and age with both learning context and outcomes in terms of proficiency. Similar to DeKeyser (2000), they found differential correlations between aptitude and age of arrival, with memory-related abilities predicting success in attainment for pre-critical period learners, whereas analytic abilities showed positive correlations with proficiency for learners who arrived as adults after the critical period. In order to rule out the possible influence of the type of formal instruction experienced by adult learners (which was largely analytic in the first study), Harley and Hart (2002) conducted a second follow-up study and again found positive correlations between analytic ability and proficiency attainment for adult learners under naturalistic conditions. Such studies appear to demonstrate that adult learners who score lower in analytic aptitude fair worse in the attainment of proficiency, which has clear implications for a potential cause of variability in learner success.

Other studies within the cognitivist tradition report similar results, and appear to demonstrate strong positive correlations between various measures of aptitude and success in adult L2 learning. Reves (1983) investigated 11th and 12th grade Arabic L1 learners of both English and Hebrew. The study’s participants had been largely immersed in Hebrew from about grade 5, with classroom instruction in the language beginning around the same time. Several aspects of the participants’ proficiency, including oral fluency and morphological accuracy in both Hebrew and English were rated. A predictive analysis using multiple regression identified measures of aptitude as providing the greatest proportion of variance for all proficiency measures in both English
and Hebrew (Reves, 1983). Robinson (2002c) conducted a study of Japanese L1 learners of Samoan and tested for correlations between working memory and grammar learning under incidental learning conditions. Participants were required to learn Samoan words by rote, followed by exposure to 450 sentences, after which they answered test questions related to comprehension. All participants received scores in comprehension above 95%, and were later assessed in post-tests that occurred at intervals right after training, a week after training, and six months later, which included grammaticality judgments and the production of sentences. Significant positive correlations were found between all aptitude measures and the incidental learning of grammar in Samoan, with the strongest correlations seen between WM and accuracy in grammatical usage (Robinson, 2002c).

A number of other studies likewise indicate strong positive correlations between measures of cognitive abilities (aptitude) and success in L2 proficiency, with findings that are applicable to explaining the high degree of variation seen among adult learners, such as studies by de Graff (1997b), Ehrman and Oxford (1995) and Horwitz (1987). Skehan (1980) examined the relationship between aspects of memory and adult success in learning L2 Arabic. He found that scores for “memory for text” and what he has called “response integration” or memory used for acquiring unknown structures, showed greater validity in predicting L2 success than other memory-related measures, such as the paired associates (PA) subtest of the MLAT. A later analysis of the data (Skehan, 1986) found that there were two different profiles for learners who had been successful: younger learners relied on memory abilities, while older learners depended on analytic types of aptitude, similar to results found by DeKeyser (2000), Harley and Hart (2002) and Wesche (1981). The existence of such significant positive correlations between measures
of aptitude and adult proficiency in an L2 is clearly applicable to explaining the high degree of variability in success experienced by adult learners, and the findings related above are cited as evidence supporting Bley-Vroman’s Fundamental Difference Hypothesis (see discussion by DeKeyser, 2003; Robinson, 2002b; Skehan, 2002).

**Variation and Fossilization: Other Perspectives**

Although there are a number of other perspectives on variation and fossilization besides the cognitivist paradigm, the study of these two phenomena has brought about a convergence of the different views in the very place where research on variation of adult L2 performance began: cognitive resources related to language learning aptitude. For example, within the generativist theoretical perspective much of the past research in this area was focused on questions related to a learner’s access to UG and the role of L1 transfer, with five possible variations of their relationship discussed in the literature, including ‘Full Access-Full Transfer’ (Schwartz and Sprouse, 1996), ‘Partial Access-Full Transfer’ (Schachter, 1989), ‘Full Access-Partial Transfer’ (Vainikka and Young-Scholten, 1994), ‘Partial Access-Partial Transfer’ (Eubank et al., 1997), and ‘Full Access-No Transfer’ (Epstein et al., 1996) (Han, 2004). The greatest challenge associated with these positions relates to how their predictions fit the empirical evidence for the nature of final state grammars, including the high degree of variability in success consistently observed (Bley-Vroman, 1989), and the possibly universal tendency for learners to experience fossilization in various aspects of the IL grammar. With their emphasis on the access/transfer question and orientation towards the acquisition of particular grammatical structures rather than the attainment of L2 proficiency, many UG studies inevitably fail to address the issue of learner variability or fossilization. This situation has begun to
change, however, and a growing number of generativist researchers have also attempted to explain both the lack of success and the variation in overall attainment commonly observed among L2 learners in ways that appeal to other variables outside of the L1 transfer-UG access paradigm.

Bley-Vroman (1989, 1990) concluded that variation in adult attainment is explained by a fundamental difference between adult and child acquisition (the Fundamental Difference Hypothesis, or FDH), and that adult learners no longer rely on UG, but rather upon their cognitive abilities. The strong version of the FDH states that an adult’s linguistic knowledge and problem-solving abilities make up for an inability to use or access the system used in childhood. The weak version argues for a partial access to UG that helps to explain the success that can and does occur (Bley-Vroman, 1989). This makes sense to many researchers since it is widely recognized that UG alone is inadequate to the task of acquisition, with extreme limits to the aspects of language that are controlled exclusively by UG (e.g. Doughty, 2003; Doughty and Williams, 1998). The FDH states that adult L2 variability must be caused by other factors, such as differences in the cognitive resources relied upon by adult learners, the effect of learning context

Other researchers similarly attribute variability in adult L2 success to external sources outside of UG, such as the nature of the language input or the type of learning environment. Doughty (2003) states that “without instruction, adult SLA is more difficult, slower, and less successful” (p. 259), implying that differences in outcomes are attributed to learning strategies and conditions. Sorace (2003) recognizes that there may be a dependence on fundamental analytic abilities for adult learners, and that such
abilities may be utilized instead of UG in L2 acquisition, particularly if UG no longer functions after the critical period. In terms of variability in ultimate attainment and the fossilization in final state grammars that is observed, it seems that a growing consensus among researchers from various perspectives is afoot: variability in success appears to be largely related to differences in cognitive resources. These cognitive resources naturally figure into any discussion of the major aspects of observed oral proficiency. A key question is whether research related to cognitive variables within the construct of aptitude can be used to shed light on the potential causes of adult L2 variability and fossilization of major aspects of the IL grammar.

Although a great number of different explanations for fossilization have been proposed, many of these potential causes appear improbable since the majority of them cannot make predictions of the phenomenon due to their relatively universal nature, including general human traits or widespread characteristics of the learning environment (Long, 2003). Selinker (1993) views fossilization as a type of linguistic process that is constrained by the characteristics of the L1, but is commonly manifested differentially by different individuals. According to Long (2003), variables that are “immutable and the same for everyone could only work as explanations (of fossilization) for the entire population of L2 learners and for all structures if they worked at all” (p. 515). He provides an extensive list of explanations that would fail such a test, including either partial or full loss of UG access, immutable characteristics of the relatedness of the L1 to the L2 (the Multiple Effects Principle), learner-wide cognitive resources that are variable in degree alone, and maturational constraints considered to be universal. It would appear that fossilization, if demonstrably proven, may turn out to be impossible to explain.
However, by using a process of elimination to rule out other proposed causes on both empirical and logical grounds, Long (2003) concludes that only one likely candidate remains that cannot be ruled out because of the above concerns: sensitivity to input. He argues that the nature of input is highly stable regardless of the learning condition, and that L2 acquirers exhibit highly variable sensitivity. As a component of learner aptitude that has been well documented in the SLA literature (see Dörnyei and Skehan, 2003), this particular type of learner ID is described as being involved in three of the original components of aptitude proposed by Carroll: phonetic coding ability, inductive or analytic learning ability and grammatical sensitivity. It turns out that research within the cognitivist tradition, especially the aptitude-related interactionist models proposed by Robinson (2001, 2002) or Skehan (2002) may therefore hold the key to discovering the cause of this elusive issue in SLA.

2.2 The Nature of the Aptitude Complex of Abilities

As stated earlier, the development of the tests devised by Carroll, Pimsleur and many others was based upon a great number of different test trials with thousands of subjects. Clearly a psychometric approach, Carroll’s research model has inevitably led to what Dörnyei (2005) has called an “atheoretical, assessment-driven view” of the aptitude construct (p. 36); the tendency of traditional aptitude tests to differ greatly from one another is a natural result of the assessment-driven, trial-and-error method that was typically used to develop them, and directly reflects widely disparate views of the nature of the construct. Some tests, like the MLAT, rely mainly on an assessment of innate language abilities (with some exceptions), while others include largely experience-driven components of ability, including an evaluation of an individual’s motivation or other
‘soft’ aptitude factors, which is the case with the PLAB. Additionally, the tendency of early researchers to equate aptitude with intelligence has been the source of considerable confusion in SLA research; a number of researchers have evidently based their conclusions about learning conditions and the nature of acquisition upon this view, having utilized IQ-related IDs to test for correlations with L2 acquisition under different learning conditions (e.g. Reber et al., 1991; Maybery, Taylor and O’Brien-Malone, 1995; McGeorge, Crawford and Kelly, 1987). Due to the common confusion between aptitude and intelligence, a logical first step in any discussion of the nature of aptitude is to examine the relationships between aptitude and intelligence.

**What is the relationship between aptitude and intelligence?**

The study of human intelligence has had a direct bearing on our understanding of language aptitude, and early aptitude researchers often equated aptitude with intelligence. In fact, this view still persists, and a number of researchers base their claims about the predictive validity of aptitude on studies that compare measures of intelligence on standard IQ tests with L2 acquisition in different learning conditions, e.g. Ellis, Katz and Williams (1987), Maybery, Taylor and O’Brien-Malone (1995) and Reber (1991). Such confusion demonstrates the importance of intelligence to this discussion.

Views of intelligence have changed from the early concept of a unified general (g) intelligence, to the current belief that intelligence consists of a variety of different cognitive abilities. Most scholars today consider intelligence to be divided into different types or components, with the most distinct division being between g or “crystallized intelligence,” and Gf or “fluid intelligence”. It is important to note that crystallized intelligence represents what is generally learned or “crystallized.” It is this crystallized
form of intelligence that is generally evaluated on IQ tests. Fluid intelligence is considered to be a distinct sub-domain of crystallized intelligence, and relates to a person’s capacity to use cognitive reasoning abilities to solve unexpected problems without prior knowledge. Jaeggi, Buschkuehl, Jonides and Perrig (2008) (Jaeggi et al.) point out that $G_f$ is broadly recognized as crucial to learning and to the accomplishment of a great variety of tasks. Fluid intelligence has been shown to be a good indicator of academic and professional success, and there is substantial agreement that this type of intelligence is quite impervious to effects associated with training or education (Jaeggi et al., 2008). Research shows some aspects of language aptitude to be closely related to fluid intelligence, while other components may be linked to crystallized intelligence.

The differences between aptitude and general intelligence are pointed out in research that demonstrates high levels of second language success in individuals who score below average or average on IQ tests, including studies that appear to indicate that those who have a natural talent for learning an L2 possess specific cognitive abilities that distinguish them from other learners (Altarriba and Basnight-Brown, 2009). For example, Novoa, Fein and Obler (1988) found that in the case of one American teenager who possessed only average to slightly above average IQ, high proficiency with native-like performance was attained in several different foreign languages, while the individual scored highly on a measure of the ‘Raven’s Progressive Matrices’ test, a non-verbal task that measures as individual’s pattern recognition ability (Obler and Gjerlow, 2002).

One similarity between aptitude and intelligence is that both constructs appear to consist of different and distinct components. A number of researchers, including Gardner (1983, 1999) and Sternberg (2002) have developed concepts of “multiple intelligences,”
theories that represent a paradigm that diverges from the traditional, more unified \( g \) theory of intelligence. Sternberg has proposed a three-fold view of intelligence called the “theory of successful intelligence,” consisting of “analytical,” “creative” and “practical” types of intelligence (Sternberg, 2002). Recent research into the nature of crystallized intelligence provides substantial evidence that intelligence does not consist of a unified domain and neither does it exactly correlate with aptitude (Sternberg, 2002). Dörnyei (2005) has stated that there is a limited relationship between some of the components of language aptitude and crystallized intelligence, though without exact correlation of abilities between the two domains. The association of some components of aptitude with crystallized intelligence (what is learned), and others with fluid intelligence (what tends to be innate) makes an exact analysis of language aptitude a crucial part of SLA research. It is clear that though different components of aptitude may be related to types of intelligence, to equate the two with each other is at the very least overly simplistic and inaccurate. Without an exact determination of the relationships between components of aptitude and intelligence, research that attempts to correlate measures of intelligence with success in L2 acquisition cannot be relied upon, and serves to compound confusion related to aptitude. Confusion of aptitude with intelligence has caused some language teachers to avoid aptitude altogether due to fears that it only serves to discourage potential learners.

When combined with other past misconceptions about aptitude, such as the belief that it only predicts formal types of learning (e.g. Cronbach and Snow, 1977; Krashen, 1982; McLaughlin, 1980), it is no surprise the construct has received such little attention in SLA literature. In light of research that has demonstrated the relevance of aptitude as a
predictor of proficiency in any learning condition (e.g. Ehrman and Oxford, 1995; Horwitz, 1987; Harley and Hart, 2002; Reves, 1983), this failure to adequately address the effect of such IDs in L2 acquisition makes the present study a vital step in furthering our understanding of the specific roles key aspects of aptitude (e.g. working memory or analytic language ability) play in different learning contexts.

**Different views of the nature of aptitude**

A comparison of the MLAT and PLAB, the two most widely influential tests for aptitude, clearly illustrates the lack of a consensus among early researchers regarding the nature of the construct. Though there was certainly some agreement about the different components of aptitude, such as Pimsleur’s and Carroll’s similar views of “inductive language learning ability,” such wide differences between tests appear to reflect a variety of often disparate conclusions regarding the nature and components of aptitude, while total disregard of some component abilities (e.g. Carroll’s exclusion of inductive or analytic language ability) reflects different perspectives as to the relative importance of various abilities. A comparison of the MLAT and PLAB demonstrates the lack of a clear and concise theoretical basis for early researchers’ conclusions, and underscores the need for a more coherent theory of language aptitude:

1. Carroll (Carroll, 1981):
   - *Phonetic Coding Ability* - “An ability to identify distinct sounds, to form associations between these sounds and symbols representing them, and to retain these associations.” (p. 105)
   - *Grammatical Sensitivity* – “The ability to recognize the grammatical functions of words (or other linguistic entities) in sentence structures.” (p. 105)
   - *Rote learning ability* – “The ability to learn associations between sounds and meaning rapidly and efficiently, and to retain these associations.” (p. 105)
   - *Inductive Language Learning Ability* – “The ability to infer or induce the rules governing a set of language materials, given samples of language materials that permit such inferences.” (p. 105)
2. Pimsleur (Pimsleur, 1966):

*Verbal Intelligence* – “The knowledge of words and the ability to reason analytically in using verbal materials.” (p. 14)

*Motivation*

*Auditory Ability* – “The ability to receive and process information through the ear.” (p. 14)

Of particular interest in Pimsleur’s description above is the inclusion of *motivation*, which is clearly out of place in any construct of aptitude (Dörnyei, 2005). The result of this view was that the PLAB included ID measures such as interest in studying a foreign language and past success in coursework. Though IDs such as motivation or past learning experiences can certainly help to predict future learner success in SLA, the inclusion of such factors as a part of “aptitude” only serves to complicate our understanding of the exact nature of the domain. The differences between the tests are the natural result of the non-theoretical nature of the research, resulting as well in a failure to develop more effective methods of testing. It is clear that if the true nature of the different components of learner aptitude is to be accurately determined and correlated to acquisition, a rigorous, theoretically based research program is required. A number of researchers have called for exactly that (e.g. Dörnyei, 2005; Robinson, 2002a; Skehan, 2002).

While not numerous, a number of aptitude research studies were conducted during the 70s and 80s that did serve to shed some light on the various components of aptitude, including Curral and Kirk, 1986; Curtin, Avner and Smith, 1983; Eisenstein, 1980; Neufeld, 1978, 1979; Nizegororcw, 1980; Schneiderman and Wesche, 1986; Skehan, 1986; Wesche, 1981; Wesche, Edwards and Wells, 1982; Zeidner, 1986 (Dörnyei, 2005). Though such studies did advance our understanding of aptitude, Carroll concluded that such research did not motivate any significant changes to the various components that he
and others had proposed since the beginning of aptitude research (Carroll, 1991). Because of the inconsistencies and lack of agreement between early descriptions of aptitude, however, a number of researchers have attempted a reanalysis of the construct.

One of the first attempts to rework Carroll’s construct is seen in a study conducted by Skehan (1989). After an in-depth analysis of Carroll’s concept of aptitude, Skehan concluded that aptitude is comprised of three general components instead of the four suggested by Carroll: *Auditory ability, linguistic ability* and *memory ability* (Skehan, 1989). Skehan’s “linguistic ability” represents a combination of Carroll’s grammatical sensitivity and inductive language learning ability (Dörnyei, 2005). It is evident that an improved clarification of the nature of the aptitude domain would require additional research. The result has been the modern development of a number of different approaches to the construct, including those that examine the interaction of aptitude components and SLA processes, forms of instruction, and learning conditions. This last approach represents more a research paradigm that matches certain aptitude abilities to SLA processes or learning conditions than a new attempt to define component abilities that constitute learner aptitude. Examples of such ‘interactionist’ research approaches to aptitude are represented by Skehan’s (2002) “componential interactionist” approach and Robinson’s (2001, 2002) “aptitude-learning condition interactions” approach. There have been some additional attempts, however, to more accurately define the construct of learner aptitude.

One recent approach to aptitude research has resulted in a theoretically-based aptitude exam. Grigorenko, Sternberg and Ehrman (2000), whose “Cognitive Ability for Novelty in Acquisition of Language as applied to foreign language test” (CANAL-FT)
have developed an aptitude model that is related to Sternberg’s (2002) threefold model of intelligence (Dörnyei, 2005). The CANAL-FT represents the first aptitude test that is based on theory. The test is designed to measure the way in which learners deal with new and ambiguous information through the gradual introduction of an artificial language with participants required to accomplish several short tasks related to five processes involving the acquisition of knowledge in the new language (Dörnyei, 2005). These new approaches to the nature of the aptitude complex of abilities reveal a growing consensus that aptitude consists of a highly differentiated set of component abilities with varying degrees of interrelatedness between them.

Research has revealed that a distinct separation exists between various cognitive abilities. One apparent observation that relates to this separation is the fact that specific abilities predict success in specific areas of L2 acquisition, such as findings that demonstrate that musical talent, as measured by the Wing measures of musical talents (Wing, 1968), may account for variance in bilinguals’ perception and pronunciation of English phonemes, but not their knowledge of lexical items or grammatical forms, which was found by Slevc and Miyake (2006), who suggest that musical talent may be used by some learners who commonly “rely on other, nonlinguistic mechanisms and abilities to aid in L2 acquisition” (p. 679); they point out that such reliance may be motivated by the greater difficulties typically experienced by later L2 acquirers. Though musical talent has been shown to play a role in the acquisition of phonology in an L2, some research indicates that it is not a necessary requirement of overall L2 acquisition (Novoa et al., 1988).
Research has also demonstrated differences between the abilities relied upon for the incidental acquisition of semantic information and those used for the learning of grammatical forms. For example, Ross et al. (2002) report results that indicate that learners following the critical period rely upon analytic language abilities in making grammaticality judgments, while such abilities were not utilized by learners prior to the critical period (see also DeKeyser, 2000). Similarly, Harley and Hart (1997, 2002) found that memory-related abilities are linked to pre-critical period learners, whereas post-critical period learners tend to rely upon analytic language learning abilities if they are to have success in acquiring the L2. Since pre-critical period learning tends to occur incidentally with an emphasis on the learning of meaning, while post-critical period learning more commonly involves the learning of form, it is logical to conclude that such differences point to a differentiated complex of abilities that are utilized for specific aspects of L2 acquisition.

Further evidence of the separation of distinct cognitive abilities is seen in research utilizing both factor and regression analyses to determine the degree of relatedness between component abilities, such as Canner (2011), who found a distinct separation between certain memory component abilities, such as the paired associates subtest of the MLAT, and the analytic language learning ability measured in the DLAB and PLAB aptitude tests. Such findings point out the importance of research that investigates the exact nature of individual components of aptitude, their interrelatedness, and the degree to which they are ameliorable to training. Though there is still much to be learned about the extent to which component abilities are interrelated or dependent on one another,
research has revealed some interesting characteristics of the distinct component abilities that comprise the language aptitude construct.

**Analytic Language Learning Ability and Grammatical sensitivity**

Carroll’s aptitude model (Carroll, 1981) included two analytic components: *inductive language learning ability* (referred henceforth as analytic language ability or ALA) and *grammatical sensitivity*. He defines grammatical sensitivity (GS) as the “ability to recognize the grammatical functions of words (or other entities) in sentence structures” (Carroll, 1981, p. 105). Analytic language ability, on the other hand, is defined by Carroll (1981) as “the ability to infer or induce the rules governing a set of language materials, given samples of language materials that permit such inferences” (p. 105). This analytic or inductive ability has been summarized by Skehan (2002) as the “capacity to extrapolate from a given corpus to create new sentences” (p. 71). The ability has to do with taking lexical and grammatical material (patterns, rules and other information) and then extrapolating or generalizing it to form new sentences. As such, ALA is largely an active, processing type of ability. GS, on the other hand, involves an analysis of material to determine the syntactic function of parts of speech, making it a largely passive ability. Interestingly, Carroll did not include a subtest for analytic language ability on the MLAT, though such a subtest is included in the PLAB. These two abilities constitute the more analytic aspects of the aptitude construct.

Of interest to this discussion is the fact that the analytic aspects of aptitude have not received much attention in research, though the results reported by a few studies are worth noting (Skehan, 2002): Harley and Hart (1997) conducted a study of the types of language aptitude that correlate with learner outcomes in a secondary school immersion
program for French begun at either early or late ages, and they found that a measure of ALA best predicted L2 success for late immersion subjects (grade 7), while memory best predicted success for early immersion subjects (grade 1) (Harley and Hart, 2002). In a related study, Harley and Hart (2002) examined the role of both memory ability and analytic language ability in learning an L2 in a natural immersion setting. The participants were high school (grades 10 and 11) L1 English learners of L2 French who lived with French-speaking families in a three month immersion program. The subjects were initially tested for aptitude with both a “memory for text” task and an “analytic language ability” task (patterned after the PLAB’s analytic subtest), as well as a French proficiency test that included vocabulary and reading comprehension tasks, with similar tests administered as follow-up with correlations made with the aptitude assessment.

They found that both the memory for text and analytic language abilities predicted success, though the memory task was not as effective a predictor as it was for the early immersion participants of the earlier (1997) study. They also found that the analytic language task was a good predictor of adult learner success in the natural setting. Harley and Hart (2002) therefore argue that the study “offers some support for the argument that age of initial intensive exposure is a factor affecting students’ L2 learning orientation, whether inside or outside the L2 classroom environment.” (p. 326)

In a study of British military learners of Arabic, Skehan (1986) conducted a cluster analysis of the correspondence of memory-related aspects of aptitude with analytic components. His goal was to examine patterns in individual aptitude “profiles” related to success in learning Arabic, and showed that general levels of aptitude as a composite ability did not accurately correspond with the data, while finding that learners
who are successful attain a level of competency either through the use of high capacity memory ability or high analytic abilities (Skehan, 2002). Such results indicate that different learners may utilize different resources to achieve success, while it also follows that those who possess strengths in both memory and analytic abilities represent the most successful learners of an L2.

In research designed to examine the relationship of aptitude with intelligence, Sasaki (1991, 1996) used three subtests of a Japanese version similar to the short form of the MLAT, including Paired Associates, Language Analysis (it is assumed that this refers to a task similar to the language analysis task on the PLAB) and Sound-Symbol Association tasks to draw correlations between measures of aptitude, proficiency measures and both verbal and cognitive reasoning aspects of intelligence. The intelligence subtests included both verbal (predominantly g intelligence) and reasoning or cognitive (fluid intelligence) tests. Results of a first order factor analysis demonstrated a separation between aptitude and intelligence scores; a second order factor analysis, however, indicated that one of the factors could explain the difference in some variables for aptitude and the measures for intelligence: the analytic language component of aptitude. In her conclusion, Sasaki states that “a general factor of second language proficiency is related to, but not identical with general cognitive abilities…The most notable changes…were that two of the three indicators of aptitude had much stronger links with the Reasoning factor.” (Sasaki, 1996, p.134) The two indicators mentioned above refer to the language analysis and sound-association (phonetic coding ability) subtests.
It appears that analytic or inductive language ability may correlate to an extent with cognitive reasoning ability, with adults relying on such analysis to varying degrees. The above-mentioned studies conducted by Skehan (1986) and Harley and Hart (2002) indicate that adults and children rely on different abilities in learning, with later learners relying more on cognitive or analytic components of aptitude, and early learners favoring memory. The Sasaki study (1996) appears to draw a correlation between analytic language ability and cognitive reasoning abilities or fluid intelligence, with the association implying that the analytic components of aptitude represent predominately rigid or static types of aptitude. But where does Carroll’s grammatical sensitivity (GS) component fit in?

Carroll (1981) stated that though some aptitude tests other than the MLAT have not used a GS subtest per se (e.g. the PLAB-4 subtest “linguistic analysis” or the DLAB), GS ability may be represented in them in various forms. Carroll’s assumption seems to be that the GS subtest is the equivalent of other analytic language tests. The DLAB’s analytic language subtest provides the learner with the vocabulary and grammatical rules of an artificial language, and then tests the ability to apply rules and extrapolate new morpho-syntactic patterns. The PLAB provides vocabulary and short sentences (with English translations) in an artificial language and then asks the learner to extrapolate what is seen in those sentences to choose a correct new sentence out of four options. In the GS task, however, the student is given a sentence in her native language with one word underlined, and is then asked to match the underlined word with the best functionally equivalent word in a second sentence. As such, is GS the equivalent of analytic language ability? And more importantly, since the GS task relies on an
individual’s knowledge of parts of speech and their functions, is this ability affected by previous learning experiences?

Canner (2011) conducted a study that attempts to answer these questions. 35 participants were tested for both memory (including AM and WM) and analytic component abilities, which included GS and ALA. Participants ranging from having little or no experience learning a foreign language, to those with a high level of experience, were then surveyed to closely determine their experiences in language learning, including the duration, type, and intensity of the learning experience. Results of two different ANOVA’s testing for correlations between GS and either proficiency or formal learning experience as independent variables demonstrated a progressive pattern of relationship between the formal condition and grammatical sensitivity, with significant correlations between all three levels of experience and GS score, but not between GS and proficiency. Results also demonstrated no such relationship between GS and more naturalistic or incidental learning experiences.

The results provide support for the hypothesis that a learner’s grammatical sensitivity is directly influenced by language-learning experiences under formal conditions, which are commonly associated with what is called ‘focus-on-form’ instruction, defined as an instructional learning context in which the learner is consciously aware of specific forms that are targeted in the classroom for acquisition, with grammatical explanation of the form’s use and structure typically provided. The study also provided a multivariate analysis of all four components which showed separation between GS and both memory components, with no predictive interdependency between GS and either AM or WMC, although interdependency was
seen between WMC and analytic language ability. The findings appear to indicate that GS is a relatively independent ability whose development largely depends upon the acquisition of linguistic information. Other research appears to support this conclusion, such as Rota and Reiterer (2009), who found overall correspondence between GS and crystallized intelligence (IQ), while other studies, such as Sasaki (1996) appear to indicate a correlation between ALA and fluid intelligence.

GS appears to be closely associated with formal learning experiences and crystallized or learned types of intelligence, not more unconscious types of learning such as implicit or incidental learning conditions, a conclusion further substantiated in a study by Reves (1983), who found a lack of measured variance for the implicit condition associated with the Arabic version of the “Words in Sentences” (GS) task. Canner (2011) suggests a two-way, reciprocal relationship between GS and formal learning in that formal learning directly expands GS resources, while a learner’s acquired grammatical sensitivity might be used in acquisition under formal or explicit conditions. A learner who has already acquired such sensitivity during previous L2 learning experiences would naturally find learning an additional language in such a context even easier, since she would be able to readily access existing GS resources during acquisition. This might help explain the apparent increasing ease language learners often ascribe to the acquisition of additional languages following successful acquisition of an L2.

In light of findings that link ALA with Gf (e.g. Sasaki, 1996) and those showing a strong correlation between GS and past learning experiences (e.g. Canner, 2011), it appears that, contrary to the assertion made by Carroll (1981), ALA and GS are not equivalent. The confusion between the two abilities appears related to the fact that GS
would naturally depend upon a learner’s inductive ability, which would allow analytic extrapolations of observed or learned patterns of morpho-syntax to other contexts. However, an individual’s grammatical sensitivity appears to represent an ability that is built up over time and is predominantly acquired through exposure to language-learning experiences, while analytic language ability appears to be a more innate or natural ability that is typically relied upon by adult language learners (Harley and Hart, 2002; Skehan, 2002). An individual’s acquired GS resources would in turn naturally feed or bolster the more innate analytic abilities utilized during acquisition. It is evident that much remains to be learned about an adult acquirer’s analytic resources, and how they may also interact with memory abilities to have an effect on the attainment of proficiency.

**Phonetic coding, memory and fluid intelligence**

Unlike the analytic abilities, far more is known about the memory complex and the closely-related phonetic coding ability. Phonetic coding ability was originally described by Carroll as “an ability to identify distinct sounds, to form associations between these sounds and symbols representing them, and to retain these associations” (Carroll, 1981 p. 105). The retention of these associations occurs in the learner’s long term memory (LTM). Besides being needed for the learning of correct pronunciation, phonetic coding ability is also used to both associate orthography with pronunciation and learn new vocabulary words in a second language in the form of paired word associates, linking phonetic coding ability with functions of memory.

A number of studies have demonstrated links between phonetic coding ability, working memory (WM) and fluid intelligence. Rota and Reiterer (2009) devoted a study to discovering possible links between four types of cognitive abilities (empathic skills,
mental flexibility, working memory ability and intelligence) and phonetic coding ability, and a direct correlation was found between working memory capacity (WMC) and the ability to acquire phonological material. The results of their study establish a close correspondence between phonetic coding ability, WMC and the ability to attain proficiency in pronunciation in an L2. Such evidence raises a couple of interesting observations related to the nature of phonetic coding ability and its relationship with other elements of aptitude. One aspect of the ability is evident: it demonstrates strong correspondence with working memory capacity, with the results of a number of studies corroborating this connection. Of note are a lexical study conducted by Service (1992) and a study conducted by Jilka (2009a), which examined the links between pronunciation (phonetic coding ability) and cognitive traits related to WM, intelligence (both crystallized and verbal), overall linguistic aptitude and attributes such as personality and motivation. Service (1992) found that the acquisition of lexical items in a second language is based upon a learner’s capacity to store new phonological information in WM (Jilka, 2009a), while the results of Jilka’s (2009a) study also show a link between phonetic coding and WM, and seem to show that prior language knowledge or linguistic awareness are not as significant as “purely perceptual abilities” in phonetic coding ability, the results showing an insignificant correspondence between phonetic coding ability and crystallized intelligence (Jilka, 2009a).

These findings would indicate that phonetic coding aptitude and the memory functions related to it are generally impervious to the influence of environmental factors, such as experience in learning a foreign language. As such, phonetic coding ability may very well represent a rigid component of aptitude. This conclusion is further motivated
by the results of some studies that seem to point to a relationship between measured phonetic abilities in the L1 and acquired skills in an L2, whether the skill is learning new vocabulary (Adams and Gathercole, 2000) or other, related abilities in L2 acquisition (Jilka, 2009a). The results of such studies show that a learner’s phonetic coding ability in an L2 corresponds pretty consistently with the same abilities in the L1, a position also argued by Sparks, et al. (1998). Research also shows that memory components of aptitude are fairly rigid, explaining the correlation between L1 abilities to learn new vocabulary and those possessed in L2 acquisition (Jilka, 2009a). This view is further supported by experiments that were designed to test variability in the phonetic coding ability of different learners through neurological scans of the vital areas of the brain that govern language.

In a study conducted by Reiterer (2009), magnetic resonance imaging (MRI) was used to analyze the brain activity of learners who were acquiring various phonetic structures (similar results were found by Galestani, Molk, Dehaene, LeBihan and Pallier, 2007a). Results indicated a relationship between the participants’ brain white matter anatomy and IDs in phonetic coding ability, motivating the conclusion that innate differences in the morphology of the brain’s auditory cortex exist between individuals that lead to differences in the speed at which they learn phonetic structures (Reiterer, 2009). These findings are complemented by extensive research that has found strong relationships between phonetic coding ability, WMC and fluid intelligence, which helps to explain the apparent rigidity of this component of aptitude (see Miyake and Friedman, 1998; Daneman and Carpenter, 1980; Dörnyei and Kormos, 1998 and Baddeley, 2003).
The Memory Complex: The Importance of Working Memory in L2 Acquisition

The dominant view of memory espoused by early SLA researchers was limited to the learner’s ability to associate new orthographic or lexical items with equivalent items in the L1: memory was viewed as a strictly associative learning task. Such associative Memory (AM) ability, was measured in the only memory-related component of the MLAT, the Paired Associates subtest, which is limited to measuring STM with its reliance on chunking items for memory storage. While demonstrating direct correlations with success in learning a second or foreign language, Carroll himself states that the Paired Associates task has a lower predictive validity than some of the other subtests (Carroll, 1981). This was also demonstrated in research by Skehan (1982). With its emphasis on the learning of phonetic-lexical items (typically accomplished orally in language-learning), the AM task also appears to measure phonetic coding ability.

Though the Paired Associates subtest reflects the limited associational perspective of memory that was dominant among cognitive psychologists at the time of its development, it is perhaps this dual nature of the task that still provides its potency as a predictor of language learning success: since the measurement of AM on standard aptitude tests incorporates lexical-phonetic abilities, the correlations seen between this traditional component and success in L2 acquisition may relate predominantly to phonetic coding, rather than the STM aspect of the task. For example, Harrington and Sawyer (1992) found an absence of any strong correlations between STM for random words or numerals (without the learning of phonetic information) and L2 reading success, while greater capacities in working memory were found to be highly correlated with strong reading ability. The model of working memory (WM) that has replaced the older
associative view of memory represents a vastly improved understanding of the memory domain of abilities.

The interrelationships between phonetic and memory aspects of aptitude are described in the model for working memory developed by Baddeley and Hitch (1974) (see Baddeley, 1986; Gathercole and Baddeley, 1993). Their model has had a significant impact on SLA and cognitive psychology research as well as some behavioral aspects of psychology (Rota and Reiterer, 2009). Gathercole and Baddeley (1993) presented a definition of WM that describes it as a cognitive system devoted to the “temporary processing and storage of information” (Gathercole and Baddeley, 1993; see also Dewaele, 2002). The central tenet of Baddeley’s model is a dynamic relationship of memory maintenance with an ability to control attention while performing complex cognitive tasks. The model proposes that WMC is composed of four interrelated components: the Phonological Loop (PL), Visuospatial Sketchpad (VS), the Episodic Buffer (EB) and the Central Executive (CE).

The Phonological Loop, first characterized by Baddeley and Hitch (1974), is the main verbal aspect of WM that specializes in storing phonetic (sound) and verbal information, with the material being stored According to its phonetic characteristics. However, without a system of rehearsal, material in the PL can be quickly lost. According to Gathercole and Thorn (1998), the PL contains both a temporary repository of verbal material (the STM aspect of the PL) and a subvocal rehearsal system which serves to keep new information active so that it can be later stored in long-term memory. Visual data is also converted into phonological information in the PL. The Visuospatial Sketchpad is the visual counterpart to the phonological loop, and serves to coordinate and
assimilate kinesthetic, spatial and visual information into a form that is stored temporarily and utilized, as in simple reading tasks. The Episodic Buffer is the most recent addition to components of working memory. It functions as a storage facility that takes material from various sources and joins it as unified code (episode). It is assumed that the episodic buffer serves as a crucial basis for conscious awareness. The Central executive is considered to be the most important, though least understood component. It is vital in controlling and assigning attention capacities, governing the choice, initiation and completion of processing. For example, the CE acts as the system that tells the Episodic Buffer when it should join material together as code. Material from the VS, PL and long term memory (LTM) is coordinated and integrated in the CE, fulfilling cognitive processes related to making decisions, planning, mathematical calculations and other reasoning tasks, like reading and comprehension. The CE is vital to this study since its functions are believed to be the main variables that affect IDs in working memory span (Baddeley, 2003; Daneman and Carpenter, 1980).

The general capacity of a learner’s working memory is usually stated as the learner’s working memory span. As Rota and Reiterer (2009) point out, the first researcher to show that a quantifiable analysis of memory span was possible was Miller (1956), and the concept has expanded to include consideration of the lexical nature of items memorized (Hyme et al., 1995) and the recent concept of “chunking” introduced by Cowan (2001). Considerable evidence shows that WM span is an accurate predictor of a great variety of cognitive abilities. It has been shown to have high correlation with measures of reasoning ability that are associated with tasks requiring the simultaneous processing and storage of information, with such tasks commonly included in standard
intelligence tests (Dörnyei, 2005). WM span accurately predicts some vital language-related abilities, including language comprehension and vocabulary learning through incidental interaction, as well as reasoning abilities related to fluid intelligence, seen in various studies (e.g. Daneman and Carpenter, 1980; Kyllonen and Christal, 1990; Shute, 1991) (Kane and Engle, 2003).

Though often erroneously thought of as equivalent to short term memory (STM) (WM and STM show some correlation), it is evident that working memory is itself composed of two components, basic memory ability (STM) and another attention control component that according to Baddeley (2003) is part of the executive control function of WM. Many researchers now believe that it is not possible for STM to be responsible by itself for the high correlation between WM span and higher order cognition (fluid intelligence), demonstrated in a number of studies that have evaluated subjects with both WM and STM span tasks, while also testing for their fluid intelligence with nonverbal tasks, such as Engle, Tuholski, Laughlin and Conway (Engle, et al.) (1999); and Conway, Cowan, Bunting, Therriault and Minkoff (Conway, et al.) (2002). The difference between the two types of “span” tasks is that the STM tasks only require immediate recall of isolated item lists (words, numerals, or pictures), similar to the Associative Memory task found on the MLAT, while WM span tasks typically combine the memorization of items with a task that demands the added processing of material that distracts from the memorization task. Such design taps into the need for attention control.

**Working Memory: Correlations with Fluid Intelligence**

Unlike the STM word span task, WM span tasks show high covariance with measures of $G_f$, and a number of past studies appear to demonstrate that the correlation of
WM with fluid intelligence is primarily due to the controlled attention component of working memory, which is absent in STM tasks (Kane and Engle, 2003): e.g. Engle (2001, 2002), Engle et al. (1999), Kane et al. (1999), and Kane and Engle (2003). It appears that the component that correlates most highly with fluid intelligence is the attention control ability that allows the simultaneous processing of material in both the main memorization and interference tasks. The results of some recent studies appear to demonstrate that the source of the correlation of WM with fluid intelligence is this aspect of attention control, motivating the conclusion that WM span measures constitute an evaluation of “executive attention” (Engle, 2002). There is disagreement, however, regarding the extent to which WM correlates with fluid intelligence.

Ackerman, Beier and Boyle (2005) conducted a “meta-analysis” of the research literature related to working memory, and concluded that WMC is neither the equivalent of general (g), nor of fluid intelligence (Kane, Hambrick and Conway, 2005): They further concluded that WMC is for the most part a construct that is “domain-general,” and that it is far more intimately related to fluid intelligence than is STM, views that are widely supported by latent-variable research based on large samples of data. In a reanalysis of the data used by Ackerman et al. (2005), including over 3100 adult subjects, Kane, Hambrick and Conway (2005) drew similar conclusions, though they found the correlation of WMC with fluid intelligence to be far higher than the approximately 20% shared variance found by Ackerman et al. (2005). The results demonstrated a shared variance of over 50% between fluid intelligence and WMC. They argue that “WMC represents a distinct cognitive-ability construct that is strongly related to Gf (fluid intelligence) and novel reasoning…” and that attention control “is largely responsible for
the shared variance between WMC and $Gf^*$ (Kane et al., 2005, p. 69). In light of the commonality between the conclusions of the above studies, it appears evident that the memory complex of abilities, and specifically working memory and its associated functions, share important traits with the more rigid aspects of intelligence, i.e. fluid intelligence.

Of the various components of aptitude, working memory appears to possess the greatest resilience to modification from environmental factors such as education, similar to fluid intelligence. The construct’s considerable correlation with $Gf$ (and lack of correlation with crystallized intelligence) motivates further research that could serve to determine the extent to which aspects of WM are shared or related to the cognitive reasoning or analytic problem solving aspects of intelligence. In a study conducted by Sawyer and Ranta (2001) that analyzed the results of a number of studies, a direct correlation between WMC and L2 proficiency was found. Other studies have shown that WMC in the L1 corresponds with that measured for the L2 (Dörnyei, 2005). Such results appear to indicate that the strictly memory components of aptitude represented by working memory are fairly rigid. With its dynamic interaction between attention control and the ability to retain items in memory, WM represents a distinct improvement over the older and more limited view of memory represented by the associative memory tasks utilized in early aptitude tests such as the MLAT. The rigid characteristics of WM (i.e. it’s resilience to modification and relative stability between the L1 and L2) are significant to any examination of the construct’s interaction with various learning conditions and types of instruction to produce proficiency.
2.3 Models of Aptitude-Oriented Interactionist Research

Skehan (2002) has developed a “componential interactionist” research approach based on SLA process-aptitude interactions. He cites studies by Wesche (1981) and Reves (1982), who indicate a relationship between types of instruction and individual aptitude profiles (Skehan, 2002). Based on 3 main modules of L2 learning and their corresponding aptitude modules (i.e. input processing - phonetic coding ability; language processing - analytic ability; memory -memory abilities), Skehan (2002) argues for a processing stage aptitude interaction approach. He relates 9 SLA processing stages to aptitude components: 1. noticing: auditory segmentation, attention management, working memory, phonetic coding; 2. pattern identification: fast analysis/working memory, grammatical sensitivity; 3. extending: inductive language learning ability; 4. complexifying: grammatical sensitivity, inductive language learning ability; 5. integrating: restructuring capacity; 6. becoming accurate: automatization, proceduralization; 7. creating a repertoire: retrieval processes; 8. rule automatization, achieving fluency: automating, proceduralization; 9. lexicalizing, dual-coding: memory retrieval processes. This componential approach relates to evidence for a critical period (e.g. DeKeyser, 2000; Sasaki, 1996), and Skehan (2002) suggests that “foreign language learning abilities, while modular, are different in kind from the modules which exist in the first language case” (p. 83) (see also Dörnyei and Skehan, 2003; Skehan, 1998, 2002).

Another interactionist research paradigm has been developed by Robinson (2002) and is of particular interest to the current study. Robinson’s approach entails an analysis

---

1 The term “interactionist” is used by Skehan, Robinson and other researchers in this context to describe the type of research that emphasizes the relationships or “interaction” between aptitude complexes and learning conditions or processes to produce acquisition. For the purposes of this study, the term is limited to this narrow definition.
of the combinations of ID factors that have the greatest impact on success in acquisition in different learning contexts. The idea that certain “aptitude complexes”, or clusters of abilities, act to influence effective learning was first proposed by Snow (Snow, Korno and Jackson, 1996), and has been echoed by Ackerman (2003), and Corno, Cronbach, Kupermintz, Lohman, Mandinach and Porteus (2002) (Corno et al., 2002). Robinson (2002) has combined the concept of aptitude complexes with an approach developed by Cronbach that emphasizes the interaction of aptitude with “treatment” (Dörnyei, 2005). Robinson combined these two concepts to develop a theoretical framework for aptitude-related SLA research that he describes as “individual difference/learning condition interactions” (Robinson, 2002a, p. 114). He asserts that learner IDs related to cognitive abilities should be used to form individual profiles that can aid in determining appropriate forms of instruction and learning contexts, the first interactionist model of this type (Dörnyei, 2005).

Robinson (2002a) argues that an ‘aptitude-treatment’ research framework can offer much in way of furthering our understanding of a number of core issues in SLA theory, including the importance of explaining differences between child and adult language learning, explaining the acquisition processes that occur under explicit, implicit or incidental learning conditions, and the need to explain the high degree of variability seen in adult L2 learning success under various conditions or contexts of instruction. He states that “aptitude, awareness, and age are important learner variables, and any general

---

2 The use of the word “treatment” is equivalent to ‘learning condition’, an emphasis of Cronbach’s (Corno et al., 2002) ‘aptitude-treatment interaction’ approach, which relates to the interaction of aptitude abilities and learning or instructional conditions that results in learning.
theory of SLA will be incomplete without an explanation of how, and under what conditions, IDs in each impact upon learning.” (Robinson, 2001, p. 369)

Based upon recent scholarship related to various cognitive abilities and the need to connect them with L2 learning contexts (e.g. Harley and Hart, 2002; Ranta, 2002; Robinson, 2002b; Skehan, 2002), the main goal of Robinson’s (2001, 2002a) framework is to utilize individual learner IDs related to cognitive abilities to form individual profiles that can be used to determine appropriate forms of instruction or effective learning conditions. Robinson’s learner profiles are distinguished by the “cognitive resources” a learner possesses, as in attention and WMC, as well as the learner’s “primary abilities,” such as pattern recognition or processing speed (Dörnyei, 2005). He describes such resources in terms of the type of input environment learners are exposed to, including implicit, explicit (or formal) and naturalistic learning environments. The profiles are then matched to optimal learning conditions. Robinson’s model is therefore useful for drawing contrasts between different learning contexts and how they interact with profiles of learner resources to affect acquisition both in terms of its process and, of key interest to the present study, its product in the attainment of overall proficiency that arises from such interactions. It is evident that in order to determine which forms of instruction or learning conditions are most appropriate for given learner aptitude profiles, different learning contexts should be compared to one another in relation to a particular measure of success, such as the acquisition of certain forms or overall L2 performance in terms of level of attained proficiency.

Robinson’s framework fits well with the objectives of the present study. His model provides the explanatory potential and concepts necessary for discovering the
correlation of working memory capacity with proficiency under different learning or instructional contexts. The approach’s pedagogical implications also address the desire to apply what is learned in the current study to pedagogical practices. The goal of the present study is to use this framework to research correlations and interactions between a specific measure of aptitude (working memory) and an evaluation of attained oral proficiency in L2 Russian (L1 English) under four different learning conditions: a) formal instruction alone; b) naturalistic language learning alone; c) formal instruction followed by naturalistic learning; d) naturalistic learning followed by formal instruction. The present study could therefore provide a basis for future research related to the interaction of other key aspects of learner aptitude with proficiency under different learning conditions, such as analytic language ability (ALA). This type of research could also be applied to studies of L2 acquisition in other target languages.

2.4 Working memory and Its Role in the Attainment of L2 Proficiency

Research findings that show direct correlations between working memory and the attainment of proficiency (e.g. Harrington and Sawyer, 1992; Sawyer and Ranta, 2001) make the study of its exact role in L2 proficiency of critical importance. A major obstacle to such research, however, is the complex and multifaceted nature of the ability. Yet the working memory complex has been found to play a role in many different aspects of language acquisition, from the acquisition of vocabulary and meaning (e.g. Service, 1992) to the development of proper pronunciation. Indeed, some researchers have concluded that working memory represents “one (if not the) central component of language aptitude” (Miyake and Friedman, 1998, p. 340). The construct appears to be relied upon for the accomplishment of tasks ranging from analysis of grammatical
information in language learning, such as the comprehension and parsing of complex syntactic or morphological information, as found by Miyake, Friedman, and M. Osaka (1998) (Miyake, et al.), to logical problem solving abilities (e.g. Kyllonen and Stephens, 1990), and even computer programming (e.g. Shute, 1991). The pivotal component of working memory that is largely involved in language learning is the Phonological Loop and its two subcomponents which are designed to temporarily store and rehearse new information. This phonological working memory system is apparently involved in the acquisition and retention of lexical, syntactic and phonological information.

**Phonological Working Memory: A Phonetic and Lexical Interface**

The close association between phonetic coding ability and working memory naturally indicates that those who possess strong abilities in phonological WM would have a greater chance of success in overall attainment of native-like pronunciation. It has even been suggested by Atkins and Baddeley (1998) that the Phonological Loop component of WM is the structure that facilitates all language acquisition (Rota and Reiterer, 2009). The nature and role of phonological working memory in acquisition has been studied extensively, with such research indeed indicating the relationship between this phonological component of WM and the processing of new linguistic information, while also drawing connections between the role of WM in the L1 and the L2. Adams and Gathercole (2000) found that high WM pre-critical period learners display considerable abilities for the repetition of foreign words, while also showing higher L1 abilities than those whose WM is lower. It is interesting to note that variation in WM between the L1 and L2 has been linked with cross-linguistic phonological differences: Cheung, Kemper, and Leung (2000) found that among Chinese-English bilinguals cross-
language variation in working memory was largely attributed to differences in syllable structure between the two languages as phonological information is encoded in working memory. Research into phonological WM has also demonstrated a strong connection between the acquisition of phonetic information and the learning of vocabulary items.

Service (1992) concluded that the successful learning of new lexical items by adult L2 learners is based upon strong abilities in phonological working memory. Other researchers have likewise concluded that the PL must possess connections with lexical knowledge, though it appears that phonological memory is stronger for items that contain familiar sound combinations (e.g. Brown and Hulme, 1992; Gathercole, 1995; Roodenrys, Hulme, and Brown, 1993), indicating that previously acquired phonological architecture strengthens an individual’s available WM resources as well as the ability to discern speech patterns (Gathercole and Thorn, 1998). It is therefore apparent that phonological WM has some kind of interface with lexical knowledge, which might help to explain the (albeit lower) predictive validity of traditional associative memory lexical tasks, such as the paired associates subtest on the MLAT.

Such linkage between lexical information and WM has also been seen in child L1 acquisition studies, which test children’s phonological loop capacity with either a digit span task (the greatest span of digits a child can remember in the exact order heard) or a non-word repetition task, in which the child tries to immediately repeat an unknown phonological item that is uttered. In a longitudinal study designed to discover how pivotal a role phonological memory plays in the growth of vocabulary knowledge during childhood, Gathercole and Baddeley (1993) tested 80 children at ages four, five, six and eight on a number of measures, such as non-verbal intelligence, reading skills, perception
of vocabulary and a non-word phonological STM task that measured the ability to recall and repeat nonsense words. Their results showed the development of a causal relationship between phonological STM and lexical knowledge at early ages, but that the relationship changes once previously learned lexical information begins to play a role in the learning of subsequent, more complex, lexical items.

The explanation offered suggests that children who possess strong phonological STM “produce phonological memory traces that are highly discriminable and persistent”, and that such a trace “will become durable and link semantically with its referent” (Gathercole and Baddeley, 1993, p. 51). They conclude that strong phonological memory acts to alleviate the difficulties associated with vocabulary acquisition. A number of other studies have also demonstrated the development of phonological WM as children age, with particularly strong correlations between lexical knowledge and Phonological Loop capacity during childhood (e.g. Gathercole and Baddeley, 1989; Gathercole, Hitch, Service, and Martin, 1997; Gathercole, Willis, Emslie and Baddeley, 1991, 1992; Michas and Henry, 1994). Such child acquisition research is mirrored by studies of adult acquisition and the role of phonological WM.

Phonological WM has been directly implicated in the acquisition of lexical items among adult L2 learners as well. Speciale, Ellis, and Bywater (2004) analyzed the role of phonological STM in the acquisition of vocabulary among college students. The results demonstrated that phonological STM combines with the acquisition of phonetic sequences to promote learning, and that the extent to which the two combine in learning grows as students gain proficiency, based on their growing ability to identify regular phonological patterns related to previously learned lexical information (Speciale et al.,
In a study of Chinese high school students, Cheung (1996) reported results indicating that measures of non-word span best predict the acquisition of L2 vocabulary for below average students, while not for learners who possess higher ability to learn new words. Cheung (1996) proposes that there is interplay between phonological STM and phonological information stored in long-term memory to promote the acquisition of new items.

Papagno Valentine, and Baddeley (1991) conducted a number of successive experiments to examine the adult acquisition of new vocabulary while the articulation of items is simultaneously suppressed. The model of working memory (Baddeley, 1986; Baddeley and Hitch, 1974) predicts that suppression interferes with the phonological representation of encoded visual information, and therefore prevents the ability of the PL to mediate visual information, while the suppression of articulation prevents rehearsal of new material, and thus interferes with its retention (Gathercole and Baddeley, 1993). In a related sequence of experiments, Papagno et al. (1991) tested Italian adult learners’ ability to acquire foreign words (Russian) as compared to more familiar items in the native language. They reported results that support the WM model’s predictions: suppression of articulation interfered with the retention of foreign words to a much higher degree than the ability to recall previously known words, showing that the acquisition of foreign words involves the phonological loop to a much higher degree that its use for items whose semantic information has already been stored in memory (Gathercole and Baddeley, 1993).

Interestingly, one of the experiments conducted by Papagno et al. (1991) found that the phonological loop can evidently be bypassed in lexical acquisition when learners
employ other strategies, such as semantic associations with previously learned items. In an experiment similar to the above study, only with English speakers, participants effectively prevented suppression by drawing such associations between the new Russian item and the English equivalent, such as ‘throat’ – ‘gargle’ – ‘gorlo’ (Gathercole and Baddeley, 1993). The results of the study of Italian L2 learners mentioned above were duplicated in an English – Finnish experiment that minimized the possibility of such semantic associations (Papagno et al., 1991). These studies point to phonological working memory’s direct role in the acquisition of new lexical items in L2 learning. Gathercole and Baddeley (1993) have concluded that as language speakers age, the strong role of phonological working memory in the learning of vocabulary of early childhood is transformed into “a more complex interaction between vocabulary knowledge and phonological memory” (p. 67). Such an interaction raises an important question: to what extent does the Central Executive (CE) play a role in this developing lexical processing system? A number of studies have indeed confirmed that a complex relationship exists between the PL and CE in adult vocabulary learning and retention.

Daneman and Green (1986) utilized Daneman and Carpenter’s (1980) reading span test of WMC to analyze the role of attention control (a function of the Central Executive) for the simultaneous processing and retention of new lexical material. They reported strong positive correlations between the participants’ reading span scores and the precision with which they defined the new items. They concluded that individual learners rely on WM abilities to interpret and utilize cues within the speech context to comprehend new vocabulary. These findings demonstrate the importance of the executive control functions of working memory in lexical and semantic processing. One
important question that relates to the impact of WM on the attainment of proficiency is the role, if any, played by the coordination of the Phonological Loop and Central Executive (the essence of the working memory system) in the learner’s production of speech.

Though this particular area of SLA research has gained momentum in the last twenty years, compared with the study of lexical processing or comprehension, research related to speech production has been severely lacking due to the complexities associated with an assessment of the processes and product of L2 oral production (Mota, 2003). This deficiency is also seen in working memory studies. A number of studies, however, have begun to illuminate the potential role of WM in oral performance. According to an early analysis conducted by Gathercole and Baddeley (1993), the subvocal rehearsal aspect of the Phonological Loop is one of the mechanisms involved in the high-order planning of speech, but not in its active production, a conclusion that has found support in studies that show no significant correlation between spontaneous speech and phonological WM (e.g. Klapp, Greim, and Marshburn, 1981; Sternberg, Monsell, Knoll and Wright, 1978). Early research did, however, provide some evidence of a role of WM, and the Central Executive in particular, in semantic processing, such as evidence that appears to show that the CE may at least be partially involved in the use of semantic material in speech production (e.g. Powers, 1985). However, apart from studies that have elucidated the role of the CE in the WM construct (e.g. Unsworth and Spillers, 2010), information about the exact role of the CE in speech production is sketchy at best, and early conclusions that phonological WM plays no role in production have been challenged.
Daneman (1991) did find a correlation between fluency and WMC at the level of discourse and articulation in the L1. The question is whether such a link between WMC and fluency or other areas of proficiency could be established for L2 learners. In a study of the benefits associated with the rehearsal of L2 information, Ellis and Sinclair (1996) demonstrate that rehearsal (an ability associated with the phonological loop) has a positive impact on a number of areas, including the comprehension and acquisition of new language information, the learner’s metalinguistic understanding of grammatical structure, skills in pronunciation, and accuracy and fluency in speech production. O’Brien, Segalowitz, Collenteen and Freed (2006) demonstrated differences between high-proficiency and lower-proficiency learners in the characteristics of the associations between phonological STM and speech production: higher proficiency learners show stronger abilities in phonological memory than others.

In a study of speech production, Fortkamp (1999) conducted a replication study of Daneman (1991) to investigate whether WM has a role in fluency and accuracy in oral performance. Similar to the results reported by Daneman (1991), the findings of this study showed that learners who possess greater WMC had also attained a higher level of oral fluency in both discourse and articulation. Fortkamp (2000) conducted a further study of the relationship, and the results motivated the conclusion that “there is a relationship between learners’ working memory capacity and fluency, accuracy, complexity, and lexical density in L2 oral performance” (p. 41). One difficulty in drawing any overall conclusions about a direct role of WM in speech production, however, is an absence of significant correlations between WMC and oral production tasks in some studies (e.g. Mota-Fortkamp and Bergsleithner, 2007). Yet due to
significant correlations found between WMC and grammatical accuracy in production for a delayed oral production task, Mota-Fortkamp and Bergsleithner (2007) nonetheless concluded that phonological WM plays a role in speech production, albeit one that is potentially indirect.

Though research has been lacking in this area, several studies implicate a link between higher working memory spans and proficiency in production. For example, Fortkamp and Bergsleithner (2007) found that individuals with lower WMC also produced a greater number of errors in accuracy than those who score higher on the working memory speaking span task devised by Daneman (1991). Fortkamp (2000) has suggested that a minimum of four overarching analytical operations required for L2 speech performance make demands on an individual’s working memory, including 1) the stimulation of L1 and L2 knowledge; 2) the restraint of L1 material that is useless to the task; 3) the identification and retrieval of relevant L2 material; 4) the evaluation of production to prevent or correct potential errors.

It is clear that the extensive body of research demonstrating the significance of working memory in the acquisition of phonological or lexical information has greatly enhanced our understanding of the pivotal role it plays in L2 acquisition. A growing number of studies appear to indicate at least an indirect role of WM in speech production. Another area of research that has proven invaluable in promoting our understanding of the role of this important cognitive resource in the attainment of proficiency is represented by studies that examine the neurological functions of the human brain during language learning. Such research utilizes neuroimaging during tasks focused on the acquisition of new phonological information.
Using such technology, a number of studies have demonstrated that phonological processing tasks that utilize WM are concentrated in the left temporoparietal areas (e.g. Paulesu, Frith and Frackowiak, 1993; Vallar et al., 1997; Warrington et al., 1971). Fuster (2003) was able to demonstrate a separation of working memory subcomponents, with phonological processing seen in areas different from those associated with other primary functions, such as those related to visuo-spatial processes. Some neurological studies of phonological working memory have examined its role in relation to level of success in L2 acquisition. For example, Chee, Soon, Lee and Pallier (2004) examined the role of phonological working memory among bilinguals who were either equally strong in both languages or not. They found distinct differences in the activation of certain areas of the left side of the brain between the two groups: their results appear to demonstrate that more extensive use of phonological working memory was associated with the higher proficiency group. Like Chee et al. (2004), other studies have similarly found a connection between high levels of L2 success or language ‘talent’ and the left inferior parietal area (e.g. Golestani and Zatorre, 2004; Golestani, Alario, Meriaux, Bihan, Dehaene and Pallier, 2006). The authors of these studies conjecture that a learner’s brain anatomy could itself be used to predict level of success or proficiency in an L2.

A number of studies have demonstrated the important role of working memory in L2 proficiency. Reliance upon working memory resources has been shown to be directly implicated in overall success in L2 learning, including the successful acquisition and retention of new vocabulary (e.g. Atkins and Baddeley, 1998; Baddeley, 2003; Baddeley, Papagno and Vallentine, 1991; Papagno and Vallar, 1992), overall level of proficiency, especially in terms of fluency (e.g. Rota and Reiterer, 2009), and the successful
acquisition of new morpho-syntactic information (e.g. Miyake, et al., 1998), which would also have an impact on proficiency in terms of accuracy and complexity. Rota and Reiterer (2009) investigated the interaction between phonological working memory and L2 learner abilities to quickly acquire, retain and accurately produce new lexical items. They found a direct correlation between proficiency and two different measures of the rehearsal characteristics of working memory, a digit span task and a word span test. Based upon previous findings in the literature, they hypothesized that those who excel in L2 lexical pronunciation would possess high level rehearsal abilities. Their results showed a direct correlation between proficiency in the L2 (in terms of pronunciation) and enhanced working memory abilities. What is perhaps least understood, however, is the role working memory may have in more analytic language learning abilities.

Some research has explored the role of WM in the attainment of proficiency in terms of the processing of morpho-syntactic information. Miyake, et al. (1998) examined the causal relationships between the four dependent variables of L1 working memory, L2 working memory, cue preference, and syntactic comprehension. By using a type of structural equation modeling called path analysis, they tested several models of causal relationships between the different variables, and found that the model that best described the data indicates a direct effect of L2 WM on cue preference, with these two factors acting together to directly impact syntactic comprehension. An indirect influence of L1 WM on both cue preference and syntactic comprehension was also found. They conclude that word order is an important cue that places a proportionately high requirement on working memory resources.
The above-mentioned studies point out the importance of the construct of working memory in relation to successful acquisition of an L2; among the great array of factors that influence acquisition, working memory appears to play a highly significant role in the attainment of proficiency in a second or foreign language, and may very well represent the overriding component of learner aptitude as Miyake and Friedman (1998) suggest. Studies also show that WM plays this role irrespective of a learner’s general intelligence (as tested with standard IQ measures) or the presence of inabilities or impairments in articulation or hearing (e.g. Gathercole and Baddeley, 1990). Research also demonstrates correlations between WM and both fluency and accuracy (e.g. Daneman, 1991; Fortkamp, 2000; Mota-Fortkamp and Bergsleithner, 2007). Such research motivates further examination of how working memory may interact with various forms of instruction or other learning conditions to influence the successful development of the major elements of L2 proficiency, including accuracy, complexity, and fluency. Research related to working memory will need to determine whether the ability plays a differential role in the development of different linguistic subsystems, while the exact nature or extent of the role of working memory under different learning conditions and the processes associated with them, such as the semantic processing nature of the naturalistic condition, remains to be determined. The present study therefore has the potential to fill an important gap in our understanding of the role of this key component of learner aptitude in the successful acquisition of an L2 in general and of L2 Russian in particular.
2.5 Challenges Related to the Acquisition of L2 Russian

Adult learners of Russian are faced with a great number of challenges in the attainment of proficiency. Learners face significant difficulties on a number of different levels, including the acquisition of verbal aspect or the system of motion verbs that so commonly plagues the adult learner. In terms of the key components of proficiency, morphological accuracy and syntactic complexity, among the most challenging obstacles to native-like proficiency are the difficulties associated with Russian morphology seen in its highly developed inflections for case, gender and number. For a learner to attain native-like accuracy in Russian relative clauses such as those that are targeted in the present study, mastery of all case inflections is necessary, along with the proper use of prepositions. Such morphological difficulties are further compounded by highly variable patterns of syntax, with some constructions showing remarkable flexibility in word order, and others great rigidity. Certain forms and constructions are both widespread in their common usage and are highly representative of the morphological and syntactic complexities of the language, presenting learners with some of the most challenging obstacles to the attainment of native-like proficiency. Though the literature related to such structures is limited, this category includes various anaphoric constructions such as reflexive pronouns, the various relative clause constructions, such as those targeted for elicitation in the present study, and other constructions such as the very common subject-less dative constructions, or the conditional complementizer chtoby, ‘so that, in order that’, the accurate use of which requires learning both specific patterns of syntax and verbal past tense morphology.
The Russian Subjunctive Complementizer *chtoby* and Verbal Morphology/Aspect

The literature related to the Russian subjunctive mood appears to be limited to descriptive studies related to Russian conditional constructions, while several studies have explored child acquisition of conditionals related to the subjunctive particle *by*, including the inseparable form *chtoby*, which is targeted for elicitation in the present study. However, there is one aspect of adult acquisition research that is directly related to the acquisition of Russian conditionals. Since Russian conditional expressions with the particle *by* require use of the l-participle form of the verb, which functions in modern Russian as the past tense, the adult acquisition of past tense morphology is of interest to the present study. Due to their close association and the tremendous difficulties that come with the acquisition of verbal aspect, tense-related acquisition research in L2 Russian also strongly relates to the aspectual distinctions of the Russian verb system.

Although a fairly healthy body of descriptive literature concerning Russian verbal aspect and tense does exist, tense/aspect research specifically related to L2 Russian acquisition is limited to just a few studies, and there is much that we still do not know about the adult acquisition of tense and aspect in the language. Though not directly related to adult acquisition research, the study of child L1 Russian acquisition of tense and aspect does have some bearing on the subject. The earliest work on the acquisition of Russian aspect was conducted by Gvozdev (1961), who meticulously examined the speech of his son, Zhenya, and found that children appear to completely acquire verbal aspect from an early age. Consequent research appears to show that children learn both Russian aspects at the same time. Another interesting finding in some research is that tense and aspect seem to be acquired simultaneously (e.g. Gagarina, 2000; Kievzak-
Mandera, 2000), with similar findings reported for L1 Polish (e.g. Weist, Wysocka, Witkowska-Stadnik, Buczowska, and Konieczna, 1984). Such findings contradict the aspect-before-tense hypothesis (Antinucci and Miller, 1976): the imperfective is more common in child speech, though past tense has not been found to be mostly limited to the perfective as the hypothesis would predict. The opposite appears to be true, however, for adult learners.

One of the dominant theories in adult L2 tense/aspect acquisition research is the Aspect Hypothesis (AH) (Andersen and Shiraj, 1994; Bardovi-Harlig, 2000). This hypothesis states that the use of tense/aspect markers enables L2 learners to detect lexical aspect: perfective-past marking is typically connected to telic verbs and imperfective marking most often to atelic verb types. The cross-linguistic literature on the subject reveals a number of studies that seem to provide evidence for the AH, such as Bardovi-Harlig and Reynolds (1995), Cadierno (2000), or Collins (2002). Other researchers have suggested what is called the Default Past Tense Hypothesis (DPTH) (e.g. Salaberry, 1999; 2002), and argue that during early development of the IL learners tend to apply a default form of the past tense in all aspectual contexts (e.g. the perfective in Russian). The DPTH argues that learners only begin to conform to the Aspect Hypothesis when they develop their proficiency to the point that they are able to account for lexical aspect as they utilize verbal tense/aspect marking in speech. Such a barrier would have a direct effect on proper use of the conditional complementizer chtoby, since an inaccurate command of aspect impacts the meaning of subjunctive verbal complements.

Though the Russian-related literature is sparse, several studies provide support for the AH in relation to adult L2 Russian. Slabakova (2005) used a cloze-style
interpretation task in which participants evaluated different potential interpretations of written sentences. The results indicated that regardless of the level of proficiency, acquirers could grasp the fact that telic verbs (indicating results) correspond best with verbs marked for the perfective, which adheres to the AH prediction that perfective marking would be most associated with telic type verbs. Another example is provided by Nossalik (2008), who used grammaticality judgments to determine the learners’ interpretations. The study found that, as predicted by the AH, Russian perfective verbs are interpreted as not being compatible with verbs indicating duration. Though they were not designed to test the AH per se, these two studies appear to provide some supporting evidence for the Aspect Hypothesis since it was found that the perfective corresponds best with the use of telic verbs.

In a study designed to directly test the Aspect Hypothesis in regards to L2 Russian, Leary (2000) examined 40 subjects who represented four distinct levels of oral proficiency. Participants were asked to produce narratives in writing after viewing a silent film. The results indicated that the lexical aspect of the verbs used by the learners drove the choice of grammatical aspect in the exercise since use of the imperfective was associated with conditions or states, whereas the perfective was used to describe the results of the action or accomplishments. Leary (2000) concluded that L2 learners of Russian beyond the second year choose the imperfective or perfective predicated upon the verbs’ lexical aspect category. The study also showed that early Russian L2 learners (level one) had very low comprehension of the past tense, with none of the three level one learners in the study choosing any past tense forms. The results of this study are interpreted as directly supporting the Aspect Hypothesis.
The DPTH was tested by Martelle (2011), who appears to find support for the idea of a differential acquisition of the tense/aspect system between early and more proficient stages of adult L2 development. She examined the extent to which L2 Russian learners representing different proficiency levels actually conform to the Aspect Hypothesis, while investigating the extent to which the AH holds under certain conditions. Using both elicited speech data and proficiency data that was examined for tense/aspect marking and lexical aspect, her results indicated supporting evidence for the AH especially with the elicited oral data. The results also appear to demonstrate that learners prefer the imperfective for past tense contexts at the earlier stages of proficiency, which is interpreted as lending credence to the default past tense hypothesis (Martelle, 2011). Although they are few in number, the above studies help to elucidate some important aspects of the adult L2 acquisition of Russian verbal tense/aspect. It is interesting to note that such research also demonstrates some distinct differences between child and adult patterns in the acquisition of verbal aspect. While L1 Russian children appear to acquire both the imperfective and perfective aspects simultaneously with tense (Contra the tense-before-aspect hypothesis), adult learners appear to begin with a particular lexical aspectual category (e.g. imperfective) and then use it as a platform for consequent acquisition of other categories. Such findings help to explain the common observation among Russian FL teachers that adult learners appear to favor one aspect over the other in the past tense, which has a direct impact on their use of Russian conditionals in verbal complements, causing miscommunication.
Difficulties in the Acquisition of L2 Russian Morphosyntax

Russian L2 acquisition research related to morphology is dominated by verb morphology studies connected with morphological processing, while some research has been conducted on the acquisition of case morphology, which is relevant to the present study due to the characteristics of the relative pronoun structures that were targeted for elicitation, i.e. the to, chto ‘that which’ and kotoryj ‘who, that, which’ constructions. It appears that the majority of L2 Russian research on this topic is associated with connectionist or functionalist perspectives of SLA, with studies commonly emphasizing frequency, probability, and cue preference effects on the acquisition of verbal or case paradigms.

An example of a case-marking study is seen in research conducted by Kempe and MacWhinney (1998), who examine the acquisition of case morphology by adult L1 English learners of L2 German and Russian in order to compare two models of learning under the Connectionist theoretical paradigm: the associative acquisition model infers that acquisition of case is governed by the cue strength of particular inflections, while the rule-based model suggests that acquisition depends on the intricacy of the inflectional paradigm. One finding of this study was that the English L1 learners of Russian acquired case-marking faster than German L2 learners even though the German L2 learners were exposed to a greater amount of language input. The results of analysis appear to indicate that Russian L2 learners tend to rely on case marking far earlier than L2 German learners, while the German L2 learners use animacy in order to augment the considerably lower strength cue provided by German case-marking. The second part of the study was devoted to developing a connectionist acquisition model that simulated the results of the
experiment fairly accurately. Kempe and MacWhinney conclude that although some of
the results were mixed for the model simulation, adult L2 acquisition is associative in
nature and is predominantly governed by cues provided in the input.

**Russian L2 Acquisition of Relative Clause Structure**

The limited literature on the L2 acquisition of morpho-syntactic properties in
Russian appears to be largely devoted to specific lexical elements and their syntactic
relationships within the sentence, such as studies on the government and binding
characteristics of reflexives or the acquisition of relative clauses and their associated parts
of speech, including the various relative pronouns commonly used in the language. Of
the types of Russian acquisition research discussed in this review, studies of the
acquisition of these particular characteristics of the language are most applicable to the
current study’s emphasis on accuracy in morpho-syntax as a major component of oral
proficiency.

English speaking L2 learners of Russian face considerable difficulties in acquiring
the structure of Russian relative clauses. Relative clauses have played a significant role
in syntax-related research since they provide an excellent case of “long-distance
dependencies ... (that) have two crucial characteristics: first, the expressions filling the
head and tail points of the dependency differ in their articulation; second, the positions
are separated by a number of unrelated segments” (Polinsky, 2011, p. 4). The idea of a
universal hierarchy relevant to types of relative clauses was established by Keenan and
Comrie (1977), who describe two major types of languages: those with ‘right branching’
relative clauses place the clause following the noun phrase (NP) it modifies, while
languages that are left branching have pre-nominal relative clauses. Like English,
Russian is a right-branching language and relative clauses occur post-nominally.

Although the basic structure is similar between the two languages, the lexical features and morpho-syntactic characteristics of Russian make the acquisition of Russian relative clauses a thorny issue for the second or foreign language learner.

In Russian any interrogative word may be used to introduce a relative clause, including any pronoun, adjective or adverb. Examples of relative pronouns include kotoryj ‘who, that, which’, the interrogative pronoun chto ‘what’, the adjective kakoj ‘which’ and any adverb that is used as a relative adverb. Although Russian is typically described as an SVO language, the language has a considerably elaborate inflectional system that is used to indicate the grammatical function of virtually all verbal complements, while permitting great freedom in word order. The extremely free word-order structure in Russian allows any arrangement of the major constituents of a clause. Such syntactic freedom makes Russian a particularly good candidate language for syntax-related research of relative clause structure (Polinsky, 2011). The relative clause (RC) research literature includes cross-linguistic descriptive studies on the structure of relative clauses in a number of different languages, including Russian, several studies devoted to comparisons between relative clause structure among heritage speakers and child and adult L1 Russian speakers, and studies related to relative clause acquisition among adult L2 learners of Russian.

One type of study that has helped to illuminate the issues surrounding the acquisition and comprehension of relative clauses tests the interpretation of relative clauses among heritage speakers of Russian living in English-speaking countries. Polinsky (2008, 2011) examined the comprehension of relative clause structure on the
part of heritage speakers of Russian in the United States. Such research can help to inform our understanding of adult L2 acquisition of relative clauses since English is the dominant language for these heritage speakers of Russian; such research can therefore help to illuminate the role of transfer from the dominant language to the target structure.

The main objective of the first study (2008) was to identify any disparities between heritage speaker interpretations of subject relative clauses and object clause types, while also testing for the role of both frequency and English transfer effects in their comprehension. Polinsky predicted that there would not be any influence of incomplete acquisition and that heritage speaking adults would show the acquisition of forms on par with adult monolingual speakers. She also predicted that since the heritage speaker is English-dominant, inflectional cues related to case would be misinterpreted when there is no match between the word orders of English and Russian relative clauses, with such conditions promoting transfer from English to Russian. Due to an influence of English relative clause structure, Russian object relative clauses would be interpreted as subject relative clauses by heritage speakers.

The results demonstrated that the adult heritage speakers performed in ways significantly different from both the monolingual adults and the child heritage speakers. Though they displayed considerable accuracy in their interpretation of subject relative clauses, their performance with object relative clauses was unpredictable, and this unstable performance with object clauses was consistent throughout different relative clause word orders. Similar results are reported in a follow-up study (Polinsky, 2011) which was likewise designed to examine adult heritage speaker knowledge of relative clause structure. In the second study, Polinsky attempts to separate the impact of two
phenomena as potential causes of the differences seen in relative clause interpretation by adult heritage speakers: incomplete acquisition and attrition. The results of this study show that when it comes to relative clause interpretation, incomplete acquisition of the structures can be ruled out as a cause of deficiency in relative clause comprehension, whereas attrition could not be eliminated. She concludes that within the area of relativization, adult heritage speakers undergo attrition. This leads to a further conclusion that incomplete acquisition does not affect all areas of the grammar. The study also demonstrates that research of this kind can differentiate between the influence of language transfer and attrition effects. Examining such effects on the processing of Russian relative clauses has also been the focus of adult L2 acquisition research.

Polinsky (2011) reports results of a study conducted by Levy, Fedorenko and Gibson (2011) that represents perhaps the first research on online comprehension of Russian relative clauses and attempts to differentiate predictions made by different syntax theories that describe common difficulties in processing relative clauses in L2 acquisition. This study included four reading experiments on L2 learner comprehension of relative clauses in Russian, and also provided a comparison of the experimental results with relevant corpus data. The analysis involved two types of relative pronouns, *kotoryj* ‘which, who’ and *chtò* ‘that, which’. In contradiction to the anticipated outcomes that would be expected to follow from the corpus data, the study found that patterns of L2 learner comprehension of Russian relative clauses provide evidence for either of two competing theories that attempt to explain processing difficulties: the data showed that memory-related theory appears to explain deficient reading times in verbal relative-clause contexts, while expectation-related theory (e.g. word-order frequency) explains
processing difficulties in accusative NP relative clauses that begin with a relative pronoun. These findings are interpreted to suggest that the intuitions provided by both memory-related and expectation-oriented theoretical frameworks can be integrated to formulate a comprehensive theory of syntactic complexity.

Some research associated with the acquisition of relative clauses in L2 Russian has attempted to examine the effects of learning context on their acquisition. As with most of the research literature mentioned above, research related to the influence of learning conditions or context upon the acquisition of Russian is severely lacking. For example, Dunn (2007) states that at the time she conducted her own research, while instruction-related acquisition research had been connected with “ESL and some other languages (Chinese, Japanese, Persian, German, French, and Italian)…No research regarding relative clause acquisition and the effect of instruction for Russian as a second language was found.” (p.2) Dunn (2007) investigated the effects of instruction on the acquisition of Russian relative clauses, and specifically attempted to ascertain whether learners of L2 instructed Russian are able to generalize the acquisition of marked relative clauses to those that are unmarked as well as their ability to generalize instruction-based acquisition of unmarked relative clauses to those that are marked. The research specifically targeted relative clauses with the relative pronoun kotoryj ‘that, which, who’ and describes findings in light of the predictions made by the Accessibility Hierarchy (Keenan and Comrie, 1977), which predicts that unmarked features are acquired prior to marked features.

The experiment included 54 college-age participants who took a pretest followed by three days of instruction on relative clauses, which was in turn followed by a posttest.
Both tests included a combination task in which participants were given two sentences and tasked with combining them into one using the relative pronoun, but without losing any information, and a grammaticality judgment task that included 30 incorrect uses of the relative clause divided into three types of errors, and ten correct examples. Subjects were divided into two groups, one that received instruction only for less-marked relative clauses, and one only for marked relative clauses. She reports results that show that all subjects improved in their performance with all types of relative clauses, with greatest improvements made with the types of relative clauses for which they were trained. Those who received training with more marked relative clauses were also able to generalize the information to less-marked contexts, while those who were trained with less-marked forms were able to generalize their knowledge to the more-marked relative clauses. In agreement with findings by Croteau (1995) for L2 Italian, Dunn concludes that learners who are instructed in a particular area will make greater progress in that specific area, while the ability to generalize acquired knowledge of relative clause structure from marked to unmarked or unmarked to marked operates in a multidirectional fashion.

The positive impact of instruction on the acquisition of Russian relative clause structure found in the above study further substantiates similar conclusions seen in research for other target languages, such as ESL from a variety of different L1 backgrounds (e.g. Ammar and Lightbrown, 2005; Doughty, 1991; Eckman, Bell and Neslon, 1988; Gass, 1992; Hamilton, 1994), L2 Italian (e.g. Croteau, 1995), and L2 Japanese (e.g. Yabuki-Soh, 2007). The effects of instruction on L2 Russian acquisition have been examined in relation to other aspects of SLA, such as the attainment of proficiency.
2.6 Challenges in Acquiring L2 Russian Proficiency

Past research has specifically examined attained proficiency in Russian, including studies that are conducted to inform the development of new college textbooks (e.g. Pavlenko, 2006), proficiency test design research aimed at providing the US government Interagency Language Roundtable (ILR) with more linguistically appropriate (i.e. lexical, morphological, syntactic or phonological) inventories of areas corresponding to proficiency-related categories (e.g. Long and Gor, 2008). Of particular interest to the current study, and therefore this review, however, are studies that examine levels of attained proficiency following traditional formal instruction in a college curriculum, special immersion programs conducted in the US, and research designed to examine attained proficiency following study abroad or other in-country immersion programs.

In research designed to evaluate gains in proficiency resulting from a study abroad experience, Brecht, Davidson, and Ginsberg (1993) assessed the proficiency of adult students of Russian both before and after the study abroad in Russia. Participants had between two and three years of Russian in a university program prior to the experience. Before the study abroad over 57% of the students had achieved a mid-intermediate oral proficiency level, and over 20% of them tested at the high-intermediate level, indicating that over 78% of the participants had achieved an intermediate level of proficiency (according to ACTFL criteria) prior to the study abroad. Following the semester-long study abroad experience, Brecht et al. found that of the students who began the program at an intermediate level of oral proficiency, only about 40% achieved an advanced level of proficiency by the end of the study abroad. Such rates of achievement can be readily compared with proficiency test scores resulting from both standard formal
classroom training and special immersion programs that are conducted in the United States.

In a study on the assessment of L2 Russian proficiency related to the four major areas of reading, writing, listening and speaking, Thomson (1996) examined data related to the levels of proficiency that are attainable by L2 learners under standard academic conditions. The study endeavored to answer four major research questions related to proficiency in L2 Russian: a) What level of proficiency is attained in Russian in the four areas after one, two, three, four, and five years of study? b) Does attained proficiency level show significant positive correlations with the amount of study? c) Do proficiency levels in the four areas show any significant positive inter-correlations? d) What are the problems that arise from the establishment of ACTFL proficiency standards for FL study? The results demonstrated that, while the number of low proficiency learners decreases and the number of high proficiency learners increases with length of study (as expected), the range of proficiency was found to overlap between different levels of experience, without any definite correlation between proficiency levels in the four areas and the amount of experience. She also found that inter-correlations between the four areas were not very strong. She concludes that such findings suggest that learners develop along different, unparalleled trajectories and that the use of proficiency guidelines to set learning outcomes and goals may not reflect developmental realities. It is certainly possible that what appears to be different ‘trajectories’ is merely a reflection of individual differences (IDs) in cognitive and other factors.

In a follow-up study, Thomson (2000) evaluated similar data among students who studied at Middlebury’s Intensive Summer Russian School following varying amounts of
experience studying Russian under formal instruction in college classrooms. Prior to the
Middlebury program, the subjects had studied for anywhere between one and five years.
Thomson reported that on average, students attained an advanced level of oral
proficiency after five years of study followed by the immersion experience. Such results
have been interpreted to indicate that in spite of course titles that are often used for
college classes (i.e. ‘Advanced Russian’), the second and third years of formal instruction
actually constitute intermediate levels of language (Pavlenko, 2006).

In a study similar to the research conducted by Thomson (1996, 2000), Rifkin
(2005) collected and examined more recent proficiency data for adult learners of Russian.
The study attempted to acquire information about three specific areas: a) determine if any
significant correlations exist between the number of contact hours in traditional
classroom learning and proficiency in the four main areas of listening, reading, speaking
and writing; b) learn whether a major difference in “acquisition” exists between a short-
term (9-week) immersion experience and traditional FL classroom learning; c) examine
whether there is any correlation between proficiency level in the four skills and
grammatical accuracy. To explore these three areas, Rifkin devised specific proficiency
tests that were based on ACTFL Guidelines for the different participants in the study.
Participants ranged from being new to Russian study to those who had over five years of
experience in various colleges in the US and experiences in Russian-speaking countries.
Testing included both pretests (for those with prior experience) and posttests, consisting
of both the standard proficiency test format and a grammatical short-answer section (a
cloze-type task), and an essay (for upper intermediate-advanced learners) and an oral
interview. The grammar test was not related to the ACTFL proficiency guidelines, which only examine the skills of listening, speaking, reading and writing.

For analysis, a straightforward statistical comparison of achieved level of proficiency in the five areas per number of contact hours in both traditional college classroom programs and the Middlebury immersion program was made. The study reported results that do not differ in any significant way from the results of studies conducted by Brecht et al. (1993) or Thomson (1996, 2000). The key results of this study seem to indicate that the number of contact hours provided in traditional college classes is inadequate for promoting an advanced level of proficiency, unless such learners also experience an immersion program such as that offered by Middlebury College in Vermont.

Rifkin proposes that study in the traditional college environment is “constrained by a ceiling just below the advanced level” (p. 13), and that those students who do not experience an immersion program will have great difficulty acquiring an advanced level of proficiency. Based on reported data from research related to study abroad programs, in which grammatical ability has been seen to correlate with attainment of higher levels of overall proficiency (e.g. Brecht et al., 1993), Rifkin also concludes that in comparison with such a formal type of immersion experience in the United States, immersion in-country prior to formal classroom training may also create a similar type of ‘ceiling effect’ to the attainment of more advanced levels of proficiency, and that this is potentially due to the predominantly naturalistic nature of a study abroad experience. Such observations raise intriguing questions about the role of different learning conditions (e.g. formal instruction versus naturalistic learning) in the attainment of
proficiency and more specifically, the potential role that different learning condition sequencing may play in the development of fossilized forms in the IL grammar, which is a major object of investigation in the current study.

Although the nature of past research related to general adult L2 Russian proficiency is limited to broad ACTFL (or other guideline) descriptions of proficiency rather than an analysis of attainment in terms of specific forms that represent native-like accuracy or concrete measures of fluency, results of Russian proficiency studies, such as Rifkin’s (2005) assertion that naturalistic experiences may have a limiting effect on grammatical accuracy, provide the impetus for a close examination of the relationships between learning conditions and the nature of attainment of L2 proficiency in Russian.

2.7 Research questions and hypotheses

The current study focuses on one general question related to the development of proficiency: For learners categorized according to similar profiles in working memory, what are the differences in attained levels of oral proficiency in L2 Russian between three learning conditions and their possible combinations: 1. strictly formal or instructed learning followed by predominantly naturalistic learning experiences in country; 2. strictly naturalistic learning in country followed by formal or instructed learning; 3. strictly naturalistic learning experiences with little or no formal or instructed learning. Through an extensive interview process and a background questionnaire, participants whose learning was predominantly naturalistic were identified. The first two combined condition categories above were deemed necessary in order to isolate potential effects of learning condition sequencing related to the differences between these contexts: the orientation of one condition, e.g. the focus on form (FonF) nature of formal instruction,
may have an influence on the learning of meaning (morphosyntactic effects on semantic processing), and vice versa. Taking into account the nature of second language proficiency (SLP), the study’s main objective motivates the following research questions and hypotheses (H1, H2, and H3):

1. To what degree does WM predict general proficiency, fluency, and accuracy in the acquisition of L2 Russian under the most common learning conditions?

2. To what degree does WM predict fossilization in the acquisition of L2 Russian under such learning conditions?

3. To what degree does WM interact with learning conditions to affect proficiency in L2 Russian, including accuracy, fluency and the presence of potentially fossilized forms?

4. What are the effects, if any, of the different learning conditions and their various sequencing combinations on the major aspects of attained proficiency and the presence of fossilized forms?

Specifically speaking, the above research questions relate to proficiency in terms of both fluency (as measured via an analysis of the average number of meaningful syllables uttered per minute (Yuan and Ellis, 2003), and accuracy, as compared to the performance of native speaker controls. To test for accuracy (and complexity), three specific structures that represent the morphological and syntactic nature of native-like speech have been identified for elicitation in the present study, including the relative pronoun kotoryj ‘who, which, that’, the relative pronoun to, chto ‘that which’ and the subjunctive/conditional complementizer chtoby ‘so that, in order that’. These structures are among those that characterize superior or advanced level speech and demand accurate control of both morphological (accuracy) and syntactic (complexity) aspects of the language. The three structures also tend to be difficult to master consistently, while in native speech they are so prevalent and required in certain contexts that elicitation tasks
can be designed in which native speakers respond consistently with the targeted structures. An additional aspect of accuracy is also examined in the present study in relation to measures of WM and the three learning conditions: the presence of potentially fossilized forms in participant speech. For the purpose of this study, a ‘potentially fossilized form’ refers to the occurrence in speech of a specific pattern of inaccuracy that recurs consistently (at least three times in a continuous monologue) whenever an attempt to utter the structure is made. For example, a speaker may consistently apply accusative animate case marking to inanimate masculine direct objects (inanimate masculine nouns appear in the nominative form as direct objects), or may consistently use imperfective verbal aspect for one-time completed actions for which the perfective should be utilized.

In light of the difficulties associated with the acquisition of L2 Russian proficiency in terms of accuracy, and the evidence from previous research related to the predictive validity of WM in a variety of learning conditions, the above research questions motivate the following three hypotheses:

H1. It is predicted that WM will demonstrate stronger positive correlations with measures of proficiency under strictly naturalistic learning conditions than under conditions that rely more heavily on formal instruction.

H2. It is predicted that the interaction between WM and the three different types of learning conditions will demonstrate differential effects in both overall proficiency in L2 Russian, and among the different linguistic subsystems of the IL grammar.

H3. It is predicted that the amount of formal learning experience will demonstrate a significant negative correlation with the presence of potentially fossilized forms in the IL grammar, with learning condition sequencing also having a significant effect.
CHAPTER III: RESEARCH METHODS

3.1. Challenges Related to Experimental Design

The current study is inspired by the call of a number of researchers (e.g. Dörnyei, 2005, 2009; Robinson, 2001, 2002; Skehan, 2002) for a research program that explores various aptitude – learning environment interactions in acquisition. The study was therefore designed to explore proficiency in L2 Russian as the product of the relationships between working memory and two specific learning conditions that occur in various combinations, formal instruction and naturalistic learning. The specific conditions targeted included: strictly naturalistic learning experiences, formal instruction followed by naturalistic learning, and naturalistic learning followed by formal instruction.

The attempt to discover the interactions between working memory and these different learning conditions presented a number of key challenges. First of all, the study’s experiment evaluated the working memory capacity of learners who have already acquired a level of proficiency as a result of previous learning experiences. The first challenge therefore related to the need to verify the generally static nature of working memory as a construct, which is attested in much of the literature as an ability that is generally not ameliorable to previous language learning experiences (e.g. Jilka, 2009a; Reiterer, 2009; Rota and Reiterer, 2009). It was therefore necessary to conduct a pilot test of the WM assessment instrument prior to its use in the experiment.
A further challenge was related to measuring the learner’s actual second language proficiency (SLP). Based upon previous research related to the nature of proficiency and its measurement (e.g. Skehan, 1998; Robinson, 2001), this study elicited and measured indicators of two major aspects of oral proficiency: accuracy and fluency. Though an additional aspect of proficiency, complexity, is often attested in the literature (e.g. Skehan, 1998), Robinson’s framework (2001) relies on a multiple-resources view of the learner’s capacity, and when proficiency is evaluated carefully, complexity and accuracy can be conflated into one aspect of proficiency (Robinson, 2001). The elicitation tasks chosen for this study were therefore designed to collect data that reflect both the morphological accuracy and syntactic complexity of the participant’s oral proficiency.

Another challenge encountered in designing the current study involved determining the nature of the different learning conditions experienced by study participants. Any characterization of past learning required a careful examination of the learner’s experiences on the study’s survey instrument. Three common types of learning contexts were identified for the present study: strictly naturalistic (incidental) learning, formal instruction followed by naturalistic learning, and naturalistic learning followed by formal instruction. In order to accurately define the learning conditions that were experienced by study participants, great care was taken to ensure an exact classification of each learner’s past experiences as well as the amount of time learners had devoted to developing competence in L2 Russian under each condition.

3.2 Study Participants

37 adult participants were recruited for the present study, with an additional three native-speaking adults who provided samples of native-level proficiency for comparison.
Of the 37 main study participants, 32 were recruited and interviewed in three different cities within two Russian-speaking countries, Ukraine and the Russian Federation. Data were collected during an extended period of travel specifically designed for this study. The cities that were targeted included Kiev, Ukraine (14 participants), Moscow, Russia (13 participants), and Ufa, Russia (5 participants). These cities are characterized as major Russian-speaking centers, where the linguistic forms targeted for this study are standard features of everyday speech. All 32 participants who were interviewed in Russia and Ukraine are expatriate English L1 learners who are currently either actively engaged in learning Russian, or who have completed formal learning experiences and are now involved in working within the local economy. A number of these overseas participants (n = 10) began their Russian L2 learning experiences at the intensive Russian language Institute (RLI) in Columbia, South Carolina, a program consisting of four hours of intensive formal instruction and conversation practice per day. Such training was followed by either continued formal training in-country or more naturalistic learning with conversation partners or in the course of everyday life and work.

The remaining five participants were recruited in the United States, and have lived for extended periods of time in Russian-speaking countries. Three of these remaining subjects were tested during training at RLI in Columbia, South Carolina. All three participants had extensive naturalistic exposure prior to the intensive program and have since returned to their lives and work in two large metropolitan cities in Ukraine, Kiev and Odessa. Like the other locations involved in this study, the Russian spoken in Odessa is characterized by widespread use of the linguistic forms targeted for this study. The two remaining subjects are married to native Russian speakers and spend a
considerable portion of each day actively using Russian. These learners speak little
English at home since their spouses are more comfortable in their native language. Both
native-speaking spouses served as native-speaking controls. These two participants also
take periodic trips to Russian-speaking countries with their spouses and families. One of
the key characteristics of the target population included extremely high motivation for
learning, since all participants in this study shared the goal of becoming as proficient in
Russian as possible for the purpose of life or work in a Russian-speaking environment.

**Participant Age of Arrival and Amount of Exposure:**

The majority of the participants were over the age of 21 when they began
immersion in L2 Russian, with one participant who began immersion at the age of 17.
Age of arrival was determined to mean either the age at which immersion in-country was
first experienced (for those who began their learning naturally), or when immersion
in an intensive formal program in the United States was first begun (for those who began
their learning under formal conditions). Age of arrival for all 36 participants ranged
between 17 and 55, with the median age of arrival at 31.8 yrs. The study included one
participant who was under 21 (17) when first immersed, twelve participants who were
between the ages of 20 and 29, 19 participants between the age of 30 and 39, and four
participants whose age of arrival was past the age of 40. Two participants were in their
mid fifties when immersion was begun, one who was 54 and the other 55.

Approximately four months of intensive formal instruction or immersion at about
four hours of experience per day is, on average, the minimum amount of time in language
learning necessary to have adequate exposure to the forms and structures that were
targeted for elicitation in this study (Dr. Curtis Ford, personal communication, April 6,
Therefore, in order to establish a baseline of measurable acquired proficiency and to ensure that participants had been exposed to the targeted forms, the minimum amount of exposure required for participants was set at ten months of total experience for either the combined learning contexts (formal followed by naturalistic or naturalistic followed by formal instruction) or for the strictly naturalistic condition. In order to categorize participants according to amount of exposure, the overall amount was calculated by considering 20 hours to be one full week of exposure, with 80 hours comprising a month’s worth of exposure to L2 Russian. Based on these criteria, five participants of the study (13.5%) reported having had less than a total of 20 months of exposure, with the lowest amount being 15.75 months. Ten participants (27%) had had between 20 and 80 months of exposure, another ten subjects (27%) reported between 80 and 140 months, and six participants (16%) reported having between 140 and 200 months of exposure. Six additional participants (16%) described over 230 months of exposure, with the highest amount being 259 months. The median length of exposure in this study was 106 months or approximately eight years.

Types of Participant Learning Experiences

Besides the criterion of amount of exposure, participants were selected and grouped according to the nature of learning strategies or conditions, and were divided into three specific categories representing three independent variables: 1) Learning experience was begun under formal instruction and was followed by naturalistic learning experiences (FN); 2. Naturalistic learning experiences were followed by formal instruction (NF); 3. Learning has occurred under strictly naturalistic conditions (NC) with virtually no explanation of the grammar. Conditions that were characterized as formal
instruction included four different types of experiences: a) classroom instruction (either stateside or in-country); b) formal intensive study (stateside learning that involved at least 15 hours per week of language study); c) in-country intensive study (also involved a minimum of 15 hours per week of in-class study); d) grammar-related tutoring (in-country). Four additional conditions reported by participants were characterized as naturalistic learning experiences: a) incidental everyday interaction; b) conversation partner or tutor; c) in-country conversation classes (non focus-on-form); d) self study involving the learning of vocabulary. In addition to the above experiences, some participants reported the independent study of grammar using a textbook, which was included as a type of formal self-instruction.

The learning conditions categorized as formal instruction above included only those strategies that are associated with a conscious focus on form, not so-called ‘tutoring’ that involves conversation practice without explanation with a native speaker. Conversely, those strategies categorized as naturalistic did not include any focus-on-form instruction or implicit form-oriented activities such as CLT, and were associated with learning that occurs in the course of focused conversation (as with a native-speaking conversation partner or ‘tutor’) or incidental interaction that occurs during everyday life or work activities. Participants provided estimates of the average number of hours spent in each type of activity per week (level of intensity) as well as the number of months and/or years spent engaged in each type of learning experience. For the purpose of data analysis, participants were grouped according to similar amounts of experience under the three learning conditions targeted in this study.
Of the 37 participants in the study, 13 (35%) reported having begun their learning of L2 Russian under formal instruction which was consequently followed by naturalistic learning experiences (FN group). Another 12 participants (32.5%) reported having begun with naturalistic learning strategies which were then followed by formal instruction (NF group). The remaining 12 participants (32.5%) had relied on strictly naturalistic strategies (NC group) in their attempts to acquire the language. Of the participants within the FN group, the median amount of exposure was 112 months or 9 years, while 69% of the participants in this category have had more than 50 months (four years) of exposure. For those included in the NF category, the average amount of exposure was 72 months or 6 years, with 58% of these subjects reporting over 50 months of exposure. Lastly, the median amount of exposure for those learners who reported largely naturalistic experiences (the NC group) was 145 months, or 12 years, while 83% of these participants had at least 50 months of exposure prior to testing. A breakdown of the study’s participant pool by learning condition and amount of exposure is provided in Table 3.1.

**Table 3.1 Study Participants**

<table>
<thead>
<tr>
<th>Learning Condition</th>
<th>Number of Participants in Terms of Months of Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;50 months</td>
</tr>
<tr>
<td>Formal/naturalistic</td>
<td>4</td>
</tr>
<tr>
<td>n = 13</td>
<td></td>
</tr>
<tr>
<td>Naturalistic/Formal</td>
<td>4</td>
</tr>
<tr>
<td>n = 12</td>
<td></td>
</tr>
<tr>
<td>Naturalistic</td>
<td>1</td>
</tr>
<tr>
<td>n = 12</td>
<td></td>
</tr>
</tbody>
</table>
3.3 Materials

Prior to testing and completion of the questionnaire, subjects were first provided with an information sheet describing the research and any benefits for participation. This was followed by a brief initial interview that was conducted to ensure that the prospective participant had both the kinds of learning experiences sought for the study, as well as the minimum amount of exposure needed for data collection. Following this short interview, those learners who fit the study profile were provided with a consent form to sign. This form assured participants of complete anonymity and also emphasized that taking part in the research was completely voluntary. Testing for both working memory and proficiency was administered following completion of the consent form.

Participant Questionnaire

Although in similar research it is common to administer background questionnaires prior to testing, in this study subjects first completes the WM and proficiency tests so that the testing experience could help inform the participant’s understanding of both the nature of proficiency in general, and his or her individual proficiency prior to providing a personal assessment of abilities. A post-test questionnaire can help to elicit a more accurate assessment of the individual’s past experiences with proficiency or aptitude testing, while also prompting the individual to think carefully about past learning experiences, and was found to promote additional questions and clarification during the questionnaire interview process. Once testing was completed, the language background questionnaire (see appendix A) was used to elicit data related to three major areas: a) general information about the participant’s experiences learning Russian, including the age when study or immersion was begun, the goals for learning
Russian, whether the participant was currently active in learning.; b) the types of learning strategies that were used (e.g. formal classroom training (either in-country or in the home country), intensive language programs, tutoring, conversation practice, naturalistic learning through interaction, including duration of each experience in months or years, the average number of contact hours per week, locations and dates; c) a self assessment of current level of proficiency with additional questions related to the participant’s understanding of the meaning of proficiency and any other relevant experiences, such as whether their proficiency had ever been tested and past experiences with other languages.

Once the participant completed the questionnaire, the researcher read it carefully with the participant present. This was followed by a brief interview designed to ask for clarification on certain items. This was especially important for those questions that asked for information related to the reported learning conditions; their nature, amount of time spent under each condition, and everyday experiences with the language. Any determination of the length of study demanded an exact quantification of the number of hours spent in each type of learning activity in terms of total years and months, the number of hours per week, and a further qualification of the extent of immersion in terms of the approximate number of hours spent conversing with nationals in the target language each week.

In order to address the limitations of a reliance on self-reports, in order to classify learners into categories of equivalent experience, the questionnaire paid particularly careful attention to both the amount and nature of language learning experiences, with definitions of all major terms provided for the participant. The duration of language courses or the amount of time spent in interaction (including the approximate number of
contact hours each week) was emphasized, as well as the exact types of language activities experienced with tutors. The information provided was also often augmented via the post-questionnaire interview, depending on the clarity of the information provided by the participant. Participant data was identified on the questionnaire via a participant code that was assigned to each learner. This code was the only means by which individual participants could be tied to the data.

**Aptitude Test for Working Memory**

The test that was used to assess WMC was a memory span task based on Osaka and Osaka’s (1992) reading span task. Following pilot testing of the working memory task instrument, it was found that on a few task items participants were able to rely on additional cues for memorization, such as when the sequence of targeted words could be combined as a potential phrase to aid in memory. These items were therefore switched with others and all the slides were controlled for quality so that no such strategies could be used. A final version of the task was then used as the main aptitude assessment instrument for all regular study participants. Prior to testing, each participant was provided with an explanation of the exact nature of the task, the number of items it contains, and its overall purpose in the study. The task was usually conducted prior to the proficiency interview. Approximately 20 minutes were required to complete the reading span task. A description of the task with examples of subtest questions follow:

b) *Working Memory Capacity Reading Span* task (approximately 20 minutes): For this test, a laptop computer was used. The task included a series of slides, each of which contains a different number of sentences (varying from 2 to 6 sentences per slide), which consist of between 7 and 12 words. Each sentence contained one word in bold, the position of the bold word varying and randomized from sentence to sentence. During the test each sentence appears long enough for the subject to read it, and then
disappears, followed by the next sentence on the slide. Following practice with a similar, one-slide practice example of the task, the subject is asked to read each sentence out loud as it appears, while simultaneously silently memorizing the word that appears in bold. It was explained that no additional strategies should be used during the test, such as repeating the target word to oneself or using voice inflection to emphasize the word. At the end of each slide, the proctor states “end,” and the participant is asked to recite the words memorized (though not necessarily in the same order of their appearance). The total number of items on this task is 48. The WM score is the total number of items successfully recalled and is recorded as a percentage.

Test Items:

Slide 7:

Eighty five years ago there was a field here.
He will wear a strange suit he found in the attic.
Your arrival with the box of books was timely.
The essence of her argument is obviously false.

Oral proficiency test

The oral proficiency test consisted of three distinct sections that were developed to evaluate three aspects of a learner’s oral proficiency: a) a measure of fluency in terms of the number of meaningful syllables uttered per minute, b) 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 c) overall ability in terms of both command of vocabulary and effective use of major structures commonly used in the language. The total time that was required to complete the entire proficiency interview process was between 40 and 70 minutes, depending on the abilities of the participant. All three major aspects of the oral interview were recorded with a digital recorder.
The first section of the test was used to assess fluency, and was designed to give the participant an opportunity to freely share a specific interest of theirs in a comfortable, relaxing setting. This fluency task was conducted first; it was decided that by giving the subjects an opportunity to begin the interview with a topic that was relatively comfortable for them, they would be more relaxed for the remaining parts of the interview. To facilitate a monologue of at least 3 minutes, the participant was asked to speak about a topic that they felt was the most familiar to them and which they could discuss for that length of time, such as their family, work, background, etc. Each monologue was analyzed in two respects as per Yuan and Ellis (2003): 1. The total number of syllables uttered per minute was counted with the final number of syllables in each monologue divided by the total number of minutes that was required to fulfill the task (number of syllables produced per minute of speech); 2. The total number of syllables, words, that were restated or reformulated divided by the total number of syllables uttered, which was represented as a percentage (percentage of syllables that were reformulated). The final fluency score consisted of the total number of meaningful syllables uttered per minute, which was calculated by subtracting the percentage of syllables that were reformulated from the total number of syllables uttered per minute. Fluency scores were assigned numerical values from one to five, with a level five being within the range of uttered syllables of a typical native speaker (see Section 3.5 for discussion).

The second section of the oral proficiency interview was designed to test for mastery of specific native-level structures common to everyday speech. In particular, this section attempts to verify the level of accuracy that has been attained by focusing on the elicitation of three structures that are representative of native-level accuracy in
morphology and syntax. This was accomplished by presenting the participant with specific situations followed by questions that were designed to elicit three targeted structures that are highly representative of both the complexity and accuracy of native-level discourse. The subject was first provided with specific instructions and told to limit answers during this part of the interview to the specific context given in each situation. Each context or situation involves individuals who are either having a conversation or a person making a statement. Following each context, participants were asked a question, which could only be answered accurately and correctly by using the targeted item. Correct answers to the questions demanded the proper (and native-like) use of the targeted structures (see Appendix B for examples). Section Three contained a total of six scenarios, with two scenarios presented for each targeted form, yielding a total of 12 items.

The three structures targeted for elicitation included two relative clause constructions, including *kotoryj* ‘who, that, which’, and *to, chto* ‘that which’, and one subjunctive or conditional clausal expression consisting of the complementizer *chtoby* ‘so that, in order that’. The Scoring of this section of the test utilized the same nine numerical values used in Section Two. Two test items per structure were included, for a total of 34 possible points on this portion of the test. Depending on the number of elements that comprise the targeted structure, either five or six possible points were scored for each item in this section of the test. The two relative clause constructions consisted of six possible points related to major elements used in these expressions, including two lexical items, two grammatical inflections, one point for correct syntax, and one point for meaning. The subjunctive clause construction targeted for this study
was scored for five possible points related to the elements required, including one point for the lexical item, one for a case inflection, one point for a past tense verb inflection, one for syntax and another point for meaning. Any variation in the misuse of these structures was qualified in determining the degree to which a speaker had mastered the various aspects of the morphosyntax specific to these constructions (see Section 3.4 for a description of experiment data coding). Oral proficiency question examples, scenarios and their translations are provided in Appendix B.

The third section of the oral proficiency test was designed to assess participants’ overall or general level of oral proficiency, as well as the presence of any potentially fossilized forms in their speech. This portion of the interview was based upon the Russian Oral Proficiency Interview developed by Canner and Gavrilyuk (2008). This part of the test was designed to elicit evidence of: a) overall proficiency in terms of level, and b) general command of commonly-used vocabulary items and constructions, such as verbs of motion, verb aspect, and use of prepositions. At the time of testing, the majority of the study’s participants had already attained a minimum of an intermediate or advanced ability in the language. The interview consisted of a series of questions representing levels of proficiency ranging from low intermediate to native-level. To determine the appropriate point at which to begin the interview (e.g. level 3) observations were made during both initial interaction and the fluency section of the test. The levels were based on guidelines designed for the Defense Language Proficiency Test (DLPT) used by the US Defense Department. Each level of ability in this section was comprised of a set of three or four questions designed to elicit examples of the lexical items, commonly used constructions, and grammatical structures associated with the level.
To evaluate the presence of potentially fossilized forms during the interview, samples of continuous participant speech lasting for a total of at least four minutes were analyzed using a procedure similar to that used for measuring fluency. All structural inaccuracies in the speech stream were recorded, including morpho-syntactic errors related to a number of areas, such as verbal morphology and aspect, case morphology, important verbal paradigms such as motion verbs. Only those error patterns that were repeated a minimum of three times in a participant’s speech (without any conscious effort to correct them) were counted as potentially fossilized forms and then the number of fossilized forms per 100 words uttered was calculated. Such data was elicited only from those participants who had been learning or studying L2 Russian for at least four years. This experience limitation was adopted in order to ensure that the presence of such repeated forms was related to overall L2 experience in terms of the reinforcement through communicative usage of inaccurate elements in the IL grammar, rather than the result of current classroom or other experiences common to more novice learners. The minimum amount of continuous speech was acquired by combining monologue answers to various questions conducted during the general interview, including the fluency monologue portion of the test. The final score for potential fossilization was later used in data analysis to determine possible correlations between its occurrence and the different learning conditions as well as scores of working memory.

For the purpose of data analysis, the scoring for the general interview was tabulated numerically. Although the majority of participants began this section either at level 2 or level 3, the test contains a total of nine possible values that represent nine levels of proficiency: A score at level one represents the Novice level of proficiency; level two
the Novice Plus level of proficiency; level three the Intermediate level of proficiency;
level four the Intermediate Plus level of proficiency; level five the Advanced level of
proficiency; level six the Advanced Plus level of proficiency; level seven the Superior
level of proficiency; level eight the Superior Plus level of proficiency; and a score of nine
represents a Native Level of proficiency. Level Nine speakers can be classified as fully
native-like, and possess both native level fluency and accuracy in morpho-syntax.
Furthermore, an additional level, Level Nine Minus, exists between levels eight and nine,
and describes speakers who have attained a near-native level of proficiency (see
Appendix B for examples of proficiency ratings).

**Native Speaker Participants**

Results for the fluency and elicitation sections of the proficiency interview
process described above were compared with similar interviews conducted with three
native speaking (NS) participants. As with the study’s main participants, for the fluency
section of the interview native speakers were asked to describe a topic of interest with
which the speaker was comfortable. Speakers were also told to speak at a normal rate of
speed for them in a relaxed conversational setting. All three NS participants spoke for
the minimum three minutes and each monologue was evaluated in terms of the number of
meaningful syllables per minute that were uttered. NS participants were also tested with
Section Two, the elicitation task. All NS participants readily produced the targeted
items; however, one of the participants exhibited some unusual characteristics associated
with use of the relative clause *kotoryj* in one of the contexts, with the form finally
produced following a specific explanation of the context (see section 4.1 for discussion).
NS scores for these two aspects of the interview were used as a baseline for determining the fluency and accuracy of all main study participants.

3.4 Pilot Testing for Aptitude and Proficiency Assessment Instruments

Prior to pilot testing for the WM component task, the task had also been used extensively during thesis research conducted in 2010 and 2011. For that study the WMC reading span task (as per Osaka and Osaka, 1992) and an analytic language ability (ALA) task (Petersen and Al-Haik, 1976; Pimsleur, 1966) were administered to 34 subjects who had varying amounts of experience learning a second language. These tasks (taken together or separately) demonstrated a high positive correlation with success in L2 learning, and demonstrated the task’s validity in predicting proficiency.

The testing of WMC for the present study was conducted after a level of proficiency had already been acquired in L2 Russian. Prior to its use, the task was therefore pilot-tested with participants both before and after an intensive language program. The pilot test was designed to evaluate the reliability of WM in terms of its resilience to improvement as a result of language learning experiences. Testing was conducted with seven participants who were recruited at the intensive Russian Language Institute held in Columbia, South Carolina. Initial testing with these students occurred in the fall, 2011, prior to their beginning a six-month intensive training experience in L2 Russian. The program consisted of four hours of coursework per day, half of which was related to formal instruction and practice of grammatical forms, new lexical items and the learning of proper sentence structure. The other two hours of instruction were characterized by more implicit and incidental learning activities, including topical or task-based conversation classes and naturalistic interaction with instructors and fellow
students. Following completion of the six-month program, participants were tested for their working memory span a second time. To prevent any possible long-term memory of the items, for the second trial a different version of the task was developed, with about half (23) of the items remaining the same though not in the same order of appearance, and 25 entirely new sentences added and incorporated into the task. The new items conformed to the type of sentences they replaced in terms of approximate number of syllables, and word length. Both versions of the test contained a total of 48 task items.

The pilot test for WM was successful, and appears to demonstrate no significant improvement in working memory span for any of the participants: of the seven learners tested, two achieved slightly higher scores (by a total of one or two items), three actually showed a slight decrease (two participant scores were a total of one point lower, and one by a total of three points). The remaining two subjects achieved scores that were identical to initial testing results (See Appendix B).

**Proficiency Pilot testing**

The test questions (Section A) used to test general oral proficiency in L2 Russian for the present study have been extensively tested and used as an effective tool for diagnosing all major elements of proficiency. These questions represent part of the Oral Proficiency Interview (OPI) that has been used for a number of years as a diagnostic tool for both higher level course placement and final exit testing conducted at the Russian Language Institute (in conjunction with stateside and overseas programs), a year-round intensive language program located at Columbia International University in Columbia, SC. The OPI is based upon the DLPT assessment format used for the past 40 years at the Defense Language Institute. Over 300 participants have participated successfully in this
testing with the procedure yielding consistent and accurate assessments of learner abilities in reading, writing and speech, and the test has been compared favorably with similar ACTFL and European Framework proficiency tests.

More specific to the goals of the present study in terms of morphological and syntactic accuracy, an additional pilot test was also conducted for the elicitation task. The task questions were piloted with two advanced-level L2 Russian acquirers and three native speakers of Russian, one who is from Ukraine and three from Kazakhstan. Prior to pilot testing some modifications to the test questions were made following an interview with one of the native speakers: the original design utilized photographs that the participant was to describe following a specific question, although it was found that regardless of the narrowed context of a photo, a learner could always use visual cues to circumlocute around the structure in the response. The modified test demonstrated that the questions used for this task are well constructed to elicit the targeted forms, with results showing that in each narrowly-defined question context, the speaker was forced to utilize the targeted construction in order to properly answer the question. Feedback on the task was also elicited from the native-speaking participants, who confirmed that use of the targeted structures was the only method available to properly answer each item.

Three of the four native-speaking participants responded exactly as predicted on all items used in the task, receiving scores at level 5. Due to researcher error and at least one other factor, one NS participant used a different structure in her attempt to answer one of the sub-questions, although the answer was corrected when further clarification was given. The pilot test was therefore also instrumental in providing the researcher with
practical guidelines related to the exact manner in which the test should be conducted to ensure the proper elicitation of the target items.

3.5 Data Coding and Analysis

In preparation for analysis, all working memory and proficiency-related data for this study were collected, scored and stored on a password-protected laptop computer. All proficiency interviews were digitally recorded, with the data then coded and analyzed by the researcher. Fluency data were measured as per Yuan and Ellis (2003): 1. The data were first analyzed for the number of syllables uttered per minute by dividing the total number of syllables uttered during the monologue by the number of seconds the participant spoke and then multiplying by sixty; 2. The number of meaningful syllables uttered per minute (MSM) was calculated by repeating this procedure but subtracting the total number of repeated or reformulated syllables. The final results were compared with NS data collected from the four NS participants, and were tallied as a percentage of the average number of MSM uttered by an average native speaker (see Table 3.2). The number of MSM uttered by NS participants fell within a range between 210 and 270 (the lowest NS participant scored 220 MSM, and the highest 258 MSM). For example, a speaker with level three fluency utters an average of between 30 and 40 percent of the minimum number of MSM uttered by a native speaker, while a level four speaker utters between 40 and 50 percent. A level eight speaker utters between 80 and 90 percent of the minimum number of syllables observed in NS speech. Any participant scoring within ten percent of a native speaker was described as ‘near native’, and assigned a score of 8.5 for coding purposes. The performance of a speaker who scores within five percent of the range of NS participants could therefore be described as ‘native-like’.
Table 3.2 Participant Fluency as a Percentage of Native-level Fluency

<table>
<thead>
<tr>
<th>% of Native Level</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
<th>Above 90%</th>
<th>Near-native</th>
<th>Native Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>8.5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Range (MSM)</td>
<td>21-41</td>
<td>42-62</td>
<td>63-83</td>
<td>84-104</td>
<td>105-125</td>
<td>126-146</td>
<td>147-168</td>
<td>169-188</td>
<td>189-209</td>
<td>210-270</td>
<td></td>
</tr>
</tbody>
</table>

The coding of the accuracy data for Section Two (the elicitation task) was achieved by assigning a total number of points for each elicited construction. Points related to specific morpho-syntactic and semantic aspects of the learner’s performance. Six points were assigned to each relative clause structure and five points were assigned to the subjunctive construction with *chtoby*. In addition to the number of items and inflections assigned to each structure, an additional point was related to syntax, and another related to semantic content. Examples are shown in 3-1 – 3-3.

(3-1) Relative pronoun *kotoryj*

O kakikh amerikantsakh govorit Valerij?  
About which(LOC/PL) Americans(LOC/PL) speak(3S.PR) Valerij?  
‘Which Americans is Valerij talking about?’

1 (infl.)  1 (item)    1 + 1 (1 item + 1 infl.)   1 (syntax) 1 (meaning)

ob amerikantsakh, s kotorymy On pogovoril vchera.  
about Americans (LOC.PL) with who (INS.PL) He speak (3S.PF.PT) yesterday  
‘He is talking about the Americans with whom he spoke yesterday.’

(3-2) Relative pronoun construction *to, chto*

O chyom khochet rasskazat’ Andrej Vitaliyu?  
About what (LOC/SG) want(3S.IM.PR) tell(INF.PF) Andrew Vitaly(DAT/SG)
‘What does Andrew want to tell Vitalij about?’

1 + 1 (item + infl.) 1 + 1 (item + infl.) 1 (syntax) 1 (meaning)
o tom, about that(LOC/SG) 1 (synta x) 1 (meaning)
chem what(INST/SG) Zanimaetsja Boris
‘…about (that) what Boris is doing’

Total possible points for each test item above: 6

(3-3) Subjunctive complementizer *chtoby*

Chto khotet Larisa?
What(NOM.SG.) want(NOM.SG.) Larisa
‘What does Larisa want?’

1 (item) 1 (infl.) 1 (Past) 1 (syntax) 1 (meaning)
Ona khotet, chtoby Masha priekhala na dachu
She want(3SG.PR) so that (SUBJ) Masha(NOM) come(PAST/F)
‘She wants Masha to come to the dacha.’

Total possible points for the test item above: 5

The items above were scored differently (six possible points for relative clause constructions and five for the subjunctive construction with *chtoby*), since relative clauses in Russian often have a greater degree of complexity, as seen with the presence of an additional item in each of these examples (the preposition *s* in the *kotoryj* clause, and the demonstrative *to* in the *to, chto* construction). Two additional points were added to each structure, one for native-like syntax and one for retention of meaning. For example, if a participant responded by omitting the correct inflection on *kotoryj*, the basic syntactic structure remains correct, while the meaning has changed. Two points were therefore subtracted (one for a missing inflection and one for a change in meaning), and the item is scored as four out of six. In the event of a missing targeted item, such as the complementizer *chtoby* or the relative pronoun *kotoryj*, two points were immediately subtracted from the score, one for the missing item and one for a change in meaning,
provided the syntactic structure remained, e.g. *Ona khochet, chto Masha priekhala na dachu* ‘She wants that Masha came to the dacha’ would be scored as 3 points, with one point taken off for the missing item and one for the change in meaning. A total of 34 possible points were included in this section of the test and scoring utilized the same 9-level numeric system. As with scoring the fluency monologue, to calculate the level of accuracy, the total number of points scored was divided by the total number of points possible (34) to arrive at a percentage of native-level accuracy; the result was scored as a percentage of native-level accuracy as demonstrated by the NS participants of the study, i.e. 10 – 20 percent as level 1, 20 – 30 percent level two, 30 -40 percent level three, and so on. (the score level ranges in terms of percent of native-level accuracy are identical to those for fluency depicted in table 3.3). Scores ranging within ten percent of native-level accuracy were assigned a score of nine minus or ‘near-native’ accuracy.

The general proficiency or ability score resulting from Section Three of the interview was rated according to the level criteria provided in the oral proficiency guideline (see Appendix C), and as with sections one and two, nine possible levels (plus an additional, ‘near native’ level) were included in the assessment criteria for this section of the test. This part of the interview was not included in the statistical analysis related to the interaction between working memory and learning conditions to produce accuracy and fluency, due to the general nature of the score on this part of the interview. The final overall proficiency or ‘native-likeness’ score was calculated as an average of the combined fluency, accuracy, and general ability measured in part three of the test. The resulting overall degree of native-likeness was then used as one of three dependent...
variables (fluency, accuracy, overall proficiency) as part of the analysis for the predictive validity of working memory under different learning conditions.

The second purpose for the general interview portion of testing, namely recording the number of potentially fossilized forms uttered by participants, was coded in a manner similar to the analysis of fluency. For this procedure only those participants with over four years L2 Russian experience were included in order to ensure that participants would have a minimum amount of exposure to the types of structures considered for the analysis. To attempt to eliminate the possibility that the learner was in the process of learning an observed erroneous structure, any items that learners attempted to reformulate or that did not recur at least three times in speech were eliminated from consideration. Due to the phonological nature of Russian unstressed and stressed inflections as well as verbal morphology and use, it was not necessary to control for any potential phonological effects on accuracy in indentifying potentially fossilized forms: phonology does not interfere with Russian morphosyntax as it does with other L2s, such as English. To determine the degree or amount of potential fossilization in participant speech, the total number of forms uttered repeatedly by a speaker was divided by the total number of words uttered in the sample and then multiplied by 100. The final repeated error rate was recorded numerically. For example, a sample of participant speech that includes 380 words and evidences a total of eight different types of repeated errors would yield a total score of 2.1 potentially fossilized forms per 100 words.

Prior to any analysis of the data, qualifying candidates were divided into three groups related to learning conditions reported on the Language Background Questionnaire: the strictly naturalistic condition (NC), formal instruction followed by
naturalistic learning (FN), and naturalistic learning followed by formal instruction (NF). Based on the goal of comparing subjects who have invested similar amounts of time in their pursuit of proficiency, each group was then further divided into subgroups based on the duration of learning experiences. The learning condition groupings consisted of three subgroups, which were divided in terms of amount of experience as follows: a) below 50 months of exposure (less than four years); b) 50-110 months (between four and nine years); c) 110-170 months (between nine and fourteen years); d) 170-230 months (between 14 and 19 years) of exposure; and e) over 230 months (19 years). The data was then tabulated and recorded in Microsoft excel to prepare for analysis.

For the Naturalistic/Formal (NF) and Formal/Naturalistic (FN) conditions, an additional procedure was utilized in order to further isolate any potential effects of the amount of formal experience on components of proficiency and potential fossilization rate: participants were divided in terms of the percentage of overall experience represented by formal training. For this procedure, participants were categorized according to the following percentages of formal instruction: a) between two and eight percent of total experience; b) between eight and fourteen percent; c) between 14 and 22 percent; c) over 22 percent of experience comprised of formal training.

Prior to statistical analysis, a general analysis was performed in order to arrive at an overall picture of four major characteristics: 1) the level of attainment achieved for the different components of proficiency for all learners collectively; 2) the differences in attainment of components of proficiency between the three learning conditions in relation to the amount of experience and the extent, if applicable, of formal learning experiences (as a percent of overall experience; 3) the predictive relationships between WM and the
major aspects of proficiency for all study participants grouped according to amount of experience; and 4) the predictive validity of WM for the specific aspects of proficiency with study participants divided by the three learning conditions in terms of the amount of learning experience.

Based on these procedures, participants were first grouped together in order to evaluate in general the average attainment of all components of proficiency for specific amounts of experience, as follows: a) between one and four years; b) between four and nine years; c) between nine and fourteen years; d) between 14 and 19 years; and e) over 19 years of experience. Furthermore, for each of the specific learning condition-related analyses, study participants were divided in three ways: 1) into the three different learning condition groups; 2) in terms of the overall amount of L2 learning experience, with groups consisting of individuals with similar amounts of total learning experience, and 3) into subgroups according to general profiles of working memory as well as experience in order to discern predictive relationships between WM and the components of learner proficiency for learners with similar amounts of experience, as well as differences between the conditions. For example, learners whose average working memory was level 2 and whose experience was between one and four years under naturalistic conditions were compared with learners who also scored level 2 in working memory and whose amount of experience was similar under a different learning context.

**Statistical Analysis Procedures**

Each aspect of proficiency was scored numerically with nine possible scores for regular study participants, ranging from level one to level nine (native-level proficiency) for each of the three categories of fluency, accuracy, the general ability proficiency score,
and the overall combined native-likeness measure. Learners could therefore receive combined accuracy/fluency scores of different values (e.g. 3/2). For the general ability score in Section Three, to be assigned a particular level, a learner had to possess a majority of the characteristics of that level according to the test guidelines. Scores for accuracy and fluency were regarded as separate dependent variables for statistical analysis. The general ability score was analyzed separately solely for the purpose of determining the fourth, native-likeness score, which consists of the average between fluency, accuracy and general ability scores. The derived native-likeness score was only used in analyzing the predictive relationships between WM and overall proficiency, since the criteria for the underlying general ability score based on Section Three of the test are more general than the fluency and accuracy scores, though the derived NL score normally closely reflected the combined accuracy/fluency scores achieved.

The score related to the presence of potentially fossilized forms was also analyzed separately as the fourth dependent variable, potential fossilization rate (FR). This variable was included with other dependent variables in both the regression UNIANOVA to test for the predictive relationship between WM and measures of SLP, and in the moderated regression designed to test for potential interactions between WM and learning conditions. Independent and dependent variables are depicted in Table 3.3.

**Table 3.3 Independent and dependent variables**

<table>
<thead>
<tr>
<th>INDEPENDENT VARIABLES</th>
<th>LEARNING CONDITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Working Memory (WMC)</td>
<td>Learning Condition</td>
</tr>
<tr>
<td>(Coded as a percentage)</td>
<td>(/ = followed by)</td>
</tr>
<tr>
<td>Low</td>
<td>Naturalistic Learning</td>
</tr>
<tr>
<td>Average</td>
<td>Naturalistic/Formal</td>
</tr>
<tr>
<td>High</td>
<td>Formal/Naturalistic</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DEPENDENT VARIABLES</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy Levels 1 – 9</td>
<td>Fluency Levels 1 – 9</td>
</tr>
<tr>
<td>Native-likeness average</td>
<td>Fossilization Rate</td>
</tr>
</tbody>
</table>

114
The objective of the three research questions in this study involved a) an attempt to discover predictive relationships between specific measures of WM and attained levels of key aspects of proficiency as well as potential fossilization, b) the effects on the major aspects of attained proficiency of an interaction of WM with different learning conditions, and c) the effects of the different learning conditions and the two different types of sequencing of conditions on both the major aspects of proficiency and the occurrence of potentially fossilized forms. In light of past research that indicates strong predictive correlations between WM and proficiency (e.g. Daneman, 1991; Harrington and Sawyer, 1992; O’Brien, Segalowitz, Collenteen and Freed, 2006; Sawyer and Ranta, 2001), the first procedure adopted in this study was to test for the predictive validity of WM in each learning condition. To examine these predictive relationships between WM and the different aspects of proficiency, logistic regression was utilized. Participant scores for working memory were recorded as a percentage and analyzed in two different ways: first, WM scores were classified as either low, average or high and loaded as factors, and secondly, the scores were loaded as covariates with learning experience. These two different methods were utilized in order to both compare the results to look for any inconsistencies and to compare and verify whether the patterns found in the general analysis were reflected in the regression. The dependent variables for participant proficiency consisted of the three components that were tested individually: an accuracy score related to the elicitation task, a fluency score derived from the fluency monologue, and a score related to the overall proficiency or “general ability” derived from the general oral proficiency interview section of the test (Section Two). A fourth dependent variable related to overall proficiency, called a “nativelikeness” score was also derived by
averaging the fluency, accuracy and general ability scores together. Finally, WM was also tested for predictive relationships with the dependent variable related to the number of potentially fossilized forms occurring in participant speech, or fossilization rate (FR). A separate regression analysis was performed to test for effects of WM on the frequency of occurrence of such forms.

Another goal of the present study was to describe the interaction between working memory, which is a predictive or explanatory variable, and other environmentally-conditioned variables, namely learning conditions. Such an analysis suggests the idea that the strength of the effect of WM is potentially affected by learning condition. In other words, the analysis tests whether the strength of the effect of WM on proficiency is modified or moderated in any way by the nature of the learning condition variable, or potentially vice versa. Along the lines of previous aptitude interaction research, an analysis of the effects of interactions of the two independent variables on the major elements of proficiency therefore required a moderated multiple regression procedure (MMR) (Overton, 2001). Furthermore, according to Overton (2001), the need to provide corrections for heterogeneous error variances also demands a follow-up analysis, and Tukey post-hoc test was therefore performed following initial analysis of the data.

The third goal of this study was to test for potential correlations of the different learning conditions and their two different combinations with major aspects of participant oral proficiency, including the occurrence of potentially fossilized forms. Two different analysis methods were used for this portion of the experiment. First of all, the amount of overall participant experience under the two combined learning contexts was analyzed for correlations with the tested measures, and second the amount of formal training as a
percentage of overall experience was loaded to check for correlations between such percentages and test results for the three components, the nativelikeness score and the observed rate of potentially fossilized forms. Multivariate regression was also used for this procedure. Combined with both the analysis of the predictive value of WM and the analysis of potential interactions between WM and learning conditions, the results present an intriguing picture of the effects such relationships have on the nature of proficiency in L2 Russian.
CHAPTER IV: RESULTS

This chapter is divided into five main sections that follow a general introduction. The introduction provides an overview outlining statistical patterns of proficiency seen in the data for all study participants. Section 4.1 corresponds to the first research question and describes results for WM and its predictive relationship with fluency and accuracy. This first section begins with such results in terms of all participants combined and then reports findings related to the three different learning conditions, including the logistical regression for WM and these two aspects of proficiency. Section 4.2 relates to the second research question, and reports results for the relationships between WM and observed fossilization. This section also begins with findings for all participants, followed by results for the three conditions, and ends with a report of the logistical regression analysis of any relationship between WM and fossilization rate. Section 4.3 reports the data related to the third research question, with findings for the moderated regression designed to examine potential WM-learning condition interactions for elements of proficiency in L2 Russian. Section 4.4 provides results related to the impact of the three conditions as independent variables on L2 Russian proficiency and fossilization. This section reports results pertaining to the two sequences of formal and naturalistic contexts (in the two combined learning conditions) in terms of proficiency, with a regression analysis for the effect of formal instruction. Section 4.5 revisits the research questions and reports the results in terms of the three hypotheses.
For all UNIANOVA regression analyses, significance occurs at or below a p-value of .05. The analysis for WM relationships with components of proficiency involved two steps. The goal of the first step was an examination of the overall predictive relationship between working memory and the elements of proficiency without consideration of learning experience, and for this analysis all participants were grouped together. The second step in the analysis involved a test of the predictive relationships within each individual learning condition, and for this part of the analysis learners were divided by learning context. In order to isolate other key variables as potential factors, both aspects of the experiment also tested for any predictive relationships between the amount and type of experience and attained levels of proficiency measures. For example, for both the FN and NF conditions, an additional experiment was conducted to test the effect on aspects of proficiency of the amount of formal training as a percentage of overall experience (see Section 4.4). To establish the differences among means, an additional Tukey post-hoc test was utilized when possible for any results that demonstrated statistical significance.

**Observed patterns in the data for all participants**

The data collected in this study demonstrate a number of general characteristics. The most apparent aspect is a correspondence between amount of exposure/learning experience and overall level of proficiency, with the highest average increase with experience occurring for fluency under all three conditions. There are also differences in overall levels of fluency, accuracy and general proficiency or native-likeness for individuals among the three conditions for given amounts of experience. A third characteristic is the extent to which participants have successfully acquired native-like
proficiency in terms of overall averages of fluency, accuracy and general ability between the three conditions. Lastly, there were distinct differences among conditions in the average number of fossilized forms recorded during the proficiency interview.

**Results for fluency, accuracy, native-likeness, and fossilization rate**

When the data for this study are examined collectively for all participants regardless of type of learning condition, results show an increase in average fluency over time. To categorize all learners together in terms of the amount of learning experience, participants were divided into six subgroups: 2-4 years, 4-8 years, 8-12 years, 12-16 years, 16-20 years, and over 20 years of experience. An overview of the data appears to show a correlation between the amount of experience and average fluency among participants whose experience ranges between two and twenty two years: there was an average fluency of 5.9 (between 50 and 60% of native-level fluency) for learners who have between two and four years of L2 Russian experience (no participants in this study possessed less that two years of experience in L2 Russian); 7.4 (between 70 and 80% of native-level fluency) for those with between four and eight years of experience; an average fluency of 8.5 (between 80 and 90% of native-level) for learners with between eight and twelve years of experience. In this data average fluency rates appear to drop off for learners with over 12 years of learning experience: those learners who have between 12 and 16 years experience averaged a fluency rating of 8.1 (80 – 90% native level), an average of 7.7 (70 - 80% of native-level fluency) was found for learners with between 16 and 20 years, and 6.7 was recorded for learners with over 20 years of L2 Russian experience. Averages for participant accuracy, overall native-likeness and potential fossilization show similar changes over time.
For participants collectively average proficiency in terms of accuracy as measured on the elicitation task was 5.9 (approximately 60% of native level accuracy) for learners with between two and four years of experience; 4.9 (approximately 50% of native-level accuracy) for those with between four and eight years; 6.9 (approximately 70% of native-level) for learners with between eight and twelve years; 6.7 (60 – 70% of native level accuracy) for learners with between 12 and 16 years, and 6.0 for those with between 16 and 20 years of experience. The data shows that the average accuracy for learners with over 20 years of experience drops to 4.0 (approximately 40 percent of native-level accuracy). The data related to the general proficiency interview, which was used along with fluency and accuracy data to derive the overall general ability or native-likeness score, indicated that learners with two to four years of experience received an average score of 3.9; those with four to eight years of experience averaged 4.7; those with between eight and twelve years 4.4; an average of 5.3 was recorded for both those with between 12 and 16 and 16 and 20 years of experience; and an average of 4.7 was found for participants with over 20 years of learning experience.

When the three test scores of fluency, accuracy and the proficiency interview are averaged together to derive the native-likeness (NL) score, the data shows that learners with two to four years of experience received an average score of 5.2; those with four to eight years of experience averaged 5.7; those with between eight and twelve years 5.6; an average of 6.7 was recorded for those with between 12 and 16 years, and a score of 6.3 demonstrated by learners with between 16 and 20 years of experience. For learners with over 20 years of learning experience, the average native-likeness score dropped to 5.1.
Last of all, combined participant data for the observed rate of potential fossilization (FR), which was measured for learners with over four years of experience, demonstrated the following characteristics for the same experience-related groups: a) average FR for participants with four to eight years of experience was 1.7; b) the average for those with eight to twelve years was 1.6; c) FR averaged 1.8 for participants with between 12 and 16 years of experience; d) the observed FR for those with from 16 to 20 years was 1.9; and e) average FR increases to 4.0 for participants with over 20 years of L2 Russian experience. Average proficiency scores for fluency (FL), accuracy (AC), native-likeness (NL) and fossilization rate (FR) for levels of experience are depicted in Table 4.1.

Table 4.1 All Participants: Average proficiency scores

<table>
<thead>
<tr>
<th>Experience (years)</th>
<th>FL</th>
<th>AC</th>
<th>NL</th>
<th>FR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 - 4</td>
<td>5.9</td>
<td>5.9</td>
<td>5.2</td>
<td>-</td>
</tr>
<tr>
<td>4 - 8</td>
<td>7.4</td>
<td>4.9</td>
<td>5.7</td>
<td>1.7</td>
</tr>
<tr>
<td>8 - 12</td>
<td>8.5</td>
<td>6.9</td>
<td>5.6</td>
<td>1.6</td>
</tr>
<tr>
<td>12 - 16</td>
<td>8.1</td>
<td>6.7</td>
<td>6.7</td>
<td>1.8</td>
</tr>
<tr>
<td>16 – 20</td>
<td>7.7</td>
<td>6.0</td>
<td>6.3</td>
<td>1.9</td>
</tr>
<tr>
<td>Over 20</td>
<td>6.7</td>
<td>4.0</td>
<td>5.1</td>
<td>4.0</td>
</tr>
</tbody>
</table>

4.1 Working memory as a predictor of fluency and accuracy

This section first provides a description of WM-related data for all participants regardless of learning conditions. This brief section is followed by a description of findings for predictive relationships between WM and the two measures of fluency and accuracy for all three learning conditions. The section concludes with results for the UNIANOVA regression experiment designed to examine the extent to which WM predicts fluency and accuracy under the three learning conditions.
All Participants: WM as a predictor of fluency and accuracy

When all participants are taken collectively, differential results were found in connection with the potential relationships between working memory and each of the major aspects of proficiency. First of all, results for the native-likeness (NL) or general proficiency score derived by averaging the scores for fluency, accuracy and the general interview show no consistent relationship with WM, although a positive correlation appears to exist within certain experience categories. The NL score shows an increase with increasing levels of WM for learners with between one and four years of experience: low WM learners averaged 4.0, average-level WM participants 4.5, and those with high levels of WM averaged 5.9 on the Native-likeness measure for this group. A similar increase in the NL score occurs with increasing WM ability for those with between 9 and 14 years and those with over 19 years of experience, whose average NL score was 5.7 and 5.1 for average WM learners, and 6.6 and 6.3 for those with high levels of WM, respectively. Such an increase in the NL score with increasing levels of WM was not observed, however, for learners who had between four and nine years of experience: the average NL scores for these learners were 6.4 for average WM participants, and 4.7 for those with high levels of WM. Lastly, the average Native-likeness score for the average-level WM learners with between 14 and 19 years of L2 Russian experience was 6.3 (all learners in this experience group had average level WMC).

In terms of the predictive relationship between WM and fluency, for all participants combined no consistent relationship between them was found: the results demonstrate that for a considerable number of participants with similar amounts of L2 Russian experience, fluency actually decreases with greater levels of working memory, while
fluency increases for others who have higher levels of WM. More specifically, results for fluency indicate that learners with between one and four years of overall L2 Russian experience show a decrease in average fluency with increasing working memory: average fluency for those with Low levels of WM was 7.5, but was 5.6 for those with an average level and 6.0 for those with a high level or working memory. Similar results are seen for learners with four to nine years and between nine and fourteen years of experience. Those with between four and nine years of experience and an average ability in WM averaged a fluency of 7.8, while those with high WM averaged a level of fluency at 6.8. For those with between nine and fourteen years of L2 experience, average fluency for learners with an average ability in WM was 8.2, while for those with high WM resources, the average was 7.8. All of the participants who have been learning L2 Russian for between 14 and 19 years showed an average level of WM ability. Average fluency for these learners was 8.0. Learners who have over 19 years of experience learning Russian do appear to show an increase in average fluency with higher levels of WM, and average fluency for learners with an average level of WM was 6.9, while those with high WM ability averaged 8.0. The results for participant accuracy show similar patterns as those seen above.

The results for accuracy as measured on the elicitation subtest also indicate that WM does not appear to be a reliable predictor of this aspect of proficiency when all participants are grouped together. For learners with between one and four years of experience, the accuracy rate does increase with increases in WM: average accuracy was 3.0 for those who have a Low level of working memory, 4.5 for learners with an average ability in WM, and 7.1 for those with a high level of WM resources. The average
accuracy for learners with between four and nine years of L2 experience, however, drops from 6.8 for those who have an average WM ability to 3.2 for learners with a high degree of WM (there were no learners in this category with Low levels of WM). For learners who have between nine and fourteen years of experience, average accuracy was 7.0 for those with average WM ability, and 6.3 for learners with high working memory. As noted above, all study participants with between 14 and 19 years of L2 Russian were tested to have an average level of ability in WM. The average accuracy for these learners was 6.0. Lastly, the average accuracy score for learners with over 19 years of experience does increase with an increase in WM, with 4.0 the average for those who tested average in WM and 5.5 for those with high WM ability.

The lack of any apparent consistent predictive relationship between working memory and either fluency or accuracy for all study participants irrespective of learning contexts was confirmed through a logistic regression for each of these proficiency components. The results indicated confirmation of the null hypothesis for fluency data (F(2) = 2.185, p. = .141). In terms of participant accuracy, no significance for a predictive relationship of working memory was also found for this measure (F(2) = .591, p. = .564). The native-likeness score likewise showed no significant relationship at F (2) = 1.013, p. = .383. To determine if correlations exist between the amount of experience and fluency or accuracy, an additional regression for both total and naturalistic experience was also performed, with results that showed no significant relationship between total experience and either fluency (F(2) = .550, p. = .586) or accuracy (F(2) = .167, p. = .848). These results indicate that WM does not predict these two measures of
proficiency for this study’s learners without consideration of the type of learning experience they may have had (see Table 4.2).

### Table 4.2 All Participants: AC, FL and NL by WM

<table>
<thead>
<tr>
<th>Experience (yrs.)</th>
<th>WM Level</th>
<th>FL</th>
<th>AC</th>
<th>NL</th>
<th>FR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-4</td>
<td>Low</td>
<td>7.5</td>
<td>3.0</td>
<td>4.0</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>5.6</td>
<td>4.5</td>
<td>4.5</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>6.0</td>
<td>7.1</td>
<td>5.9</td>
<td>n/a</td>
</tr>
<tr>
<td>4-9</td>
<td>Average</td>
<td>7.8</td>
<td>6.8</td>
<td>6.4</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>6.8</td>
<td>3.2</td>
<td>4.7</td>
<td>2.0</td>
</tr>
<tr>
<td>9-14</td>
<td>Average</td>
<td>8.2</td>
<td>7.0</td>
<td>5.7</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>7.8</td>
<td>6.3</td>
<td>6.6</td>
<td>1.5</td>
</tr>
<tr>
<td>14-19</td>
<td>Average</td>
<td>8.0</td>
<td>6.0</td>
<td>6.3</td>
<td>2.2</td>
</tr>
<tr>
<td>Over 19</td>
<td>Average</td>
<td>6.9</td>
<td>4.0</td>
<td>5.1</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>8.0</td>
<td>5.5</td>
<td>6.3</td>
<td>2.3</td>
</tr>
</tbody>
</table>

**WM as a predictor of fluency and accuracy among the three conditions**

A general analysis of the relationship between WM and scores for fluency and accuracy for the three learning conditions reveals two main patterns. First, there are differences among the three learning contexts in the extent of a possible relationship between the ability and these components of proficiency: for fluency the results are highly variable within conditions and indicate a lack of any consistent predictive relationship between WM and the measure among the targeted learning conditions. The second observation involves differences between the different conditions in terms of a potential relationship between WM and accuracy, with a predictive relationship for accuracy revealed in the naturalistic condition, but not for the other contexts. The following overview of the WM-proficiency component data will also include general proficiency or native-likeness, and approaches the topic by describing the observed trends.
in terms of the different amounts of experience for each learning context in relation to three major levels of working memory: High (.78 – 100 percent), Average (.56 - .77) and Low (below .56). The relationships between WM and component scores will be reported in relation to the highest, lowest, and average scores for each experience-defined group. The same procedure is used for fossilization data in Section 4.2.

**WM and the native-likeness score among the three learning conditions**

As an average of the three major aspects of the proficiency interview, fluency, accuracy, and the general interview score, the native-likeness (NL) score attempts to determine the overall degree to which learner proficiency conforms to native-like patterns of speech. As such, the NL measure is perhaps the best general indicator of overall participant proficiency measured in this study. In terms of the degree to which working memory acts as a predictor of the NL measure, the results appear to show a limited and inconsistent level of predictive validity for WM in terms of average proficiency as measured by the NL score, except in relation to individual test scores: with only a few exceptions, the individuals who demonstrated the highest level of WM within each experience condition also scored highest in the NL measure.

With respect to overall averages for WM and the NL score, the following patterns are seen in the data for learners with between two and four years of L2 Russian experience: a) the average NL score for NC learners was 3.1 with average WM (.63); b) NF learners demonstrated an average NL score of 5.2 while their average WM was high at .83; c) for FN learners the average NL score was 6.1 while average WM for this group was .78 (average). When learners are classified according to level of WM, the following results were obtained for this experience category: a) the average NL score for NC
learners whose WM was average was 2.0, while the average was 4.3 for the one participant whose WM was measured as Low; b) NF learners showed an average NL score of 5.2 with average WM for all participants being high (.83); c) results for FN learners showed an average NL score of 5.7 for learners whose WM was measured as average, and 6.5 for those with a high level of WM ability.

The individual test scores for this level of experience reveal the following results. Among naturalistic learners the individual with the lowest level of measured WM also scored highest in all aspects of proficiency, including the overall NL measure. For NF and FN learners within this experience group, the individuals who have the highest WM score also received the highest NL scores: in the FN group, the highest NL score was 6.5, which was achieved by a learner with high WM (.85), and in the FN group the highest NL score (7.3) was also obtained by the learner with highest WM (.87).

For learners who have between four and nine years of experience, average WM and NL scores for the three conditions were as follows: a) among NC learners average NL was 5.6 with average WM (.77); b) NF learners showed an average NL score of 4.5 while average WM was .76; c) in the FN group the average NL score was 6.8, with average WM at .70. When learners are classified by level of WM the average NL score for all NC learners (average WM) was 5.6. NF learners with high WM ability averaged 4.3 in native-likeness, with the one participant with Low WM obtaining an NL score of 5.0. All the learners within the FN condition demonstrated an average level of WM (average of .70), and the average NL score for these learners was 6.8.

Individual test score results for learners with four to nine years of experience show that within the NC group the highest NL score (6.3) was achieved by the participant with
the highest WM (.79). Among NF learners the highest achiever similarly received a 6.2 NL score with High WM measured at .87. In the FN group it was the individual with the second highest NL score (7.5) who tested highest in WM (.77, while the highest score for this measure was given to the individual with the lowest level of WM (.62).

Participants who had been learning L2 Russian for between nine and fourteen years demonstrated the following results for the NL measure: a) NC participants obtained an average NL score of 5.0 (average WM of .71); b) the average NL score was 6.3 for NF learners (.75 WM); c) learners in the FN category showed an average NL score of 6.4 with WM at .77. With participants divided by WM level, the NL average score for the Naturalistic group was 7.0 for participants with a high level of WM, and one learner whose WM ability was tested as Low at .37 (the lowest level of WM among all study participants) received a score of 1.0 for the NL measure. The scores for WM for all NF learners in this experience group were within .11 of each other, while the average NL score for these learners was 6.3. FN learners who scored average in WM averaged 6.6 for the NL measure, while High WM participants averaged 6.1. Results found for individual participants in this category among the three contexts reveal that the highest NL score for naturalistic learners was made by the individual with the highest WM (.94). The same pattern was found for both NF and FN learners, with the highest NL scores for these participants being 6.8 (.81 WM) and 7.5 (.87 WM), respectively.

The limited number of individuals classified as having between 14 and 19 years of experience limit the analysis for this experience category, especially in terms of dividing participants according to similar levels of working memory. However, results indicate that for NC learners who showed a high average level of WM at .84, the average NL
score was 6.6. The NL score was 6.3 for the one NF learner whose WM was tested at .71 (Average). Individual test results demonstrate that the highest NL score among NC learners (7.0) was achieved by the participant who scored highest in WM (.96).

Learners with over 19 years of L2 Russian experience yielded the following results for the native-likeness measure: a) NC learners with an average WM that was Average (.72) yielded an average NL score of 5.0; b) FN learners also averaged an average level of WM at .66, with average NL at 6.3. When divided by level of WM, NC learners with Average WM averaged an NL of 4.5, while the one learner with High WM scored a 5.7 on the NL measure. For the FN learners, the average WM was Average while average NL was 6.3. Individual NL test results demonstrate that among learners with over 19 years of experience the individuals who received the highest WM score in both the NC and FN groups were also found to score highest on the NL measure. Average NL scores by the level of WM for the three learning conditions are shown in Table 4.3.

### Table 4.3 Average native-likeness by level of working memory/experience

<table>
<thead>
<tr>
<th>Condition</th>
<th>WM level</th>
<th>2-4 yrs</th>
<th>4-9 yrs</th>
<th>9-14 yrs</th>
<th>14-19 yrs</th>
<th>&gt;19 yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC</td>
<td>Low</td>
<td>4.3*</td>
<td>-</td>
<td>1.0*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>2.0</td>
<td>5.6</td>
<td>-</td>
<td>-</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>-</td>
<td>-</td>
<td>6.7</td>
<td>5.0</td>
<td>6.0*</td>
</tr>
<tr>
<td>NF</td>
<td>Low</td>
<td>-</td>
<td>5.0*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>-</td>
<td>-</td>
<td>6.5</td>
<td>7.0*</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>5.2</td>
<td>4.3</td>
<td>6.5</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FN</td>
<td>Low</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>5.7</td>
<td>6.8</td>
<td>8.0</td>
<td>-</td>
<td>6.0*</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>6.5</td>
<td>-</td>
<td>5.5</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* single participant data

Table 4.3 shows an inconsistent relationship between WM and NL scores for the different conditions. Participants across the three conditions show irregular patterns for
NL scores. For example in the NC condition those with high working memory average 5.0 in overall proficiency in the 14-19 year category, but 6.7 for those with between 9 and 14 years of experience. Similar inconsistencies are evident for both NF and FN learners: NF learners with both high and Average WM and between 9 and 14 years of experience demonstrate average NL scores of 6.5, while FN learners in the same experience category show average NL at 8.0 for Average WM learners, and 5.5 for those with high measured WM. The absence of a predictive WM-related relationship with general proficiency was verified via a logistical regression which accepted the null hypothesis for any predictive relationships between WM and NL scores for any of the three conditions.

**WM and learner fluency among the three learning conditions**

In terms of learner fluency, results appear to indicate no substantial predictive relationship between WM and fluency across the three conditions. For learners with between two and four years of experience, the following overall average scores for working memory and fluency were obtained for the three learning conditions: a) the average WM score for learners within the naturalistic condition was .63 (Average level of WM), while the average fluency was 4.2; b) average WM for Naturalistic/Formal learners within this group was .83 (high), and the average fluency was 5.0; c) average WM for Formal/Naturalistic learners was .78 (high), with an average fluency of 7.2. When participant data were further categorized according to level of WM (Low, Average or High), those NF learners who had a High level of WM also had an average fluency of 5.0. The single NF participant with Low WM showed a fluency of 8.0. Within the FN condition, those learners with an Average level of WM averaged a fluency of 7.5, while learners whose WM was high scored an average of 7.0 for fluency. NC learners were
excluded from this aspect of the analysis since the number of participants within this experience category was limited and divided among different levels of WM. These results would indicate no apparent predictive relationship between WM and proficiency. When individual data is examined more closely, a number of characteristics appear to further demonstrate a lack of consistent relationship between working memory and fluency within the three conditions for all experience-related subgroups.

Within the condition subgroups consisting of learners with between two and four years of experience, the following results were found: a) the highest fluency score for NC learners (6.0) was achieved by the learner with the lowest WM score (.50 or Low WM; b) The NF learner with the highest WM score within this experience category (.86, or high WM) also achieved the highest fluency (6.0); c) for those learners in the FN condition, two learners achieved WM scores of .87 (high WM), and one of them also received the highest fluency score (8.0). With participants categorized according to levels of WM for this experience group, the following results were found: a) WM scores for NC learners were distributed between Low and Average ability, with Average WM learners scoring far lower in fluency (2.2) than the single learner with Low WM ability (6.0); b) all NF learners scored high in WM with average fluency at 5.0; c) FN learners whose WM was Average averaged 7.5 in fluency, while fluency for those with high WMC averaged 7.0.

The following averages for WM and fluency were found for learners who had between four and nine years of L2 Russian experience at the time of testing: a) the average WM for those whose experience was naturalistic was .77 (Average), with an average fluency rating of 7.1; b) the averages for NF learners were .76 for WM (Average) and 6.8 for fluency; c) the average WM score for FN learners was .70 (Average), with an
average fluency of 8.2. When learning condition groups are subdivided according to level of WM, NC learners were all characterized as having an Average level of WM, with the average fluency for such learners 7.0. Among NF learners, those with a high level of WM (ranging from .79 to .87) had an average fluency of 6.5. NF learners who had a Low average WM, showed an average fluency of 7.7. In the FN group the WM level was Average (between .62 and .77) for all participants, with average fluency measured at 8.2.

Among individual learners in this experience group, the highest fluency rating for naturalistic learners was level 8, achieved by the learner with the highest WM (.81 or High WM). For NF learners the highest fluency score was 8.5 (near native level fluency) achieved by the learner with highest overall WM, which was measured as .87 (high WM). Among FN learners the highest fluency rating was level 9 (native level fluency). In this case, the learner who achieved this score had the lowest score for working memory (.62, or Average level) within this experience category.

Learners who had between nine and fourteen years of experience demonstrated the following averages for WM and fluency scores: a) NC learners had an average WM that was .71, and scored an average of 6.0 for fluency; b) average scores for NF learners were .75 for WM, and 8.7 for fluency; c) the average scores for WM and fluency for FN learners were .77 and 7.4, respectively. When learners within this experience group were categorized according to level of WM (Low, Average or high), the fluency averages found for the three learning conditions were as follows: a) NC learners’ average fluency was 8.5 for those whose working memory was classified as high (an average of .88); b) NF condition learners showed an average WM that was Average (.75), and demonstrated an average fluency of 8.5 (near native level fluency); c) in the FN condition those
learners whose WM was High (.86 average) showed an average fluency of 6.7, while those with an Average level of WM (.68 average) showed an average fluency of 8.0.

In terms of the fluency scores for individual learners within the three learning conditions, in the naturalistic group one learner demonstrated a Low level of working memory (.37 or 18/48), and exhibited a fluency of 23 words per minute, or level one. The naturalistic learner with the highest fluency score (level 9, or native-level fluency) also possessed the highest WM (.94, High WM). For NF learners, two participants received the highest scores for fluency (8.5). Of these individuals, one received the highest WM score (.81 high WM), while the other achieved an Average WM score (.69). Lastly, within the FN condition group, the individual with the highest WMC (.87) also received the highest score in fluency, with a score of 8.5, or near-native fluency.

There were a limited number of participants in the experience category of 14-19 years, with just one participant whose experience was naturalistic/formal (NF). This learner’s WM was measured at .71 or Average, with a fluency score of level 7. NC learners in this experience category had an average working memory of .84 or High, and were divided between Average and high WM categories. The average fluency for NC learners with between 14 and 19 years of experience was level 9 (native-level fluency). There were no FN participants with between 14 and 19 years of experience.

Learners with over 19 years of experience constituted a fairly limited category. There were no NF participants within this experience category, while average WM for the FN condition was Average at .66. Average fluency for the FN group was 7.0. The average WM for NC learners was also Average (.72), with an average fluency of 6.7. When divided into WM-related groups, all FN learners were tested to have an Average
level of WM (.71), with an average fluency of 7.0. Naturalistic participants who demonstrated an Average level of WM showed an average fluency of 6.5, while the one learner with high ability in WM received a fluency score of 7.0. Individual scores within this experience group indicate that for NC learners, the individual who scored highest in fluency possessed an Average level of WMC (.72).

**WM and fluency by learning condition: Results for the regression UNIANOVA**

The above results appear to indicate that there is no readily apparent predictive relationship between WM score and fluency for any of the learning conditions targeted in this study. This finding was confirmed via the regression analysis which accepted the null hypothesis for any predictive relationship between WM and participant fluency for all three learning conditions. Average fluency scores by the level of WM for the three learning conditions are presented in Table 4.4.

**Table 4.4 Average fluency by level of working memory/experience**

<table>
<thead>
<tr>
<th>Condition</th>
<th>WM level</th>
<th>2-4 yrs</th>
<th>4-9 yrs</th>
<th>9-14 yrs</th>
<th>14-19 yrs</th>
<th>over 19</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC</td>
<td>Low</td>
<td>6.0*</td>
<td>-</td>
<td>1.0*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>2.2*</td>
<td>7.0</td>
<td>-</td>
<td>-</td>
<td>6.5</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>-</td>
<td>-</td>
<td>8.5</td>
<td>9.0</td>
<td>7.0*</td>
</tr>
<tr>
<td>NF</td>
<td>Low</td>
<td>8.0*</td>
<td>7.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>-</td>
<td>-</td>
<td>8.5</td>
<td>7.0*</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>5.0</td>
<td>6.5</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FN</td>
<td>Low</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>7.5</td>
<td>8.2</td>
<td>8.0</td>
<td>-</td>
<td>7.0*</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>7.0</td>
<td>-</td>
<td>6.7</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* single participant data

The fluency data shown in Table 4.4 demonstrate an inconsistent relationship for WM and fluency for learners in the three conditions. For example, for NF learners both Average and High-WM learners with nine to fourteen years of learning indicated average
fluency scores of 8.5, while for FN learners in the same experience category average fluency for Average-WM learners was 8.0, while it was 6.7 for high-WM participants.

**WM and learner accuracy among the three learning conditions**

Learner accuracy appears to demonstrate different characteristics among the three conditions, with results that appear to indicate an insignificant predictive relationship across conditions. Average scores for accuracy in relation to learner WM demonstrated the following characteristics when L2 experience was between two and four years: a) for naturalistic learners the average accuracy was 3.0 with the average WM at .63; b) average accuracy for NF learners was 6.7 with average WM at .83; c) FN learners within this group showed an average accuracy of 6.6, with an average WM of .78. When divided by level of working memory, NC learner accuracy in this experience category was 2.5 for participants with an Average level of WM, while the accuracy for the one learner with a Low level of WM was measured at 4.0. All NF learners showed a high level of WM (.83 average), with an average accuracy of 6.7. For FN learners the average accuracy was 5.7 for those with an Average level of WM, and 7.5 for those with high WM.

In terms of the observed relationships between individual WM scores and level of accuracy in this experience category, within both the NF and FN groups those with the highest levels of WMC also scored high in accuracy: the top NF learner achieved an accuracy of 8.5 or near-native ability with the highest WMC within this group (.85), while two of the FN learners achieved the highest accuracy, with a score of level 8 for the learner with highest WM (.87), and an accuracy of 8.5 for one participant who also scored fairly high in WM (.77). An exception to this pattern is seen among NC learners, among whom the highest accuracy score was 4.0, which was achieved by the participant
with the lowest WM, measured at .50. The lowest accuracy score in the NC condition was found for the learner with the highest level of working memory (.78).

For learners with between four and nine years of experience, the following results for average accuracy were found: a) for NC learners in this category the average accuracy was 7.0 with an average WM of .77; b) NF learners demonstrated an average accuracy of 2.2 with an average WM of .83; c) results for FN learners showed an average accuracy of 7.1, with an average WM of .70. When participants are divided into groups according to levels of working memory, accuracy for the Naturalistic participants (Average WM) was 6.0. For NF condition learners, all participants except one demonstrated a high level of WM. The average accuracy for these participants was 2.3, while the one learner who scored Low in working memory received an accuracy score of level 2. Among FN learners all participants demonstrated Average WM (.70) and average accuracy was 7.1.

In the four-nine-year experience category, individual accuracy test scores demonstrate the following characteristics: a) Among NC learners the participant with the highest WM score also possessed the highest accuracy, with a score of 7; b) of the NF learners, the highest accuracy score was 5.0, again achieved by the participant with the highest WM ability (.87) in this group; c) the highest score for accuracy in the FN group was level 9, or native-level accuracy, and was achieved by the learner whose WM ability was statistically lowest (.62), while the second highest score (8.5) was found for the participant with the highest measured level of WMC (.77). All learners within the FN group were found to possess an Average level of working memory (between .62 and .77).

Average accuracy results for learners with between nine and fourteen years of experience showed the following: a) NC learners had an average accuracy of 4.5 with the
average for WM being .71 (Average); b) NF learners’ average accuracy was 6.5, with WM was also Average at .75; c) average accuracy for FN learners in this category was 6.5, with Average WM (.77). The following results were found when participants were divided by level of working memory: a) all NC learners except one participant exhibited a high level of WM and showed an average accuracy of 6.7, with the one naturalistic learner with a Low level of WM (.37) scoring 0 on the accuracy elicitation task; b) among NF learners, participants were divided between high and Average WM, but were within .11 percent of each other, with an average accuracy of 6.5; c) for the FN group those learners showing High WM (.86 average) had an average accuracy of 5.5, while the average accuracy for those with an average level of WM was 7.5.

The following characteristics were found for individual test scores for accuracy in the 9-14 year experience group: a) Among NC learners, the participant who scored the lowest in accuracy also demonstrated the lowest level of working memory, with a WM score of .37 (Low); this learner exhibited the lowest level of WM and the lowest accuracy of all participants in the current study; b) the individual among NC learners with the highest accuracy score (8.5 or near-native accuracy) showed the highest level of WM (.94) in this group and the second highest among all study participants; c) in the NF learner group the highest score for accuracy was 7.0, which was achieved by the learner with the highest level of WM (.81); d) among FN learners the individual with the highest level of WM (.87) also achieved the highest score in accuracy, which was 7.0.

Results for average accuracy for learners with between 14 and 19 years experience revealed the following: a) NC learners exhibited an average WM of .84, with an average accuracy score of 4.9; b) WM for the one NF participant in this experience group was
and the accuracy score was 7.0. There were no participants in this experience category for the Formal/Naturalistic learning condition. The number of participants within this experience category was also not adequate for classifying learners according to levels of working memory, with single participants spread among the different WM levels. The individual test score results within this experience category show that the NC learner who had the highest level of WM (.96) also scored highest in accuracy at 5.0.

Learners with over 19 years of experience revealed the following averages for WM and accuracy: a) NC learners showed an Average level of WM (.72), and average accuracy of 4.0; b) WM for learners in the FN context was also Average (.66), with an accuracy of 6.2. There were no NF participants within this experience category. When learners are divided in terms of level of WM, NC learners with an Average level of WM showed an average accuracy of 3.0, while the one High WM learner received a score of 6.0. For the FN condition, participants were divided among different levels of WM. The individual test scores in this experience category show that the NC learner who scored highest in accuracy also scored highest in WM (.96), and had the highest level of WM of all study participants. Among FN learners, the same pattern was seen, and the highest accuracy score (6.4) was achieved by the individual with the highest level of WM.

Although patterns indicating some kind of relationship between working memory and accuracy appear in some aspects of the data, such as that seen for this experience category, the relationship is inconsistent for different levels of experience, indicating that any predictive relationship is likely too variable to be significant. Results for average accuracy by the level of WM for the three learning conditions are shown in Table 4.5.
Table 4.5 Average accuracy by level of working memory/experience

<table>
<thead>
<tr>
<th>Condition</th>
<th>WM level</th>
<th>2-4 yrs</th>
<th>4-9 yrs</th>
<th>9-14 yrs</th>
<th>14-19 yrs</th>
<th>&gt;19 yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC</td>
<td>Low</td>
<td>4.0*</td>
<td>-</td>
<td>0.0*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>2.0*</td>
<td>6.0</td>
<td>-</td>
<td>-</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>-</td>
<td>-</td>
<td>6.7</td>
<td>5.0</td>
<td>6.0*</td>
</tr>
<tr>
<td>NF</td>
<td>Low</td>
<td>2.0*</td>
<td>7.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>-</td>
<td>-</td>
<td>6.5</td>
<td>7.0*</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>6.7</td>
<td>2.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FN</td>
<td>Low</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>5.7</td>
<td>7.1</td>
<td>7.5</td>
<td>-</td>
<td>6.0*</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>7.5</td>
<td>-</td>
<td>5.5</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* single participant data

The data in Table 4.5 indicates relatively inconsistent relationships between WM and accuracy for the three conditions. The most consistent relationship appears to be for the naturalistic condition, in which individuals with relatively higher levels of WM have acquired higher levels of accuracy, although more data would be required to determine the relationships between some of the groups, and therefore the extent of any predictive relationship between WM and naturalistic learning for accuracy. The data for both the NF and FN contexts do not appear to demonstrate any significant patterns that could be interpreted as a relationship between WM and accuracy for these contexts.

WM and accuracy by learning condition: Results for the regression UNIANOVA

The results for the regression experiment testing for predictive relationships between working memory and accuracy for the three conditions indicated a lack of significance for two of the three conditions: the procedure accepted the null hypothesis for the NF and FN conditions. Significance in the relationship was found, however, for the naturalistic condition, with the level at $F(1) = 12.133$, $p. = .007$. The Tukey post-hoc test revealed, however, that significance lies between Average and high levels of WM, and not between either Low and Average or Low and High WM. A lack of significance
was found for experience as a predictor of accuracy (F(1) = .378, p. = .554). The WM and accuracy relationship for the naturalistic condition is depicted in Figure 4.1.

![Figure 4.1: WMC by Accuracy for Naturalistic Learners](image)

**Figure 4.1: WMC by Accuracy for Naturalistic Learners**

### 4.2 Working memory as a predictor of fossilization rate

The examination of potential relationships between working memory and the observed rate of fossilization yielded a number of interesting results. The first part of this section will evaluate the relationship by considering all participants collectively. The second part provides results for the relationship in terms of the different learning contexts. For the purpose of fossilization data, only those learners who reported over four years of experience were included in the analysis. Fossilization rate (FR) is defined as the number of repeated potentially fossilized forms that occur per 100 words of continuous participant discourse. Learners were again grouped into experience categories including a) between four and nine years; b) between nine and fourteen years; c) between fourteen and nineteen years; and d) over 19 years of L2 Russian experience.
As seen with fluency and accuracy data, with all participants combined, WM demonstrates no apparent predictive relationship with the observed rate of potential fossilization (FR). Learners demonstrated the following averages for the FR score: a) learners with between four and nine years of experience demonstrated an average FR of 1.2 for those with an Average level of WM, and 2.0 for those with High WM; b) those with between nine and 14 years averaged a fossilization rate of 1.7 for Average, and 1.5 for High; and c) FR averages for participants with over 19 years of experience were 4.0 for those with an Average level of WM, and 2.3 for those with High WM. The average rate for the Average-ability WM learners with between 14 and 19 years was 2.2. These results appear to indicate no apparent general predictive relationship between WM and potential fossilization for learners of various types of experience. The average fossilization rates (FR) for all participants with different levels of WM ability are provided in Table 4.6

<table>
<thead>
<tr>
<th>Experience (yrs.)</th>
<th>WM Level</th>
<th>FR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-4</td>
<td>Low</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>n/a</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>n/a</td>
</tr>
<tr>
<td>4-9</td>
<td>Average</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>2.0</td>
</tr>
<tr>
<td>9-14</td>
<td>Average</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>1.5</td>
</tr>
<tr>
<td>14-19</td>
<td>Average</td>
<td>2.2</td>
</tr>
<tr>
<td>Over 19</td>
<td>Average</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>2.3</td>
</tr>
</tbody>
</table>
In Table 4.6 the apparent lack of predictive relationship between WM and fossilization is apparent: among participants with similar amounts of experience, fossilization either decreases (learners with between nine and fourteen and over 19 years of experience), or increases (thos with between four and nine years of experience). These results for all participants combined were substantiated by logistic regression. The procedure accepted the null hypothesis for any WM-fossilization relationship \( F(2) = .722, p. = .506 \). Results for an additional regression for both total and naturalistic experience showed no significant relationship between total experience and fossilization with significance level at \( F(2) = 1.15, p. = .349 \). The relationship between the amount of naturalistic experience and fossilization was also not significant \( (F(2) = .489, p. = .625) \).

**WM and observed potential fossilization among the three learning conditions**

A general analysis of the relationship between WM and fossilization reveals variability between the different conditions in terms of any relationship between WM and FR. The results reveal what appears to be a negative correlation between working memory ability and the rate of potentially fossilized forms within the three distinct learning conditions, and the strength of the relationship between WM and FR appears to be highest for learners whose L2 experience has been naturalistic: when participants are divided according to amount of experience, in each experience category NC learners with the highest level of WM exhibited the lowest rate of potential fossilization. This tendency is also somewhat true for both NF and FN learners.

Results for the average fossilization rate for the three conditions among learners who have between four and nine years of experience were as follows: a) NC learners in this category demonstrated an average WM of .77 and an average FR of 2.2; b) results for
NF learners also showed an average FR of 2.2 with the WM average at .76; c) the FN group learners revealed an average FR score of .92 (the WM average was .70). When learners are grouped in terms of level of WM the average FR score for NC learners with High WM was 1.8, and the one learner in this condition who tested as having Average WM obtained an FR score of 2.6. The average FR score for NF learners with High WM was 2.0, while a single participant received a Low WM score and an FR score of 2.5. Lastly, the average FR score for all FN learners was .92, as stated above (FN condition learners all demonstrated an Average level of WM).

The characteristics observed for individual test results demonstrate that: a) the NC learner who achieved the lowest FR score (1.8) also scored the highest in WM of all NC participants in this experience group (.79); b) the two NF learners who received the lowest FR rate scores were within 10% of each other, while both learners tested at a High level of WM; c) among FN learners in this category, all participants demonstrated an Average level of WM, although the learner with the lowest WM score (.62) demonstrated the lowest rate of potential fossilization, with an observed FR of .2. This learner received the highest scores in all components, and demonstrated the most native-like speech. Her experience was also characterized by a very high level of formal instruction as a percentage of overall experience, a subject that is addressed in Section 4.4.

For learners across the three conditions whose L2 experience was between nine and fourteen years, the following averages for potential fossilization were found: a) for NC learners average WM was .71 (Average) and average FR was 2.8; b) NF learners with an average WM of .75 showed an average FR score of 1.1, and c) those learners in the FN group (average WM = .77) demonstrated an average FR score of .82. When learners are
divided by level of working memory, Naturalistic learners with High WM ability averaged an FR of 2.8, while the one NC participant with Low WM received a score of 10.3. As stated above, although NF learners were split between Average and High levels of WM, their WM scores were within .11. The average FR was 1.1 for these learners. Among FN participants, those learners whose WM tested at Average (average of .68) showed an average FR of .75, while the average score for those with High WM was .9.

Individual data within this experience category show that within each learning condition except the Formal/Naturalistic group, the participant with the highest level of WM also demonstrates the lowest rate of potential fossilization. The lowest FR score among NC learners (2.6) was received by the participant with the highest WM, which was .94. Likewise, for NF learners the lowest FR observed was .24, with this participant also showing the highest level of WM (.81). Lastly, among FN learners, the individual with the highest level of WM demonstrated an FR of .9, while the lowest FR score was .7, which was observed for a learner with an average WM level at .66.

Among learners with between 14 and 19 years of L2 Russian learning, the following results were found: a) NC learners showed an average FR of 2.7, while average WM for the NC group was high at .84; b) the NC learner with the highest level of WM (.96) also achieved the lowest FR, with a score of 2.3; c) the one NF learner demonstrated a potential fossilization rate of 1.0 (WM was measured at .71).

The average FR scores for participants with over 19 years of experience were as follows: a) among Naturalistic learners the average FR was 4.4 while average WM was found to be .72; b) FN learners showed an average fossilization rate of 1.4, with average WM measured at .66 (Average level working memory). With participants divided
according to levels of working memory, those NC learners with Average WM showed an average FR of 5.3, while high WM participants had an FR of 2.4. Among FN learners, the lowest level of FR was found to be 1.1. All participants in the FN condition group tested as having an Average level of WM. Overall results for WM and fossilization rate for learners with over four years of experience are shown in Table 4.7.

### Table 4.7 Average FR by level of working memory/experience

<table>
<thead>
<tr>
<th>Context</th>
<th>WM</th>
<th>4-9 yrs</th>
<th>9-14 yrs</th>
<th>14-19 yrs</th>
<th>&gt;19 yrs</th>
</tr>
</thead>
<tbody>
<tr>
<td>NC</td>
<td>Low</td>
<td>-</td>
<td>10.3*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>2.2</td>
<td>-</td>
<td>-</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>-</td>
<td>2.8</td>
<td>2.7</td>
<td>2.4*</td>
</tr>
<tr>
<td>NF</td>
<td>Low</td>
<td>1.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>-</td>
<td>1.1</td>
<td>1.0*</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FN</td>
<td>Low</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>.92</td>
<td>.75</td>
<td>-</td>
<td>1.4*</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>-</td>
<td>.9</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* single participant data

Table 4.7 shows some apparent relationships between the three conditions and rate of fossilization. The rate is first of all lower for the two combined conditions than it is for naturalistic learners. It is also apparent that for the naturalistic context the rate decreases for learners who possess higher levels of WM, while such a relationship does not appear to exist for either NF or FN learners. A third observation is that within the combined conditions learners with similar amounts of experience reveal different rates of fossilization, with lower levels seen among FN learners than for the NF condition.

**WM and FR by learning condition: Results for the regression UNIANOVA**

Results for the regression experiment for WM and fossilization rate yielded mixed results among the three conditions. For the naturalistic context a significance of $F(2) = 30.893, p. = .002$ was found for WM and FR. Since one of the WM level groups had less
than two cases, a post-hoc test was not performed. Significance in the relationship was found for both groups, with p. values of .001 for low working memory learners, and .045 for learners with an Average level of WM (for tabulated results see Appendix C). The additional test for relationships between the amount of naturalistic experience and FR for the naturalistic condition found no significance (F(2) = 1.78, p. = .261). The null hypothesis for a relationship was confirmed for NF and FN conditions, however. Of interest is that for FN learners a relationship was found between formal training and FR (see Section 4.4). It is apparent that WM demonstrates significant predictive validity for fossilization rate for the naturalistic context. Further examination of the relationship with a larger sample might confirm the extent of the WM-FR relationship for naturalistic learners with all potential working memory profiles. The relationship between WM and fossilized forms found in this study for naturalistic learners is depicted in figure 4.2.

![Figure 4.2: WMC by FR for Naturalistic Learners](image_url)
4.3 Working memory-learning condition interactions

The second experiment was designed to test for any possible interactions between working memory and learning conditions. As such, it was important to characterize the amount of experience in a given learning condition, with an especially close examination of the extent of formal training as a percentage of overall experience for the two combined conditions. To test for interactions, a moderated regression was utilized. This procedure examines any moderating effect upon WM by a given learning context. Consequently, the moderated regression may yield results that indicate significant interactive relationships for aspects of proficiency whereas a standard logistic regression may yield no significance between WM and the same measure.

For the moderated regression only data specifically related to the three distinct learning conditions was evaluated, with tests for interaction conducted between WM and both overall experience and the extent of formal experience for the NF and FN contexts. Since the native-likeness measure represents an overall average of the two main components tested (accuracy and fluency), and not a specific aspect of proficiency, the main proficiency measures that were considered necessary for the interaction experiment were accuracy, fluency, and fossilization. However, for the sake of comparison, an additional moderated regression was performed on NL data. Results appear to indicate a significant interaction occurring between WM and one main aspect of proficiency within the naturalistic condition, and one of the components for learners in the FN condition. Significance was also found for the additional regression conducted with NL data for the FN context. No significant interactions were found for the NF learning condition.
For naturalistic learners, results indicate no significant interaction between WM and two of the three proficiency measures analyzed, including accuracy (F(1) = 1.290, p. = .282) and fossilization rate (F(1) = 1.080, p. = .329). A significant interaction between working memory and naturalistic experience was found for NC learners for fluency, however, with a significance of F(1) = 5.704, p. = .038. The Tukey post-hoc test further showed, however, that significance was only found between low and high levels of WM; the level of significance between low and high levels of WMC was p. = .046, whereas the level was p. = .209 between Average and high WMC learners, and .292 between low and Average-level learners (for tabulated results see Appendix D).

With respect to learners who have followed up naturalistic experiences with formal training, the results of the current study indicate no significant interaction between either the amount of naturalistic experience and WM, or the proportion of overall experience consisting of formal experience and WM. The complete lack of any statistically significant relationships between WM and any component of proficiency in the NF condition, whether in terms of either predictive or interactive relationships, is an interesting finding in this study. However, a reversal in the sequence combination of these two learning conditions, with formal training preceding naturalistic learning, appears to indicate a level of interaction between WM and formal experience: for FN learners, significance was found in the interaction between WM and accuracy. Out of curiosity the procedure was also repeated for the measure of the degree to which overall speech is native-like: results for this test demonstrated a significant interaction between WM and the overall proficiency ability or native-likeness score.
For FN learners a significant level of interaction (F(2) = 37.436, p. = .001) was found between WMC and this learning condition in relation to accuracy. It is also interesting to note that when considered separately, the amount of formal experience as a percentage of overall learning experience also shows significance as a predictor of accuracy via the regression analysis: F(4) = 6.401, p. = .033, whereas WM alone does not show a significant relationship (F(1) = .440, p. = .537). The post-hoc analysis was not necessary since the data for WM under the FN condition only includes two groups. A significant level of interaction was found between several levels of WM and different amounts of formal training, ranging from very low to very high percentages of formal experience: between average levels of WM and a very low amount of formal training the significance was p. = .003, between Average WM and an average amount of formal training significance was found to be p. = .045, significance was found between Average WM and very high amounts of formal training at p. = .001, and high WM and high levels of formal training showed an interaction with significance at p. = .003 (tabulated results are presented in Appendix D). The results can be interpreted to mean that the variable of WM does not work independently of formal experience (WM alone shows no significant relationship with learner accuracy, with the level of significance at p. = .537), while it does work together with formal training to affect learner accuracy. These results indicate a fairly robust interaction between working memory capacity and the extent to which overall learner experience in the FN condition is comprised of formal classroom training, a finding that fits in with the patterns observed between both accuracy and the number of potentially fossilized forms and formal experience described in Section 4.1.
Although the measure was not included in the experimental design for the interactive analysis, for the overall measure of proficiency represented by the native-likeness score, results also appear to demonstrate an interactive relationship between WM and the amount of formal experience for the FN condition: the additional analysis indicated a significance of $F(2) = 12.088, p. = .012$. For the NL measure, neither WMC nor the percentage of experience comprised of formal training indicated any independent significant relationship with the NL score (See Appendix D for tabulated results).

**4.4 Learning Conditions and Russian SLP**

When participant data is categorized according to the targeted learning conditions, a number of contrasts between the conditions emerge. An examination of scores achieved by some individual participants reveals several patterns, including distinct differences between contexts in levels attained for the different aspects of proficiency, and what appears to be a relationship between such proficiency components and the amount and timing of formal training. When average scores for aspects of proficiency are examined, changes in accuracy, fluency, native-likeness and fossilization measures are found with increasing amounts of experience. There are differences, however, between the components in terms of the extent of change with experience and also its direction among the different conditions. In the following overview, scores representing fluency, accuracy and native-likeness represent a nine-level scale, with level nine equivalent to native-level proficiency in the ability. A score of 4.3 for fluency or accuracy therefore represents ability between 40 and 50% of native-level proficiency. For the analysis of average test scores, to ensure a minimum number of participants in each subgroup, subcategories related to the amount of learning experience within the different learning conditions were
formed by dividing participants into the following subgroups in terms of years of experience: a) 1-4 years; b) 4-9 years; c) 9-14 years; d) 14-19 years; e) over 19 years of experience. No participant in the current study reported experience in excess of 22 years.

For the two combined learning conditions (the NF and FN groups), an additional procedure was devised in order to try to isolate the potential effects of formal training on each area of proficiency resulting from overall experience for learners in these contexts. To test for this effect, the proportion of overall experience represented by formal learning was quantified with participants further categorized according to the amount of formal training they received as a percentage of overall experience. For this analysis subjects were classified into categories according the percentage of overall experience consisting of formal training, including the following: a) between two and eight percent; b) between eight and fourteen percent; c) between 14 and 22 percent; d) over 22 percent of experience. The analysis included both the extent of formal experience in comparison with the full amount of experience in years and its overall duration in years. The extent of formal training experience as a percentage of overall experience was also included as an additional fixed variable for the statistical regression analyses performed in this study.

**General proficiency or native-likeness resulting from the three learning conditions**

The general proficiency interview score was used along with measures of accuracy and fluency to determine the extent to which participant speech was native-like. Average interview scores for the naturalistic condition ranged from a low of 2.5 (learners with two to four years of experience) to a high of 6.0 (between 14 and 19 years of experience), and the overall average score for the interview was 4.0. The following results for the degree to which participant speech was native-like were found for
naturalistic learners: a) a 3.1 average native-likeness score for participants with between two and four years of experience; b) 5.6 for those with between four and nine years experience; c) 5.0 for those with between nine and fourteen years; d) a 6.6 average for learners who have between 14 and 19 years; and e) an average of 5.0 for learners with over 19 years of L2 Russian experience.

For Naturalistic/formal (NF) learners, the combined native-likeness average scores for participants were a) 5.2 for those with between two and four years; b) 5.0 for those with four to nine years; c) an average native-likeness score of 6.3 for participants with from nine to fourteen years of experience; and d) a score of 6.3 was also achieved by the one participant with between 14 and 19 years of experience. there were no learners within this learning condition with over 19 years of experience.

FN learner averages for the native-likeness score were as follows: a) an average of 6.1 was found for individuals with from two to four years of experience; b) 6.8 for those with between four and nine years; and c) 6.4 for learners with from nine to fourteen years of experience. The one learner with over 19 years of experience achieved a score of 6.3 on the native-likeness measure. Results for the NL or general proficiency score among the three conditions are shown in Table 4.8.

<table>
<thead>
<tr>
<th>Experience (in years)</th>
<th>Native-likeness</th>
<th>NC</th>
<th>NF</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 4</td>
<td>3.1</td>
<td>5.2</td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>4 - 9</td>
<td>5.6</td>
<td>5.0</td>
<td>6.8</td>
<td></td>
</tr>
<tr>
<td>9 - 14</td>
<td>5.0</td>
<td>6.3</td>
<td>6.4</td>
<td></td>
</tr>
<tr>
<td>14 - 19</td>
<td>6.6</td>
<td>(6.3)*</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Over 19</td>
<td>5.0</td>
<td>-</td>
<td>(6.3)*</td>
<td></td>
</tr>
<tr>
<td>Overall Average</td>
<td>5.0</td>
<td>5.7</td>
<td>6.4</td>
<td></td>
</tr>
</tbody>
</table>

* data is for one participant
Participant fluency resulting from the three conditions

For those learners whose learning has occurred under predominantly naturalistic conditions (NC learners), the lowest fluency rating was a score of 1.0, while the highest rating was 9.0 (native-level fluency). The average fluency scores for NC learners was as follows: those with two to four years of immersion in-country received an average score of 4.0; those with four to nine years averaged 7.0; for nine to fourteen years the average was 6.0; an average of 9.0 for fourteen to nineteen years; and those with more than nineteen years of experience averaged 6.7. The overall average fluency for the naturalistic learning condition was 5.75.

Measurements of fluency in terms of overall experience for the NF group demonstrated the following characteristics: The lowest fluency rating in the NF group was 4.5, while the highest fluency score received was 8.0. The average attained fluency score was 6.0 for participants with between two and four years of experience, 6.9 for those with between four and nine years, and 8.5 for learners with between nine and fourteen years of experience. One additional participant in this learning condition possessed between 14 and 19 years of experience, and received a fluency rating of 7.0. The average overall fluency score for the NF group was 7.14.

Fluency score results attained by learners whose L2 Russian experience has been Formal/Naturalistic (FN) range from a low rating of 5.0 to a high score of 9.0. Average fluency related to overall learning experience for this condition was 7.2 for learners with between two and four years of experience, 8.5 for those with between four and nine years, and 7.4 for learners who have between nine and fourteen years of overall experience. One additional participant reported over 20 years of combined learning
experience and achieved a score of 7.0 (approximately 70% of native-level fluency). The overall average fluency for the entire FN group was 7.6. Average Fluency scores in terms of amount of experience for Naturalistic (NC), Naturalistic/Formal (NF), and Formal/Naturalist (FN) conditions are depicted in Table 4.9.

Table 4.9 Learning Conditions: Average Fluency by Experience

<table>
<thead>
<tr>
<th>Years of Experience</th>
<th>NC</th>
<th>NF</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 4</td>
<td>4.0</td>
<td>6.0</td>
<td>7.2</td>
</tr>
<tr>
<td>4 - 9</td>
<td>7.0</td>
<td>6.9</td>
<td>8.2</td>
</tr>
<tr>
<td>9 - 14</td>
<td>6.0</td>
<td>8.5</td>
<td>7.4</td>
</tr>
<tr>
<td>14 - 19</td>
<td>9.0</td>
<td>(7.0)*</td>
<td>-</td>
</tr>
<tr>
<td>Over 19</td>
<td>6.7</td>
<td>-</td>
<td>(7.0)*</td>
</tr>
<tr>
<td>Overall Averages</td>
<td>6.5</td>
<td>7.0</td>
<td>7.6</td>
</tr>
</tbody>
</table>

* one participant

Average accuracy resulting from the three conditions

For learners under strictly naturalistic learning conditions the scores for accuracy found for the elicitation subtest ranged from a score of 1.0 to 9.0. The average accuracy scores for participants were as follows: a) 3.0 for those with between two and four years of experience; b) 6.0 for those with four to nine years; c) 4.5 for learners with nine to fourteen years; d) an average of 5.0 for participants with fourteen to nineteen years; and e) an average score of 4.0 for those with over nineteen years of experience. The overall average accuracy for all participants under naturalistic conditions was 4.5.

For NF learners, accuracy scores ranged from a low score of 1.0 (two participants) to a high score of 8.5 or near-native level (five participants). Average accuracy scores for all NF learners with differing amounts of overall experience were as follows: a) 5.8 for learners with between one and four years; b) 3.8 for those with four to nine years; and c) 6.5 for learners with nine to fourteen years of overall experience. The one learner with 14 to 19 years of experience achieved a score of 7.0 on the accuracy elicitation subtest.
For learners who have followed formal training with naturalistic learning (FN learners), scores for accuracy ranged from a low score of three up to a score of nine (native-level accuracy). Results show the following: a) average accuracy of 6.6 for learners with two to four years of experience; b) 7.1 for those with four to nine years experience; c) 6.5 for participants with between nine and fourteen years of overall experience. The accuracy score for the one learner with over 19 years of experience was 6.0. Average scores for accuracy (AC) and the overall Native-likeness (NL) measure in terms of amount of experience for Naturalistic (NC), Naturalistic/Formal (NF) and Formal/Naturalistic (FN) conditions are depicted in Table 4.10.

Table 4.10 Learning conditions: Average accuracy by experience

<table>
<thead>
<tr>
<th>Experience (in years)</th>
<th>Accuracy</th>
<th>NC</th>
<th>NF</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 4</td>
<td>3.0</td>
<td>5.8</td>
<td>6.6</td>
<td></td>
</tr>
<tr>
<td>4 - 9</td>
<td>6.0</td>
<td>3.8</td>
<td>7.1</td>
<td></td>
</tr>
<tr>
<td>9 - 14</td>
<td>4.5</td>
<td>6.5</td>
<td>6.5</td>
<td></td>
</tr>
<tr>
<td>14 - 19</td>
<td>5.0</td>
<td>(7.0)*</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Over 19</td>
<td>4.0</td>
<td>-</td>
<td>(6.0)*</td>
<td></td>
</tr>
<tr>
<td>Overall Average</td>
<td>4.6</td>
<td>4.9</td>
<td>6.7</td>
<td></td>
</tr>
</tbody>
</table>

* One participant

Occurrence of potentially fossilized forms among the three learning conditions

As seen in the results for accuracy, fluency, and the native-likeness score, among the three conditions distinct differences are also indicated in the average number of potentially fossilized forms. In determining the amount of potential fossilization, data was collected only for those participants who had been actively learning L2 Russian for a minimum of four years. The same procedure was used to categorize participants in calculating the potential Fossilization Rate (FR), and learners were subdivided by amount of experience into the following groups: a) those with between four and nine years of experience; b) learners with from nine to fourteen years; c) those with between 14 and 19
years; and d) participants who had over 19 years of L2 Russian experience. For learners in the NF and FN categories, the potential effect of formal learning conditions on the frequency of potentially fossilized forms was also examined with the same procedure discussed above: Potential fossilization was further analyzed in relation to the amount of formal training experience as a percentage of overall learning experiences. When appropriate, additional scores or averages are provided for individual learners who have had a high percentage of experience as formal training (above 22 percent).

For learners with predominantly naturalistic learning experience the observed Fossilization Rate (FR) ranged between a low of 1.8 and a high of 6 potentially fossilized forms per 100 words. The NC group demonstrated the following average fossilization rates: a) 2.2 for those with between four and nine years of experience; b) 2.8 for learners with from nine to fourteen years; c) 2.75 for those with 14 - 19 years of experience; and d) 4.4 for those participants with over 19 years of overall experience. The overall average rate of potential fossilization for the naturalistic group was 3.17.

For participants whose learning has consisted of naturalistic experiences followed by formal training (NF), the observed average FR in relation to the overall amount of learning experience was as follows: a) An average FR of 1.8 for participants who had between four and nine years of experience; b) 1.1 for those with between 9 and 14 years; and c) 1.0 for participants with between 14 and 19 years of L2 Russian experience. The overall average for potentially fossilized forms within the NF group was 1.6.

Finally, those participants whose L2 Russian experience has been characterized by formal training followed by naturalistic experiences in country (FN) demonstrated the following results for average fossilization rate in terms of overall L2 learning experience:
a) An average FR score of .92 for participants with between four and nine years of L2 Russian experience; and b) .87 FR for those with between nine and fourteen years of experience. One participant in the FN group had over 19 years of experience and exhibited a potential Fossilization Rate of 1.4. The overall average rate of potential fossilization for all participants in the FN group was .93. Average FR in terms of amount of learning experience for the three learning conditions is shown in Table 4.11.

<table>
<thead>
<tr>
<th>Experience (in years)</th>
<th>NC</th>
<th>NF</th>
<th>FN</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 - 9</td>
<td>2.2</td>
<td>1.7</td>
<td>.92</td>
</tr>
<tr>
<td>9 - 14</td>
<td>2.8</td>
<td>1.1</td>
<td>.82</td>
</tr>
<tr>
<td>14 - 19</td>
<td>2.7</td>
<td>(1.0)*</td>
<td>-</td>
</tr>
<tr>
<td>Over 19</td>
<td>4.4</td>
<td>-</td>
<td>(1.4)*</td>
</tr>
<tr>
<td>Overall Average</td>
<td>3.2</td>
<td>1.7</td>
<td>.93</td>
</tr>
</tbody>
</table>

* one participant

**Learning Conditions: Individual results for fluency, accuracy and fossilization**

The differences among the three learning conditions in terms of proficiency are further clarified through an examination of individual proficiency scores. A few key observations can be made about the specific characteristics of individual participants in each learning condition. The first is that the individual with the highest scores for all measures and consequently the most native-like of all participants had been learning L2 Russian for approximately 11 years at the time of testing, and achieved native-level scores in both fluency and accuracy, with a native-likeness score of 8.3 (her overall proficiency interview score was level seven). She was the only subject who achieved a native-level score on the accuracy elicitation task. This participant also produced the lowest rate of potential fossilization among all participants, at .2 forms per 100 words uttered. She began her study of Russian in an intensive formal setting in the US (a strategy used by a number of FN learners). In her case, the duration of intensive formal
study was unusually long – well over one year, and was followed up with immersion experience in Ukraine. It is also interesting that this subject received the fifth lowest WM score of all study participants (.62 or average WM). She is now living once again in the US and is married to a native Russian speaker: immersion in the home has continued.

Interestingly, the second most proficient participant in the study achieved similar scores to the FN learner above (except in terms of fossilization), had a similar amount of L2 Russian experience at testing (12 years), and had learned L2 Russian in a predominantly naturalistic way. This subject also scored native-like in fluency. Her accuracy rating was 8.5 or near-native, and she also scored a level seven on the proficiency interview, yielding an overall native-likeness score of 8.1. There are two significant differences between this learner and the highest proficiency learner discussed above: this participant exhibited a far higher frequency of potentially fossilized forms with a score of 2.6 such occurrences per 100 words, and her working memory span was the second highest recorded among all study participants. By way of comparison, the next most proficient naturalistic learner also achieved a native-level score for fluency with a fairly high level of accuracy (7.0). This second naturalistic subject showed a slightly lower rate of fossilization at 2.3 and exhibited a WMC measured as the highest of all study participants at .96. These two naturalistic learners demonstrated the lowest rates of observed fossilization among all naturalistic participants and also received the two highest overall working memory scores of all participants in this study. In terms of naturalistic learners, it is also interesting that the individual who exhibited the lowest scores of all subjects for all measures of the proficiency interview (1.0 in fluency, level 0
in accuracy and a fossilization rate of 10.3) also received the lowest score in working memory (.37) of all study participants.

Of the 10 most proficient participants, six had followed formal study with naturalistic learning (FN), and after the two learners described above, four of the FN learners were among the most native-like, two of whom had achieved native-like scores for fluency, and two who were near-native on the accuracy measure. All of these participants achieved accuracy scores of at least 7.0. Five of these top FN learners also exhibited potential fossilization rates lower than 1.0, with the average being .83 (one had been studying Russian for less than four years). Of the remaining ten most proficient learners, the seventh highest proficiency participant had begun L2 Russian learning naturally and followed it with extensive formal training (one and a half years). This learner also demonstrated a low level of WM (.52). The ten most proficient participants in this study included two naturalistic learners, two NF learners and six FN learners. The individuals scoring highest in overall proficiency according to the NL average of the major aspects of fluency (FL), accuracy (AC), and the proficiency interview score are depicted in Table 4.12.

Table 4.12 Highest individual proficiency component Scores

<table>
<thead>
<tr>
<th>Rank</th>
<th>Context</th>
<th>Experience (yrs.)</th>
<th>FL</th>
<th>AC</th>
<th>NL</th>
<th>FR</th>
<th>WM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FN</td>
<td>11.2</td>
<td>9</td>
<td>9</td>
<td>8.3</td>
<td>.2</td>
<td>.62</td>
</tr>
<tr>
<td>2</td>
<td>NC</td>
<td>11.4</td>
<td>9</td>
<td>8.5</td>
<td>8.1</td>
<td>2.6</td>
<td>.94</td>
</tr>
<tr>
<td>3</td>
<td>FN</td>
<td>5.8</td>
<td>8.5</td>
<td>8.5</td>
<td>7.5</td>
<td>.8</td>
<td>.77</td>
</tr>
<tr>
<td>4</td>
<td>FN</td>
<td>12.6</td>
<td>8.5</td>
<td>7</td>
<td>7.5</td>
<td>.9</td>
<td>.87</td>
</tr>
<tr>
<td>5</td>
<td>FN</td>
<td>2.2</td>
<td>9</td>
<td>7</td>
<td>7.3</td>
<td>N/A</td>
<td>.87</td>
</tr>
<tr>
<td>6</td>
<td>NF</td>
<td>7.3</td>
<td>8.5</td>
<td>8</td>
<td>7.2</td>
<td>1.0</td>
<td>.52</td>
</tr>
<tr>
<td>7</td>
<td>FN</td>
<td>11.4</td>
<td>9</td>
<td>7</td>
<td>7.0</td>
<td>.8</td>
<td>.71</td>
</tr>
<tr>
<td>8</td>
<td>NC</td>
<td>19.2</td>
<td>9</td>
<td>5</td>
<td>7.0</td>
<td>2.3</td>
<td>.96</td>
</tr>
<tr>
<td>9</td>
<td>NF</td>
<td>10</td>
<td>8.5</td>
<td>7</td>
<td>6.8</td>
<td>.24</td>
<td>.81</td>
</tr>
<tr>
<td>10</td>
<td>FN</td>
<td>6.2</td>
<td>8.5</td>
<td>7</td>
<td>6.5</td>
<td>.9</td>
<td>.65</td>
</tr>
</tbody>
</table>
Table 4.12 demonstrates that of the three contexts, the FN learners in this study have achieved higher overall levels of proficiency, as measured by the native-likeness score, than their NC and NF counterparts. Among the top ten learners, FN participants demonstrate the highest overall scores for accuracy (AC) and the lowest observed fossilization rate (FR) as well as scores for fluency (FL) that are equivalent to levels shown by learners of the other two conditions. Of these learners, FN learners appear to have achieved higher average proficiency with an overall lower average amount of L2 Russian experience (8.2 years) than either naturalistic participants, whose average experience is 15.3 years, or NF learners, who have an average of 8.6 years of L2 experience.

**NF condition: fluency, accuracy, native-likeness and Fossilization by formal training**

To determine the potential influence on SLP of the extent of formal condition experience, learners within the two combined (NF and FN) conditions were further divided according to the percentage of overall experience consisting of formal training. When categorized in this fashion, some distinctly different patterns are revealed for NF and FN learners. Such results will first be described for NF learners, followed by findings related to the FN context, including a description of highest individual results for the two contexts. This subsection concludes with a report of individual accuracy-related results in terms of the amount of formal instruction for the two combined conditions.

When characterizing fluency in the NF condition in terms of the percentage of overall experience comprised of formal training, average levels of fluency were as follows: a) 8.0 for those for whom formal training amounts to between two and eight percent of experience; b) 7.3 for those with between eight and fourteen percent formal
training; c) 6.7 for those whose formal training was between 14 and 22 percent of the total; and d) an average of 6.0 for those with over 22 percent of overall experience comprised of formal training. Greater proportions of formal training are evident with a decreasing amount of overall L2 experience within the NF learning condition.

NF condition results for accuracy and overall native-likeness in terms of the percentage of experience consisting of formal training indicated the following: a) an average of 5.0 was attained by learners with between two and eight percent of overall experience as formal training; b) 4.8 by those with between eight and fourteen percent; c) an average accuracy of 5.5 was achieved by learners with between 14 and 22 percent; and d) a 5.2 average was found for those whose formal training comprised greater than 22 percent of their total experience. The greatest amount of formal training as a percentage of the total was 27 percent. It is also important to note that none of the NF participants who reported over 22 percent formal experience possessed more than 4.3 years of total L2 Russian experience, with the minimum amount being 2.3 years. These same learners exhibited the following average scores for the combined native-likeness measure: a) 6.0 at between two and eight percent formal training; b) 5.5 for between eight and fourteen percent; c) 5.5 for between 14 and 22 percent; and d) an average of 5.0 for those with above 22 percent of experience comprised of formal training.

The analysis of the effect of formal instruction on potential Fossilization rate for NF learners obtained the following results: a) an average FR of 1.7 for learners whose formal learning experience comprised between two and eight percent of their total experience; b) 1.9 average FR for those for whom formal training comprised between eight and fourteen percent of their experience; c) an average FR of 1.5 for those with between 14 and 22
percent formal training; and d) 1.4 average FR for those whose formal training amounted to more than 22 percent of total experience at the time of testing.

In order to confirm the results for the analysis of formal instruction for accuracy-related measures, including FR, NF data were also categorized according to the total amount of formal experience in addition to the percentage of learning it represents. Results showed that for NF learners the subjects with the greatest amount of formal experience (an average of 1.6 years) also had the highest percentage of experience consisting of formal training (16%). These learners also achieved the highest average in both accuracy (7.25) and the lowest average rate of fossilization (.81). In contrast, learners with less overall formal experience (an average of .64 years and 11% of their overall experience) demonstrated an average accuracy of 3.4 and an average FR of 2.1.

**FN condition: fluency, accuracy, native-likeness and FR by formal training**

FN condition results for measured fluency in relation to the extent of formal training as a percentage of overall experience revealed the following characteristics: a) an average score of 7.0 was achieved by learners for whom between two and eight percent of total learning was formal in nature; b) an average fluency of 8.0 was achieved by learners with formal training of between eight and fourteen percent; c) 7.5 was the average attained by those for whom formal training has comprised between 14 and 22 percent of experience; and d) those participants who reported over 22 percent of their learning as formal classroom-related experience achieved an average fluency score of 7.4. The average fluency scores related to the percentage of experience comprised of formal instruction for both NF and FN learning conditions are presented with data for all major proficiency components in Tables 4.13 and 4.14.
In terms of average participant accuracy and the native-likeness score, when categorized according to the percentage of overall experience consisting of formal classroom training, the data show the following results for FN learners: a) 5.0 for between two and eight percent of total experience as formal training; b) 7.5 for those with between eight and fourteen percent; c) 7.0 for those with between 14 and 22 percent; and d) a 7.0 average for those with formal training at over 22 percent of total experience.

Participants in this last group averaged between two and 5.8 years of total L2 Russian experience, with the percentage of formal training ranging between 30 and 48 percent of total L2 experience. The average proficiency interview scores for the same FN learner groupings were: a) 5.0 (between two and eight percent formal training); b) 4.5 (between 8 and 14 percent); c) 5.7 (between 14 and 22 percent); d) 4.8 (more than 22 percent).

Finally, the average native-likeness scores for FN learners was a) 5.6 (two to eight percent), b) 6.7 (between eight and fourteen percent), c) 6.8 (between 14 and 22 percent), and d) 6.4 (learners with over 22 percent of total experience as formal training).

In terms of the effect of formal training experiences on potential fossilization, FN participants demonstrated the following averages: a) 1.4 FR for learners for whom formal training comprised between two and eight percent; b) .75 average FR for those with between eight and fourteen percent formal training; c) .67 was the average for learners with from 14 to 22 percent of L2 experience characterized as formal training. One participant reported experience in formal training amounting to 31% of his total L2 Russian experience, and demonstrated a potential fossilization rate of .80. The total amount of experience for this learner was 5.8 years.
A comparison of learners according to the overall amount of formal experience showed that FN learners with a lower average amount of total formal training (.78) as well as a lower percentage of overall experience consisting of formal experience (9.7%) had an average accuracy of 6.7 and an average FR of .93 FR. Those whose average total formal experience was 1.9 years with an average of 15% of total experience as formal training also demonstrated an average accuracy of 6.7 and had an average FR of .94. Average scores for fluency (FL), accuracy (AC), native-likeness (NL), and fossilization rate (FR) for NF and FN learners in terms of the amount of formal training (FT) as a percentage of total experience are shown in tables 4.13 and 4.14, respectively.

Table 4.13 NF Proficiency Averages: Formal Training as a Percentage

<table>
<thead>
<tr>
<th>FT (as % of total)</th>
<th>Avg. Experience (years)</th>
<th>FL</th>
<th>AC</th>
<th>NL</th>
<th>FR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 – 8</td>
<td>12.3</td>
<td>8.0</td>
<td>5.0</td>
<td>6.0</td>
<td>1.7</td>
</tr>
<tr>
<td>8 – 14</td>
<td>6.1</td>
<td>7.3</td>
<td>4.8</td>
<td>5.5</td>
<td>1.9</td>
</tr>
<tr>
<td>14 – 22</td>
<td>7.8</td>
<td>6.7</td>
<td>5.5</td>
<td>5.5</td>
<td>1.5</td>
</tr>
<tr>
<td>Above 22</td>
<td>3.1</td>
<td>6.0</td>
<td>5.2</td>
<td>5.0</td>
<td>1.4</td>
</tr>
</tbody>
</table>

Table 4.14 FN Proficiency Averages: Formal Training as a Percentage

<table>
<thead>
<tr>
<th>FT (as % of total)</th>
<th>Avg. Experience (years)</th>
<th>FL</th>
<th>AC</th>
<th>NL</th>
<th>FR</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 – 8</td>
<td>15.2</td>
<td>7.0</td>
<td>5.0</td>
<td>5.6</td>
<td>1.4</td>
</tr>
<tr>
<td>8 – 14</td>
<td>12.6</td>
<td>8.0</td>
<td>7.5</td>
<td>6.7</td>
<td>.75</td>
</tr>
<tr>
<td>14 – 22</td>
<td>9.2</td>
<td>7.5</td>
<td>7.0</td>
<td>6.8</td>
<td>.67</td>
</tr>
<tr>
<td>Above 22 %</td>
<td>3.2</td>
<td>7.4</td>
<td>7.0</td>
<td>6.4</td>
<td>.80*</td>
</tr>
</tbody>
</table>

* data for FR is from one participant

Individual results for NF and FN conditions by amount of formal training

Individual results also show some distinct similarities and differences between the combined learning conditions. Both contexts were similar in that the four subjects with the highest amount of formal instruction (both overall and as a percentage of experience) also demonstrated the lowest fossilization rates and some of the highest accuracy scores.
In the NF context the learner with the greatest amount of formal experience (over two years intensive training) had the lowest FR (.24), while in the FN context the individual with the greatest amount (about two years of intensive study) achieved the second lowest FR (.7) of all FN learners. In both contexts individual and average data indicate a negative correlation between the amount of formal experience and fossilization. There is a distinct difference in individual performance for accuracy-related measures: with similar amounts of formal learning experience, FN learners have achieved both higher rates of accuracy and demonstrate a lower occurrence of fossilization than those whose learning was characterized as Naturalistic/formal, as depicted in Tables 4.15 and 4.16.

Table 4.15 NF Learners: Accuracy-related Scores by Formal Training (FT)

<table>
<thead>
<tr>
<th>FT (years)</th>
<th>FT (percentage)</th>
<th>AC</th>
<th>FR</th>
<th>NL</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4</td>
<td>.22</td>
<td>7.0</td>
<td>.24</td>
<td>6.8</td>
</tr>
<tr>
<td>1.6</td>
<td>.22</td>
<td>8.0</td>
<td>1.0</td>
<td>7.2</td>
</tr>
<tr>
<td>1.4</td>
<td>.20</td>
<td>6.0</td>
<td>1.0</td>
<td>5.7</td>
</tr>
<tr>
<td>.94</td>
<td>.05</td>
<td>7.0</td>
<td>1.0</td>
<td>6.3</td>
</tr>
<tr>
<td>.98</td>
<td>.23</td>
<td>2.0</td>
<td>2.5</td>
<td>5.0</td>
</tr>
</tbody>
</table>

Table 4.16 FN learners: Accuracy-related scores by formal training (FT)

<table>
<thead>
<tr>
<th>FT (years)</th>
<th>FT (percentage)</th>
<th>AC</th>
<th>FR</th>
<th>NL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2</td>
<td>.19</td>
<td>9.0</td>
<td>.2</td>
<td>8.3</td>
</tr>
<tr>
<td>1.8</td>
<td>.31</td>
<td>8.5</td>
<td>.8</td>
<td>7.5</td>
</tr>
<tr>
<td>2.0</td>
<td>.14</td>
<td>8.0</td>
<td>.7</td>
<td>6.3</td>
</tr>
<tr>
<td>1.3</td>
<td>.11</td>
<td>7.0</td>
<td>.8</td>
<td>7.0</td>
</tr>
<tr>
<td>.44</td>
<td>.07</td>
<td>7.0</td>
<td>.9</td>
<td>6.5</td>
</tr>
</tbody>
</table>

Tables 4.15 and 4.16 demonstrate distinct differences between the two combined learning conditions. While the five highest accuracy-related scorers among NF learners demonstrate lower fossilization rates with increasing overall amounts of formal training, the overall percentage or amount does not appear to be as important a factor for FN learners, who also demonstrate considerably lower FR with less formal training. FN
condition learners have achieved higher scores for both accuracy-related measures and the general proficiency score represented by the native-likeness measure. The combined conditions also demonstrate some apparent long-term advantages for learners who include formal training as a strategy for learning L2 Russian, with the sequence of formal and naturalistic experience also playing a potentially significant role.

**Fossilization for individual NF and FN learners by amount of formal training**

A consideration of individual test scores related to potential FR in terms of the amount of experience comprised of formal training appears to reveal a pattern pertaining to the potential effect of the sequencing of the two main types of learning (formal and naturalistic). For the NF sequence (naturalistic followed by formal training), although the average FR does decrease slightly from 1.6 for learners with between four and fourteen percent to an average of 1.4 for those with between 14 and 23 percent of total experience comprised of formal training, results demonstrate that the percentage of formal training does not appear to have a significant impact on rate of fossilization over time. Learners who reported a higher percentage of experience as formal training nonetheless exhibit either high, average or low rates of potential fossilization, and the participant in this learning context who showed the highest FR also reported a fairly high amount of formal training (15 percent), while the participant with the lowest fossilization rate also reported a high amount of formal training (22 percent). Working memory also does not appear to explain the potential fossilization rate within this learning condition (see results in Section 4.2). The results for NF and FN individual FR scores in relation to formal training as a percentage of overall learning experience are shown in Tables 4.17 and 4.18.
Table 4.17 NF Learners: FR by formal training (as a % of experience)

<table>
<thead>
<tr>
<th>Participant</th>
<th>Experience (years)</th>
<th>Formal (percent)</th>
<th>WM</th>
<th>FR</th>
</tr>
</thead>
<tbody>
<tr>
<td>035</td>
<td>6.7</td>
<td>.04</td>
<td>40</td>
<td>2.0</td>
</tr>
<tr>
<td>038</td>
<td>12.2</td>
<td>.05</td>
<td>33</td>
<td>2.0</td>
</tr>
<tr>
<td>016</td>
<td>18</td>
<td>.05</td>
<td>34</td>
<td>1.0</td>
</tr>
<tr>
<td>020</td>
<td>6.3</td>
<td>.08</td>
<td>42</td>
<td>2.0</td>
</tr>
<tr>
<td>002</td>
<td>6.0</td>
<td>.10</td>
<td>39</td>
<td>1.8</td>
</tr>
<tr>
<td>001</td>
<td>5.8</td>
<td>.15</td>
<td>38</td>
<td>2.5</td>
</tr>
<tr>
<td>025</td>
<td>7.0</td>
<td>.20</td>
<td>30</td>
<td>1.0</td>
</tr>
<tr>
<td>026</td>
<td>7.2</td>
<td>.22</td>
<td>25</td>
<td>1.0</td>
</tr>
<tr>
<td>032</td>
<td>11</td>
<td>.22</td>
<td>39</td>
<td>.24</td>
</tr>
<tr>
<td>033</td>
<td>4.3</td>
<td>.23</td>
<td>24</td>
<td>2.5</td>
</tr>
</tbody>
</table>

Table 4.18 FN Learners: FR by formal training (as a % of experience)

<table>
<thead>
<tr>
<th>Participant</th>
<th>Experience (years)</th>
<th>Formal (percent)</th>
<th>WM</th>
<th>FR</th>
</tr>
</thead>
<tbody>
<tr>
<td>017</td>
<td>8.8</td>
<td>.02</td>
<td>.75</td>
<td>1.8</td>
</tr>
<tr>
<td>010</td>
<td>21.6</td>
<td>.07</td>
<td>.66</td>
<td>1.4</td>
</tr>
<tr>
<td>029</td>
<td>6.2</td>
<td>.07</td>
<td>.65</td>
<td>.9</td>
</tr>
<tr>
<td>028</td>
<td>11.3</td>
<td>.11</td>
<td>.71</td>
<td>.8</td>
</tr>
<tr>
<td>027</td>
<td>14</td>
<td>.14</td>
<td>.66</td>
<td>.7</td>
</tr>
<tr>
<td>013</td>
<td>12.5</td>
<td>.16</td>
<td>.87</td>
<td>.9</td>
</tr>
<tr>
<td>039</td>
<td>6.2</td>
<td>.19</td>
<td>.62</td>
<td>.2</td>
</tr>
<tr>
<td>014</td>
<td>9.0</td>
<td>.22</td>
<td>.85</td>
<td>.9</td>
</tr>
<tr>
<td>009</td>
<td>5.8</td>
<td>.31</td>
<td>.77</td>
<td>.8</td>
</tr>
</tbody>
</table>

A comparison of the two combined conditions reveals distinct differences in the FR measure. Unlike the pattern seen for individual data for the NF context, the opposite sequence of the two conditions (FN) does appear to have an effect on the rate of potential fossilization. In this learning context average FR for learners with between two and fifteen percent was 1.2, while learners with between 15 and 30 percent of experience comprised of formal training showed an average FR of .7. The FN learners with the highest percentage of formal training consistently achieved low FR scores (below 1.0), while learners with the lowest amount of formal training (below 10 percent) exhibit the highest FR scores. Although the learner with the lowest WM score (.62) also achieved the
lowest rate of potential fossilization, WM also appears to demonstrate a potential relationship within this condition, with average WM for those who showed a high FR lower (.69) than for those with lower fossilization rates (.78) (see results in Section 4.3).

**Results for the formal experience regression experiment**

Of particular interest is the finding that while for FN condition participants’ working memory demonstrates no predictive relationship with the rate of fossilization (F(1) = 3.011, p. = .133), significance was found for the relationship between the amount of formal experience and FR, with a significance of F(1) = 9.006, p. = .024. No such significance was found for NF learners, however. For FN learners there was also no apparent significance in the predictive relationship between the amount of naturalistic experience and FR. These findings appear to partially substantiate the third hypothesis: at least for FN learners, the amount of formal training appears to directly correlate with observed fossilization rate. The relationship between formal training and fossilization rate for FN learners is depicted in Figure 4.3 (see Appendix D for tabulated results).

![Image](image.png)

**Figure 4.3: Observed FR by Formal Training for FN learners**
Figure 4.3 depicts the relationship between amounts of formal training as a percentage of overall experience and the rate of fossilization. To generate this figure from the regression analysis, FN participants were classified further according to five categories of the percentage of overall experience comprised of formal training, including: 1) 2 – 7 percent; 2) 7 – 12 percent; 3) 12 – 17 percent; 4) 17 – 22 percent; and 5) over 22 percent of experience.

4.5 Research questions: A summary of the results

Research Question and Hypothesis #1

The first research question and first hypothesis were related: It was predicted that WM would demonstrate stronger positive correlations with attained levels of proficiency for naturalistic learning than for conditions that rely heavily on formal instruction. The results show that while working memory was found to be a significant predictor for the two accuracy-related aspects of proficiency, accuracy and fossilization rate, under naturalistic learning experiences, no significant direct predictive relationship was found for any aspects of proficiency for either the NF or the FN conditions. The hypothesis was therefore confirmed.

Research question and hypothesis #2

The second hypothesis, which relates to the second research question, predicted differential effects in the interaction of working memory with the different learning conditions in respect to both overall proficiency (or native-likeness), and the specific aspects of the IL grammar, including accuracy and fluency. The results of the experiment produced three significant interactions between working memory and learning conditions: with the naturalistic condition for fluency, and with the amount of formal
instruction in the FN condition for both accuracy and overall proficiency. These results can be interpreted as confirmation of the second hypothesis: since the interactions are related to different conditions and for different component abilities, including an interaction found with the FN condition for overall proficiency, there is an apparent differential relationship between WM and types of learning experience for fluency, accuracy and overall level of proficiency.

**Research question and hypothesis #3**

The third hypothesis and underlying research question related to the effect of formal experience on the presence of potentially fossilized forms in the IL grammar. A negative correlation was predicted between the amount of formal learning and the presence of potentially fossilized forms. A significant effect of the sequencing of naturalistic and formal learning contexts on fossilization was also predicted. It is evident that the first part of this hypothesis was confirmed: the experiment found considerably lower rates of potentially fossilized forms for learners in the two combined conditions than for naturalistic participants. A negative correlation between the amount of formal training as a percentage (and as a total) was also seen, although significance was found for only the FN condition. These results confirm the first part of the hypothesis. However, the second aspect of this hypothesis (i.e. that the sequencing of the two main conditions would have a significant effect) was neither confirmed nor disconfirmed. This is due to the fact that the experiment was not adequate to the task of establishing any relationships between sequencing and fossilization: although FN learners demonstrated considerably lower rates for approximately equivalent amounts of formal experience, the experiment did not provide any mechanism to examine the impact of sequencing. Additional
research will be necessary to confirm if there is a significant effect of sequencing on the rate of potentially fossilized forms.
CHAPTER V: DISCUSSION

The following discussion of experimental results follows the same general format of Chapter Four and includes four major sections. The chapter begins with an overview discussion of the data for all study participants without consideration of learning conditions. Section 5.1 is concerned with the results found for the predictive relationships between working memory and general proficiency, fluency and accuracy, including the findings for all participants and the three conditions. Section 5.2 discusses the results for the extent to which WM predicts learner fossilization, including findings for the WM-FR regression experiment. Section 5.3 is devoted to a description of the findings produced by the interaction experiment testing for the interactive relationships between WM and tested components of proficiency. Lastly, Section 5.4 provides a discussion of the findings related to the relationships between the three learning conditions and targeted components of proficiency and potential fossilization related to 1) the amount and type of learning experience, including individual and average scores for the targeted aspects of proficiency, and 2) for the two combined conditions, the effect on proficiency of the amount and timing of formal training. This last section includes a discussion of the results for the formal experience regression experiment for NF and FN contexts.

Overview of the proficiency data for all study participants

Apart from specific learning condition results, the overall participant component score results seen in this study demonstrate a fairly consistent pattern for all learners
represented in the data: the average score for each tested aspect of proficiency increases with the amount of experience until approximately the twelve-year mark, at which point scores appear to either level off or decrease, depending on the aspect of proficiency. This average reversal of proficiency appears to constitute a type of upper ceiling or ‘barrier’ to more advanced proficiency. This barrier to advanced proficiency is variable between the different tested components, and following the point of maximum average proficiency there appears to be a sharp decrease in average fluency, accuracy, and the combined native-likeness score. The observed rate of fossilization appears to show a more gradual change over time, and the rate increases progressively: unlike the other measures, there is no apparent improvement preceding an increase in the fossilization rate.

**Fluency, accuracy, native-likeness, and fossilization by amount of experience**

For all participants average fluency increased by 2.6 for learners with between two and four years of experience to approximately 85% of native-level fluency for learners with between eight and twelve years of L2 Russian experience. Between the 12 and 20-year point, average fluency then dropped steadily down to about 67% of native-level fluency for learners with over 20 years of experience in this data. This pattern is repeated for average accuracy, which improved about 10% from levels seen among less experienced learners up to approximately 69% of native-level accuracy for learners with between eight and twelve years of experience. Average accuracy then tapers off at first gradually by about 9% for learners with 16-20 years of experience, and then dramatically to about 40% of native-level for learners with over 20 years of experience. It is, of course, unremarkable that the combined native-likeness score, which is an average of the three scores for fluency, accuracy and general proficiency, similarly shows a gradual
increase for learners with from 12 to 16 years of experience, but then drops off by about 16% for those with over 20 years.

Across learning conditions the pattern observed for the components of proficiency mentioned above is repeated for the observed average number of potentially fossilized forms measured for participants with more than four years of L2 Russian experience. The average FR decreased slightly by .1, and then for those learners with more than 12 years of experience, FR increased progressively by about .3 up to the 20-year-experience mark. The rate then increased greatly to 4.0 for learners with over 20 years of experience. In terms of the points in time when the change appears to occur, the pattern for FR fairly closely mirrors the overall participant data for accuracy which at first declined gradually following the twelve year point and then abruptly for learners with over 20 years of experience. Since the observed fossilization rate is closely related to accuracy, it is likely that the rate is influenced by the same processes or factors that affect participant accuracy over time.

The patterns described above appear to indicate that on average regardless of the type of participant experience or their working memory span, the L2 Russian learners in this study have experienced a plateau effect in their development of proficiency: there appears to be a barrier to continued improvement that occurs following the twelve-year-experience mark. Additionally, the results of this study also indicate no concrete correlation between length of experience and proficiency in terms of either fluency or accuracy. These findings were confirmed by means of a logistic regression which demonstrated no significant predictive relationship between either amount of overall experience or the amount of naturalistic experience and any of the components tested,
including the rate of potential fossilization. An additional logistic regression related to any relationship between WM and observed measures of SLP also did not show any significant predictive correlation between WM and any of the component averages or observed fossilization rate. Such findings appear to confirm a type of barrier to sustained progress that functions independently of the amount of experience or WMC possessed by learners.

The observation of a peak in overall proficiency, at which many learners tend to ‘plateau’ or stagnate is generally well supported in the literature and has been one of the observations cited in favor of the perspective that the attainment of true native-like proficiency is generally either extremely rare or potentially impossible for adult L2 learners (e.g. Abrahamsson and Hyltenstam, 2000). Research conducted by Rifkin (2005) demonstrates what he refers to as a ‘ceiling effect’ that exists just below an advanced level of proficiency for both stateside college programs and immersion programs overseas. In terms of more naturalistic types of learning experience, Rifkin (2005) concluded that the barrier to higher proficiency attainment is potentially related to the nature of naturalistic learning conditions. Similar to the results for all participants in this study, in research designed to investigate the effects on proficiency of the amount of L2 Russian study in terms of overall experience, Thomson (1996) found no definite correlation between amount of experience and attained levels of proficiency. Yet contrary to this study’s results for all participants, previous research has shown significant correlations between working memory and the major aspects of proficiency, including fluency, accuracy, and lexical density (e.g. Carpenter and Just, 1989; Daneman, 1991; Daneman and Carpenter, 1980; Fortkamp, 1999).
It is important to point out that although no significance was found for a relationship between working memory and any of the tested aspects of proficiency for all participant data, past WM studies that have shown significant correlations had a far more narrowly focused subject pool in terms of learning conditions and amount of experience: subjects who are tested are typically involved in the same college level classes or immersion programs. With the diverse backgrounds of participants in the present study, it is not surprising that WM did not demonstrate validity as a predictor for all the participants taken together and it is important to remember that the goals of this study included an examination of the differences between the targeted learning contexts in proficiency attainment, any predictive significance of WM for aspects of proficiency within the different conditions, and potential interactions between WM and individual learning conditions to produce major aspects of proficiency and rate of fossilization. The most interesting aspect of the data for all participants is the apparent inability for most learners to advance in proficiency to a native-like level.

The results appear to support the existence of a barrier to more advanced proficiency for most learners, and it is possible that, as Rifkin (2005) asserts, it is related to the specific characteristics of naturalistic learning. A number of researchers have found that, unlike both implicit and formal types of learning, naturalistic acquisition is predominantly related to semantic processing rather than the acquisition of morphosyntax (e.g. Reber and Allen, 2000; Robinson, 2002b). According to Reber and Allen (2000) acquisition under incidental conditions relates to learning that occurs in a way that is “unrelated to gross measures of high-level cognitive function” (p. 238). This characteristic of naturalistic acquisition may explain the general tendency for accuracy-
related scores to potentially drop off after the 12-year mark, with an accompanying increase in potential fossilization over the same period.

Such an effect on accuracy-related data among the study’s participants makes sense especially in light of both the nature of naturalistic learning and the extensive amounts of time that had been spent learning naturally on the part of the participants who had been immersed the longest in-country. It appears that the longer one is immersed, the lower the performance in all areas tested in this study. The relationship between naturalistic conditions and accuracy is further clarified when participant data is categorized and analyzed in respect to the specific learning conditions targeted in this study (see section 5.1).

A number of studies have helped to elucidate the difficulties associated with achieving a level of native-like proficiency. Some research (e.g. Oyama, 1976; Patkowski, 1980) indicates the existence of a gradual decline in ability over a period of years that begins abruptly and continues unabated in spite of any increase in the degree to which learners are exposed to input: increased L2 exposure appears to have no effect on the decline. Such studies would indicate a ‘sensitive’ rather than ‘critical’ period following puberty. In their landmark study on the Critical Period, Johnson and Newport (1989) themselves state that “there is a gradual decline in language learning skills over the period of an ongoing maturational growth and a stabilization of language skills at a low but variable level of performance at the final mature state” (p. 97). The bulk of the research, however, examines such a decline almost strictly in terms of learner accuracy, and although a potential relationship or interaction between naturalistic learning and
accuracy may help to explain this aspect of the IL grammar, it does not explain lower average scores for fluency, which one might expect to increase steadily over time.

Although a number of studies have analyzed learner fluency in relation to measures of working memory (e.g. Daneman, 1991; Fortkamp, 1999, 2000; Mota, 2003; Mota and Bergsleithner, 2007; Weissheimer and Mota, 2011), much of the SLA research to date emphasizes adult proficiency or ultimate attainment typically in terms of accuracy in morphosyntax rather than participant fluency, with a reliance on grammaticality judgments (e.g. DeKeyser, 2000). Proficiency studies have also been predominantly preoccupied accuracy, lexical usage, or pronunciation (e.g. De Jong and Van Ginkel, 1992; McNamara, 1990), and few studies have focused attention on the characteristics of learner fluency in terms of ultimate attainment. For standard proficiency ratings the aspect of proficiency that tends to be the overriding factor is grammatical accuracy (Iwashita, Brown, McNamara and O’Hagan, 2008).

Further research into the nature of long-term acquisition beyond the 12-year mark will be necessary to confirm whether such a barrier to fluency is a common trait of proficiency over the course of time in different L2 contexts, with emphasis on exploring potential causes. One seemingly counterintuitive possibility might be that the fluency plateau effect is related to specific cognitive abilities relied upon by adult learners to develop fluency and that the nature of naturalistic learning somehow interferes with or moderates their operation, which would potentially cause the barrier to exist for those learners with lower levels of the necessary cognitive resources, whereas higher-ability learners are able to attain native-like fluency. An aptitude-related cause of this type could potentially explain the apparent ability of some individual learners to overcome this
barrier and retain fluency past the 20-year mark. More research in this area is clearly needed.

5.1 WM regression experiment: Major proficiency components

This section provides a discussion of the results for the predictive analysis of working memory for the main components targeted in this study. The first part is devoted to the observations that were made for all study participants regardless of the specific learning conditions. Following this part, the section proceeds with a discussion of the results for the distinct learning conditions, including the findings related to the WM regression experiment.

Measured proficiency for all participants by levels of working memory

Although a considerably different picture of potential relationships emerges when learners are categorized according to the three individual learning conditions and amount of experience, the results for all participant data taken collectively do not appear to support a predictive relationship between working memory and the major components of proficiency targeted in this study; outside the respective learning contexts, WM demonstrates at best a limited predictive validity for the major aspects of proficiency. As noted above, average fluency increases for all participants until approximately the 12-year mark, when a decrease in fluency is observed. Yet when working memory is compared with average measures of fluency for different categories of experience, one striking aspect of the data in this study is that average participant fluency appears to decrease overall with increasing levels of working memory. Learners demonstrating average levels of WM scored high in fluency in some experience categories, while others with high WM ability scored lower, and vice versa. The only experience category that
appeared to show any potential predictive relationship between WM and fluency consisted of learners who had over 19 years of L2 Russian experience. For these learners average fluency for average levels of WM was 69% of native-level fluency, while those with high WM ability averaged 80%. Based on these results, at the very least it can be said that there is no apparent independent relationship between fluency and working memory seen in the data for all participants. The data for accuracy show a similar relationship.

Overall average scores for accuracy and the fossilization rate also demonstrate no apparent predictive relationship between WM and these aspects of proficiency, with mixed results for the different experience groups. For example, learners with two to four years of experience appear to demonstrate that higher levels of WM are associated with higher scores in accuracy. For the next experience category, however, the pattern reverses with the average accuracy for average-level WM learners found to be 6.8, while those with higher WM ability averaged 3.2 on the accuracy elicitation subtest. The average FR for these learners was also higher than the lower WM learners. A similar pattern exists for accuracy among those with between nine and fourteen years of experience, while fossilization for these learners shows the opposite trend, and improves with higher levels of WM. Learners with over 19 years experience show an increase in accuracy for those with higher levels of WM, while FR also improves with higher levels of WM. The data therefore appears to demonstrate a total lack of any independent predictive validity of WM for either accuracy or the occurrence of fossilized forms when all participants are grouped together.
The results appear to show that working memory is a poor independent predictor of proficiency for any of the measures elicited in this study. In fact, these results appear to indicate that for all participants taken together the amount of overall experience is a significantly better predictor of proficiency measures than working memory, with overall gains in fluency, accuracy, and the combined native-likeness measure with an increasing amount of experience up to the twelve-year mark. Experience regardless of learning conditions also appears to be a better predictor of the occurrence of potentially fossilized forms, with an increase in such forms occurring with increasing amounts of overall experience (a negative correlation). Lastly, these overall findings related to WM for all participants were confirmed by means of logistic regression, which also indicates a lack of significance in any predictive relationship between amount of experience and the respective measures of proficiency.

Taken independently, the results for all participants appear to contradict previous research that shows strong correlations between WM and all measures of proficiency, including fluency, accuracy and complexity. However, it is important to remember that the majority of previous WM research showing such correlations with proficiency (e.g. Daneman, 1992; Fortkamp, 1999; Harsuiker and Barkhuysen, 2006; Mizera, 2006; Mota, 2003) was conducted among learners within similar learning conditions, such as college FL classrooms or immersion programs. As Robinson (2001, 2002b) has pointed out, the situation is evidently far more complex than a straightforward correlation regardless of learning context. A number of studies have found significant correlations between various cognitive abilities and the learning context, with some cognitive resources implicated in acquisition that occurs under formal conditions (e.g. analytic abilities such
as inductive language learning ability), and others that appear to be associated with more implicit or incidental conditions such as naturalistic learning.

Studies by Harley and Hart (2002) and Ross, Yoshinaga and Sasaki (2002) have produced evidence that learners rely on different cognitive abilities in different contexts, while other research has shown that younger learners tend to rely on memory abilities, while older learners depend on analytic types of aptitude (e.g. Skehan, 1986; DeKeyser, 2000; Harley and Hart, 2002; Wesche, 1981). Such findings point out the disadvantages of examining WM-SLP correlations without consideration of the type and nature of learning experiences. Furthermore, the analysis of WM-proficiency relationships for all participants also demonstrates the importance of discovering the potential interactions between learning contexts and working memory.

**Proficiency component scores by level of working memory and learning context**

In general, the results related to the independent variables of working memory and learning conditions and their potential relationship with proficiency-related dependent variables reveal several patterns. When participants who possess similar amounts of experience and levels of WM are compared across the three learning conditions, the first pattern that emerges is that within each learning condition and for each experience-related subgroup those participants who possess the highest levels of WMC have also usually acquired the highest levels of both general proficiency (native-likeness) and accuracy, though the pattern of relationship appears to be inconsistent. The attained fluency is also often, though not consistently, the highest. A number of possible relationships are apparent for two of the three learning conditions, including the following: 1) although an overview of the data does appear to show a correlation between
WM and the overall proficiency measured by the native-likeness score, again no significance was found in the regression analysis for any of the conditions; 2) there does not appear to be any predictive relationship between WM and fluency in any of the learning conditions; and 3) there were mixed results for the relationship of WM and accuracy, with an apparent significant predictive relationship between WM and accuracy for the naturalistic condition, and although the data do appear to show a correspondence between WM and accuracy for the combined learning conditions, no significance was found for any predictive relationship.

**Working memory and overall proficiency by type of learning experience**

The potential relationship between working memory and scores for overall proficiency represented by the native-likeness measure demonstrate no apparent predictive relationship between WM and overall average proficiency by amount of experience. Some predictive validity of WM for overall proficiency was observed, however, in terms of individual performance within the NF and FN contexts: The learners who showed the highest level of WM in almost all experience categories in these two conditions also demonstrated the highest scores in WM. This was not the case for naturalistic learners, and the data for these learners do not indicate any apparent relationship between the ability and overall proficiency within this condition. The lack of a relationship between WM and the NL score in any of the three conditions was confirmed in the regression test, which showed no significance between WM and the NL measure.

These results for overall proficiency as measured by the NL score would appear to contradict the findings in a number of studies that indicate a significant correlation.
between WM and overall proficiency, especially in terms of fluency and accuracy (e.g. Daneman, 1991; Kormos and Safar, 2008; Mota, 2003). However, as an average of the main component scores of fluency, accuracy and the proficiency interview, the NL measure is perhaps a far too general proficiency measure for such an analysis. Most WM research has focused on measures of specific aspects of proficiency such as fluency, complexity, or accuracy. A composite score such as the NL measure is perhaps ill-suited for these research purposes, though it certainly represents an adequate measure for standard proficiency ratings and related research.

**Working memory and fluency by type of learning experience**

The analysis of predictive relationships between working memory and fluency in the three learning contexts also appears to demonstrate so significant relationship between the ability and learner fluency across the learning conditions. The results show that average fluency for low WM learners was almost native-like in some experience categories, whereas average or high WM averages co-occur in others with average fluency scores that are considerably lower. Results for individual learners also showed no predictive patterns between WM and individual performance for different levels of experience. Based on the results, therefore, there does not appear to be any consistent predictive relationship in the data between WM scores and measures of fluency. This lack of correspondence was confirmed with the regression analysis, and no apparent significance was found for WM as a predictor of fluency under any of the three learning conditions.

As with the overall measure of proficiency represented by the NL score, the results for fluency would appear to contradict a number of WM-related studies, including
research by Daneman (1991), who found a significant correlation between WM and verbal fluency. Fortkamp (1999) found that learners who had higher levels of L2 WMC also exhibited higher levels of fluency in various speech tasks. Harrington and Sawyer (1992) found that those learners who possess higher reading spans in their L2 also show higher levels of performance on the TOEFL grammar and vocabulary subtests. However, studies such as Unsworth and Engle (2007) indicate that a reliance on working memory is only implicated in contexts or tasks in which competition arises between processes involving higher or lower degrees of automatization. Other studies also point out that the importance of learning context, rather than abilities such as WM, as the major factor, with some researchers falling on the side of immersion or so-called high input contexts (DeKeyser 2007a; Freed 1995b; Miller and Ginsberg 1995), and others who have found that fluency may even be more directly attributed to more formal learning contexts (Collentine and Freed 2004; Freed 2008). It is entirely possible that neither of these conclusions is correct, and that for fluency working memory is largely a context-oriented ability that interacts with some learning conditions, though perhaps not all. The results for learning conditions found in this study and others (see Collentine and Freed, 2004; Freed, 2008) appear to demonstrate that learners under more formal conditions may attain equivalent (or potentially higher) levels of fluency as those who learn more naturalistically.

**Working memory and accuracy by type of learning experience**

The results for learner accuracy in relation to WM show a more consistent pattern of working memory’s potential predictive relationship with this aspect of proficiency than with either overall proficiency or fluency, although the relationship appears to be stronger for the naturalistic condition than it is for either the FN or NF learning contexts.

186
Individual profiles of working memory often, though not consistently, show that those participants who tested highest in WM usually scored highest in accuracy. This was again largely true for the naturalistic condition, though not for the two combined conditions, with some individuals with lower levels of measured WM scoring highest in accuracy in certain experience-related groups. Average scores for levels of experience likewise demonstrate inconsistencies in any predictive relationship. The results produced by the regression experiment indicated a significant relationship between working memory and accuracy for the naturalistic condition (p. = .007), but not for either the FN or NF contexts.

The results for accuracy provide some potential insights into the relationships between working memory and the different learning conditions. Since formal types of language learning have been shown to directly benefit accuracy-related abilities, one might expect that working memory would be predictive of such abilities in such contexts. A number of researchers have reported results indicating predictive correlations between WM and accuracy along with other measures such as complexity and fluency (e.g. Mota, 2003; Fortkamp and Bergsleithner, 2007). Fortkamp and Bergsleithner (2007) suggest that working memory (as measured by a speaking span task) is implicated in the production of L2 speech, but not in the capacity to notice patterns in the input. The findings of the present study would appear to contradict previous findings related to accuracy.

However, these results are not without precedence. Findings reported by Mota and Weissheimer (2009) indicated considerable correlations between WM and both fluency and complexity, but not for accuracy. The significant influential element may
therefore be the type of learning context in which accuracy is developed. The significant relationship between WM and accuracy in the naturalistic condition and not for learners who developed proficiency from the combined conditions may indicate that learners rely upon cognitive resources such as WM differentially depending on the nature of the environment. This possibility appears to be borne out in the interaction experiment, which showed an interaction between WM and naturalistic learning in relation to accuracy (see Section 5.3).

5.2 Working memory regression experiment: Fossilization Rate

The results for both the overall analysis and the regression experiment to test relationships between working memory and the observed rate of potentially fossilized forms appear to show a strong correlation between WM and fossilization, though the strength of the relationship differs between the three conditions. Significance for the relationship was found for the naturalistic condition, while an interactive relationship between WM and formal training was found for the FN condition (see Section 5.3), but not the NF learning context. The data therefore demonstrates differential results for the relationship between the three learning conditions.

The observed rate of fossilization demonstrated some of the strongest apparent relationships with working memory among all aspects of proficiency tested in this study. The data appear to demonstrate a strong predictive negative correlation between WM and this measure for all three learning conditions, although the relationship appears strongest for learners in the naturalistic condition. Results for individual learners showed that in all three conditions the tendency was for the participants with the highest levels of measured WM to demonstrate the lowest fossilization rate. The relationship was strongest for
naturalistic learners, and in each experience-related group those NC learners who exhibited the highest WMC also produced the lowest FR. This pattern was also true for most experience groups in the NF and FN conditions, though correlations appear stronger for FN learners than for NF context participants. Relationships were also found for experience-related variables, and when NF and FN participants were divided according to the amount of formal experience, a significant predictive correlation was found between this variable and fossilization for FN learners. These observations were confirmed by regression, with results indicating a significant correlation (p. = .002) between WM and FR for the naturalistic context. Significance was also found (p. = .024) for the amount of formal training as a predictor of fossilization in the FN condition (see discussion in Section 5.3).

In light of the significant predictive relationships seen between working memory and both accuracy and fossilization for the naturalistic condition, the absence of such significant correlations for accuracy-related measures in the two combined conditions confirms the prediction of stronger correlations between working memory and measures of proficiency under naturalistic conditions (Hypothesis #2). Correlations between the construct and accuracy fit in with some previous research that found significant relationships between accuracy and WM (e.g. Mota, 2003; Fortkamp and Bergsleithner, 2007). However, the absence of such a relationship for FN and NF contexts may indicate that results demonstrating predictive WM relationships with accuracy are not solely related to a one-to-one relationship between WM and this aspect of proficiency, which would help to explain the results of some studies that found no relationship between WM and accuracy (e.g. Mota and Weissheimer, 2009).
The mixed predictive results for WM in this study may indicate that such relationships are not so clear cut: the nature of the learning condition, i.e. in this case the characteristics of formal instruction, perhaps moderates the relationship. Such a conclusion appears warranted in light of the results for the moderated regression experiment (see Section 5.3). Although research related to potential correlations between abilities such as WM or learning contexts and fossilization is limited, the findings for FR in this study support the suggestion of a number of researchers that fossilization is potentially caused by a lack of sufficient amounts of formal types of language learning (e.g. Doughty, 2003; Schmidt, 1983; Seliger, 1975). Several studies point to a lack of key elements of formal instruction as probable causes, such as corrective feedback (e.g. Lightbrown and Spada, 1999).

Any evidence linking either the duration or timing of formal learning contexts with reduced rates of fossilization may indicate the potential of both improving our understanding of the issue and perhaps its prevention. The results reported in this study related to the occurrence of potential fossilization in respect to both working memory and the nature of learning conditions provides an impetus to move the discussion of this intriguing aspect of learner proficiency forward and examine both its causes and potential strategies to limit its effects.

5.3 WM-Learning condition moderated regression

The results for the moderated regression indicate a number of interactions between working memory and elements of proficiency. Interactions were found for both the naturalistic and the FN conditions, but not for any aspects of proficiency in the NF context. The analysis of the interaction between the naturalistic condition and working
memory yielded a significant interaction for fluency in this condition. A post-hoc test indicated that levels of significance were only found for low and high levels of WM, and not between average levels of the ability and either low or high levels, indicating that the relationship is not a gradient interaction. Since no predictive relationships were found between either WM or naturalistic experience and fluency, the results indicate that although neither variable independently has an effect on fluency, the two in combination do, though perhaps not completely: working memory apparently works together with naturalistic learning to affect fluency at least for low and high WM learners.

In light of an absence of any significant predictive relationship of WM for fluency for this condition (although near significance was observed at p. = .062), as well as previous research that has shown robust correlations between WM and fluency in other learning contexts (e.g. Daneman, 1992), this finding appears to indicate that WM does not operate in isolation in naturalistic learning, but that there is a potential moderating effect of the context itself that serves to ‘activate’ working memory resources with respect to this element of proficiency. It is therefore of interest that no interaction was found between WM and either the NF or FN contexts for fluency, while there was also an absence of any predictive correlation between the ability and fluency in either of these conditions. It is clear that additional research will be necessary to further clarify the interactive (and predictive) relationships between WM and fluency in naturalistic conditions. Although no interaction was found between WM and fluency for either of the combined conditions, a different result was found in relation to accuracy.

The other significant interaction found in the moderated regression experiment was in relation to the formal/naturalistic condition. For FN learners WM was found to
interact with the percentage of experience comprised of formal training in relation to learner accuracy. In the analysis, while WM alone did not show any predictive validity for accuracy among FN learners (p. = .537), its interaction with the percentage of experience comprised of formal training demonstrated significance (p. = .001). Interestingly, the percentage of experience consisting of formal training did show significance in predicting accuracy independently of WM as well (p. = .033). These results indicate a fairly significant interaction between WMC and the extent to which overall FN learner experience is comprised of formal classroom training in respect to accuracy, which is a finding that helps to explain the relationships observed between the number of potentially fossilized forms and formal experience described in Section 5.2.

The interaction between WM and formal instruction in terms of accuracy for FN learners appears to fit the findings related to formal training and its predictive validity for fossilization under this learning context (see section 5.2). Since fossilization is directly related to accuracy, and probably constitutes a byproduct of repeated errors in morphosyntax (see Hulstijn, 1989), these two seemingly separate findings could be interpreted to indicate an intricate relationship between learning conditions and major components of both proficiency and learner aptitude. The apparent relationship between formal training and WM may also help to explain the mixed results of past research related to potential correlations between WM and accuracy, with some research indicating significant correlations (e.g. Mota, 2003; Fortkamp and Bergsleithner, 2007), and other studies finding a lack of significant correlations between working memory and accuracy (e.g. Mota and Weissheimer, 2009).
Based on the findings for the interaction experiment, it becomes clear that to more accurately determine the nature and extent of such relationships, it is important to further isolate specific learning conditions and the duration of experiences and examine the relationships both predictively and interactively. Additional interactive research related to other types of learner aptitude, such as inductive language ability, will also greatly broaden our understanding of the dynamic relationships between cognitive resources and learning context in the acquisition of the major aspects of proficiency, and will aid future research related to the twin problems of inter-learner and intra-learner variability. It is also clear that while a number of interesting findings have resulted from this experiment, a clearer picture of interactive relationships could be made by incorporating into the study greater breadth in terms of the number of participants and a further narrowing of the amount of experience for study participants.

The absence of any statistically significant relationships between WM and any component of proficiency in the NF condition is an additional intriguing finding of this study. In light of the results seen for this condition in comparison with the data for FN learners, which demonstrated significance in relationships with accuracy and fossilization in particular, it is curious that no such relationships were demonstrated for the NF condition, even when a number of participants reported amounts of formal training roughly equivalent to the amounts reported by FN learners. It is possible that the difference is related to effects associated with the sequencing of the two conditions. If naturalistic or incidental types of learning tend to activate certain resources for particular aspects of proficiency, as suggested for accuracy and its apparent effect in the rate of fossilization, it is possible that learners may be preconditioned to rely on certain cognitive
abilities in a given context, and that when the transfer is made to a different learning context a type of mismatch occurs that somehow prevents or interferes with the use of the ability in further learning. The evidence that an interaction does occur between conditions and working memory for some conditions or sequences of conditions and not others both appears to confirm the existence of differential effects in WM interactions among different linguistic subsystems of the IL grammar (Hypothesis #1) and motivates a more in-depth analysis of these relationships.

5.4 Learning condition-Russian SLP correlations

With participants divided according to the type and amount of learning experience, the results for average and individual proficiency component scores and rate of potential fossilization in this study indicate a number of intriguing relationships between learning conditions and the nature of the IL grammar. The patterns or relationships demonstrated by the results include: 1) Individual test score results that appear to demonstrate an inverse relationship between WM and the extent of formal types of language training; 2) differences among learning conditions in the effect of an observed barrier to more advanced proficiency for specific aspects of SLP; 3) differences in levels of SLP attainment between learning contexts for given amounts of experience; 4) an apparent relationship between the amount and timing of formal training and both the attainment of accuracy and the degree of observed fossilization for the two combined learning conditions; 5) differential context-related predictive relationships between working memory and performance for some aspects of proficiency and not for others, confirmed by regression analysis.
Individual performance for fluency, accuracy and rate of fossilization

The data related to individual experience and SLP show what appears to be an inverse relationship between working memory and the type of learning experience. First of all, the individual who scored highest in all SLP measures (the most native-like participant) was found to possess a rather low level working memory span (ranked 32 out of 37 participants). She (an FN condition learner) also reported an inordinately high amount of formal intensive training. Like a number of participants in this study (n = 5), this learner achieved a native-level rating in fluency. Of significant interest, however, is the fact that she was the only participant who had attained a native-level of accuracy, and also demonstrated the lowest observed incidence of potentially fossilized forms.

Similar to the FN learner described above, the next highest proficiency participant also possessed native-level fluency and scored near-native on the accuracy subtest. The key differences between the two involved the second learner’s experience being almost entirely naturalistic (she also reported a modest amount of tutoring in-country without grammar explanation), while this learner evinced a far higher incidence of fossilization (2.6) and demonstrated the second highest level of WMC of all participants. It is also pertinent that the highest level of WM of all study participants was exhibited by the second most proficient naturalistic learner (native-level fluency and over 80% of native level in accuracy). The other related observations among individual learners included: 1) the individual who demonstrated by far the lowest accuracy rating (1.0) and the highest FR (10.3) also exhibited the lowest memory span of all participants, 2) the high number of FN condition learners (n = 7) among the 10 highest proficiency participants, and 3) though their SLP performance was lower than FN condition subjects, NF learners also
demonstrated an advantage over NC learners in both accuracy and FR commensurate with the extent of formal training experiences.

These observations for individual learners appear to demonstrate three important findings. The first relates to the apparent advantage individual learners can derive from formal learning experiences, especially in the area of accuracy and its association with the occurrence of repeated errors or potentially fossilized forms. That higher amounts of formal instruction as a percentage of overall L2 experience were associated with higher accuracy scores and lower fossilization rates even when working memory resources were limited appears to demonstrate a key role of formal training in proficiency attainment. This apparent advantage is also underscored by the large number of FN learners among the 10 most proficient participants. Secondly, although those NF learners who had at least some formal training faired better in accuracy and rate of fossilization than naturalistic learners, the NF and FN participant data also seems to argue for the importance of initial formal instruction prior to immersion experiences. The distribution of participants among the various experience categories are similar for these two conditions, and a comparison of learners with similar amounts of experience and formal training reveals that NF learner speech evinces both higher rates of fossilization and overall lower accuracy scores than the speech of their FN counterparts. Such patterns seem to indicate that more native-like accuracy is best developed by means of initial formal training. Perhaps most importantly, the extent of formal training also appears to correspond with a sharply reduced rate of potentially fossilized forms, the existence of which potentially poses the greatest obstacle to native-like proficiency in L2 Russian. Lastly, the apparent inverse relationship between WM and accuracy seen among these
learners demonstrates the potential value of providing L2 learners with a working
memory profile (along with other key cognitive measures such as inductive analytic
ability) prior to embarking on any language-learning strategy.

The correspondence with accuracy and lower fossilization rates among individual
learners in this study fits in with previous research findings. A number of studies have
demonstrated the advantages of formal instruction in terms of accuracy for the individual
learner (e.g. Ammar and Lightbrown, 2005; Doughty, 1991; Eckman, Bell and Neslon,
1988; Harley, 1989), while some researchers have linked fossilization to a lack of formal
language learning experience (e.g Doughty, 2003; Schmidt, 1983; Seliger, 1975), or a
number of potential causes including variables that are essential aspects of instructed
language learning (e.g. Doughty, 2003; Lightbrown and Spada, 1999; Tomasello and
Herron, 1988). These dual issues of accuracy and fossilization will be revisited in
relation to the discussion of average scores for the tested aspects of SLP for the three
learning contexts.

**Barriers to advanced proficiency attainment: Inter-condition variability**

Perhaps the most readily observed experience-related characteristic in the data is a
reflection of the general barrier to continued progress that was observed for all study
participants. The plateau effect appears to have an impact on all major areas of
proficiency for naturalistic and NF learners, but is especially acute in terms of fluency
and accuracy, with inter-condition differences in the rigidity and evident strength of the
barrier and the point at which it occurs for the two proficiency measures. The barrier to
advanced proficiency appears to have a far lower effect, however, on accuracy-related
measures for the FN condition, while the effect is apparently less for NF learners than it
is for those who have learned Russian naturalistically. A review of the main aspects of the results will aid in describing this limitation to more advanced proficiency.

The occurrence of a plateau in proficiency is especially apparent for fluency data. For naturalistic participants, average fluency appears to peak between 14 and 19 years, when the average is actually native-like, and then drops to about 67% of native-level fluency for participants with over 19 years of experience. The peak appears progressively earlier, however, for both NF and FN learners. The highest average fluency of 85% of native level was found for NF participants at between 9 and 14 years, and then dropped to an average of 70% for learners with more than 14 years of experience. For FN condition learners, however, the peak average of 82% occurred at between four and nine years of experience, and then dropped to an average of 72% following the peak. Surprisingly, the drop in fluency appears to be more dramatic with both NF learners and NC learners, than it is for the FN condition, with a total variance of participant score averages of 2.3 for NC learners, 1.5 for NF learners, and 1.0 for the formal/naturalistic condition.

A similar plateau was observed for accuracy within the three conditions, although greater variability in its apparent strength and the point at which it occurs was found between the different conditions. Average accuracy for naturalistic and FN learners appears to peak at a point between four and nine years, while for NF learners it occurs between nine and fourteen years of experience. For NC condition learners the high average of 60% of native-level drops to an average of 45% for those with over nine years of experience and then appears to stay relatively stable with increasing amounts of experience. For NF learners the peak average score of 70% of native-level accuracy
occurs relatively late in terms of experience at between 14 and 19 years. As with the naturalistic group, the peak for FN learners (approximately 71% native-level) occurred at between four and nine years. Average accuracy decreases for this group to a low of 60% for learners with over 19 years of experience. Inter-learner variability in average accuracy was lowest within the FN condition with a 1.1 difference between the lowest average score and the highest, while the division was 3.2 for FN learners and 3.0 for the naturalistic condition. These differences can be interpreted to indicate that the strength of the observed barrier to advanced proficiency differs between the three conditions in this data, with accuracy seen among formal/naturalistic learners demonstrating a significantly lower decrease over time.

As a combined average of scores for fluency, accuracy and the general proficiency interview score, the overall native-likeness (NL) score shows a less abrupt decrease for NC learners, while there does not appear to be any significant decrease in overall average proficiency for either NF or the FN condition learners. While in the naturalistic condition average NL scores peak at an average of 66% of native level relatively late (between 14 and 19 years), followed by a 16% drop after this point (the late peak appears to reflect the high average fluency during approximately the same period), NF learners appear to have hit a ceiling to further improvement for those with between 14 and 19 years experience, but without a decrease in the score (again apparently reflective of high average scores for both accuracy and fluency at later stages of development). FN learners likewise appear to hit a ceiling (at about 68%) without a significant decrease in the average, which levels out at about 64% of native level. This apparent barrier to advanced proficiency does not
appear to be a factor, however, for the observed potential fossilization in the data for the three conditions.

The observed potential fossilization rate does not appear to demonstrate either a plateau effect or any strong directional change among the three conditions. The average fossilization rate appears to decline for the two combined conditions, whereas FR increases in an apparently gradient fashion for the naturalistic condition from a low average of 2.2 for learners with between four and nine years, to its high of 4.4 for those with over 19 years of experience. The reverse is true for both the NF and FN conditions, with NF learners showing a considerable decline from an average of 1.8 for the least experienced learners to 1.1 for learners with between nine and fourteen years of experience. The lowest average rates for every experience category are demonstrated by FN learners, whose average FR also declines progressively from .92 to .80. Overall averages for all participants in each condition demonstrate an apparent advantage for FN learners, who show an average potential fossilization considerably lower than learners in the other two conditions.

The findings for individual learning conditions appear to confirm the observation made for all participants that a plateau occurs on average at a point after approximately 12 years of experience. This observation includes three major findings: First, the plateau exists for both fluency and accuracy in two of the three conditions, including the NC and FN contexts; the absence of data for NF learners with over 19 years of experience rules out any similar conclusion related to accuracy for this learning context. The second major finding is that one exception to the plateau phenomenon appears to be the rate of potential fossilization, which increases progressively for NC learners, while it either
stabilizes or decreases for NF and FN learners. The third observation is that the effect of the barrier is apparently variable, depending on the makeup of learner experience; a distinct advantage seems to exist for those learners whose experience has included formal classroom training, but with differences between the two combined conditions.

Both FN and NF learners demonstrate an apparent decrease in the impact of any barrier to improved accuracy and rate of fossilization when compared to naturalistic learners. Moreover, it appears to be the sequence of the two types of combined experience that is perhaps the overriding factor in sustained accuracy-related abilities: FN learners in this study demonstrate higher average accuracy, a considerably lower average rate of potential fossilization and what appears to be an overall more consistent trajectory of development with greater stability and a far less significant fluctuation in accuracy over time than NF learners. These learners also reveal an apparently consistent decrease in observed fossilization rate with amount of experience. These findings further substantiate the observation stated earlier regarding individual learners: the data indicates an advantage in the long-term development of native-like morphosyntactic accuracy for FN condition participants, with a decreased effect of any barrier to more advanced levels of proficiency. Additional evidence for the correspondence between this aspect of the IL grammar and formal training is provided by the results related to the extent and timing of formal learning experiences.

**Russian SLP by amount and type of experience**

The data show considerable variability among average component scores between the three learning contexts, with the observation of distinct differences between the contexts for average fluency, the composite native-likeness score, accuracy, and the rate
of potential fossilization for given amounts and types of experience. A second observation is an apparent disparity between learning contexts in terms of average accuracy and rate of fossilization for given levels of overall experience, with test score averages showing differences in the extent of inter-learner variability for these aspects of proficiency between the three conditions. Additionally, it is also apparent that the amount of experience is not as significant a variable in participant proficiency as the type of experience. A final observation concerning the impact of type and amount of experience on SLP in Russian involves what appears to be a correspondence between the amount of formal experience as a percentage of overall experience and participant performance in accuracy-related aspects of proficiency.

**The general ability or native-likeness score by experience**

The overall proficiency or general ability represented by the native-likeness score for the three conditions generally reflects the results found for the main components of proficiency, with a possible plateau beyond which average overall proficiency does not appear to advance for FN and NC learners, while the data for the NF condition are insufficient to make a determination. An interesting aspect of the cross-condition NL averages by amount of experience is that average NL scores peak earliest for FN learners at the four-nine-year mark and latest for NC learners (14-19 years). Another characteristic of NL averages is higher NL scores for FN learners in each experience category, while the highest average is attained by NC learners with between 14 and 19 years of experience. NL score variances appear to favor learners who have had formal experience: A variance of 4.1 was found for naturalistic learners, 1.9 for NF participants, and 1.2 for the FN group. Overall averages again indicate potential advantages for the
FN context, which indicated an overall average native-likeness that was 1.4 points higher than naturalistic learners, and .7 points higher than the NF condition average.

The NL-related data appear to demonstrate an overall higher attainment of proficiency among FN learners, with average proficiency for this context measured as approximately 64% of native-level proficiency. In comparison, the average for NF learners was 57% and that for naturalistic group approximately 50% of native-level proficiency. The complicating factor, however, is the apparent existence of a barrier to more advanced proficiency beyond 12 years of experience and the existence of a higher proportion of NC learners (n = 5) with more than fourteen years of overall experience than either the NF or FN conditions, each of which were limited to only one participant with more than 14 years of experience. The fact that the NC condition included a fair number of participants within this experience category and the other two conditions did not is a limitation in the data and precludes any overall conclusions related to differences in ultimate attainment of proficiency between the three conditions. The overall trends observed in the data for nativelikeness or general proficiency are shown in Figure 5.1.

Figure 5.1 Average nativelikeness by amount of experience
Figure 5.1 shows average general proficiency or ‘nativelikeness’ for participants of the three learning contexts according to their amount of overall L2 Russian experience. The data appear to indicate that a higher level of general proficiency was attained by FN learners, although there is a rough parity between these learners and NF participants, especially after the four-nine year period. The initial lower levels of proficiency for NF learners appear to be potentially related to initial naturalistic learning and its effect on accuracy. The lowest average nativelikeness score was exhibited by naturalistic learners in those experience categories where data was available. Such lower averages appear to reflect the lower average accuracy scores achieved by these learners, for whom average fluency was high. The absence of data for the two combined learning conditions (FN and NF) for learners with over 14 years of experience (there was only one participant for each of the combined conditions in these categories) motivates additional research related to these conditions for learners with greater amounts of experience.

**Cross-condition comparisons of fluency by amount of experience**

Participant fluency across the three learning conditions targeted in this study did not demonstrate any apparent significant differences in attainment. The existence of a barrier to more advanced fluency attainment appears to play a role for the naturalistic context, while no firm conclusions can be drawn about the NF or FN conditions due to insufficient data beyond the 14-year point. The peak in fluency occurs between nine and fourteen years for the two combined conditions in the data. However, although additional research would be required to ascertain any advantages for a given learning condition, contrary to what would be expected, the results do seem to indicate at least a potential advantage for FN learners over the NC condition: Average fluency was found to
be considerably higher for each successive experience category in the FN condition. Furthermore, the overall average fluency was 11% higher for all FN learners than it was for naturalistic learners. Fluency for the FN condition also peaks earlier than it does for either NF or NC learners, while NC learner average fluency occurs latest during the 14-19 year period. Additionally, the differences between the three conditions in terms of total variance also appear to indicate a potentially more consistent development for FN learners: The highest variance for fluency was found for NC learners (6.58), and the lowest for FN learners (1.8), with total variance for NF learners in between the two (3.27). Averages across the three learning conditions for given experience categories are depicted in Figure 5.2.

**Figure 5.2: Average fluency by amount of experience**

Figure 5.2 indicates a differential peak in average fluency for learners across conditions. The peak appears to occur earliest for FN learners (at between four and nine years of experience), later for NF learners (at the 9-14 year point), and later still for NC learners, for whom the peak in average fluency occurs at between 14 and 19 years of experience. This peak is followed by a decrease in average fluency for all learning
conditions. The drop in fluency does not appear to be as significant for FN learners than for the other two conditions. Though the data for the three conditions is too limited for any definitive conclusion, it appears that average fluency may be more affected by long-term naturalistic exposure, which would explain both the higher levels for NC and NF learners at later points: there is a later peak for learners for whom naturalistic conditions either predominate or occur early on (for the NF condition).

It might seem reasonable to interpret the results for fluency as a long-term advantage for learners who receive formal training, especially for the FN condition. However, as mentioned in the discussion of native-likeness or general proficiency, the analysis is complicated by the existence of a higher proportion of NC learners with more than fourteen years of overall experience than either their FN or NF counterparts, for which the data were severely limited in this regard. In light of the observed barrier to advanced fluency generally beyond the 12-year mark, it very well may be the impact of this phenomenon that has produced the predominantly lower scores for naturalistic condition learners with over 12 years of L2 experience. To draw conclusions about any condition-related advantages for fluency, the limitation would therefore need to be overcome by including a greater number of participants for the NF and FN conditions with over 14 years of L2 Russian experience.

One of the most interesting aspects of the cross-condition results for fluency is the progressively later occurrence of the average maximum attainment among the three conditions. Average maximum attainment occurs at between four and nine years for FN learners (82% of native level), between nine and fourteen years for NF learners (85% of native level, and between 14 and 19 years for naturalistic learners (90% of native level
fluency). Such differences in the timing and the overall maximum attainment between the three conditions may be related to differences between the learning conditions. A higher average maximum for naturalistic learners would appear to make sense given the nature of more incidental types of acquisition, which are largely related to the learning of semantic and pragmatic meaning (Reber and Allen, 2000).

What may be difficult to explain is the earlier high peak for FN learners than for the other contexts. Additional research designed to determine if such a pattern is a consistent tendency for formal instruction is warranted. A number of studies have described the effects on fluency of formal instruction, and results commonly indicate a slower development of fluency for formal learners (e.g. DeKeyser, 1998). Some researchers have stated that the ideal context for the effective acquisition of lexical and prosodic information is when L2 acquirers are involved in negotiating the two types of meaning in the course of genuine speech acts (e.g. Long, 1996; Prabhu, 1987) (see also discussion by Ellis, 2005). One possibility might be an overall effect of formal learning on the development of fluency, resulting in an earlier attainment maximum followed by a consequent decline.

**Cross-condition accuracy by amount of experience**

One of the interesting aspects of average accuracy across the different learning contexts is the later average ‘peak’ in accuracy for NF learners than for either NC or FN conditions, both of which show the highest average accuracy occurring between four and nine years. Accuracy for NF learners in this data was also lower for learners with between four and nine years than it was for those with two to four years of experience. This disparity may be due to the differences between these experience groups in the
timing of initial formal training: the more experienced group had spent a greater amount of time learning naturalistically prior to formal training than those with from two to four years of experience. The data for the NF condition also shows that average accuracy for learners with four to nine years of experience was considerably lower than it was for NC learners with the same amount of experience. Such differences are difficult to explain by relying on amount of experience alone.

The data show different degrees of experience-related variability for accuracy within the three contexts, and the FN condition appears to demonstrate more consistent and potentially stable tendencies for accuracy for the different experience groups, with averages for the first three experience categories (between two and fourteen years) within .6 of each other. Average accuracy does not appear to decrease significantly for the FN context, and such learners appear to make strong initial gains in the measure which do not erode significantly. The distribution of accuracy scores was wider for both the NC and NF conditions (over 3.0), while the difference between the highest and lowest scores was 1.1 for FN learners. The overall variance was 1.87 for FN learners, 2.62 for the NF condition, and 2.32 for naturalistic learners. Overall average accuracy was lowest for NC learners at about 46% of native-level accuracy. NF learners demonstrated 49% of native-level accuracy. Learners in the FN category showed an average accuracy of 67% of native level. The data reveals a distinct advantage in accuracy for those learners who have included formal learning experiences as a major part of their overall acquisition strategy. Average accuracy across the three learning contexts for given amounts of experience are shown in Figure 5.3.
Figure 5.3: Average accuracy by amount of experience

Figure 5.3 demonstrates distinct differences between the three learning conditions in terms of average accuracy. A considerably increased ability in accuracy is apparent for those learners whose L2 Russian experience has included formal instruction, with demonstrably higher overall accuracy rates for both the NF and FN conditions when compared with naturalistic learners. In addition to a lower average accuracy for NC context learners, initial development in morphosyntactic accuracy at around the 4-9 year mark drops off considerably with increasing amounts of experience. This drop off appears to indicate a considerable negative effect on accuracy of continued naturalistic learning without formal instruction. The two combined conditions indicate what appears to be a detrimental effect of initial naturalistic learning for the NF condition, with a drop in accuracy occurring around the 4-9 year mark. However, this trend appears to be reversed in the data following an average of nine years of experience.

Several observations and potential conclusions can be drawn from the context-related findings obtained for participant accuracy in this study. The first observation relates to the relatively later increase in accuracy for learners in the NF condition. This
trend is perhaps best explained by the fact that these learners had received formal training later, after initial naturalistic learning. The early average score of 3.8 for NF learners at the four-to-nine year point may be a reflection of the effects of naturalistic learning, with the later increase reflective of later formal training. Many of the NF learners reported beginning their formal training relatively late, often after serving for four to eight years in Russia or Ukraine, at which time many had returned to the US to improve their perceived deficiencies in grammatical understanding and ability: the average point at which NF participants began formal instructional training was 4.6 years. It would appear that average accuracy-related scores for this learning condition following nine years of experience demonstrate a degree of success in this endeavor, especially when they are compared with scores achieved by strictly naturalistic learners with the same amount of experience.

The data in this study also indicate a lower variance for accuracy scores for the FN condition than for the other contexts, and although additional data will clarify these results, an apparent curtailment of any absolute decline in accuracy was indicated for learners who followed naturalistic experience with formal training. In terms of the potential effect of the sequencing of naturalistic and formal learning experiences, the higher accuracy scores demonstrated earlier on for the FN group, the apparently lower degree of decline in accuracy over time, and the overall higher accuracy average for FN learners (6.7) compared with either NC (4.6) or NF learners (4.9) seem to point to a potentially significant advantage in accuracy-related abilities for those who begin learning L2 Russian with formal training. These conclusions are further substantiated by the individual score results described earlier, which show that seven of the ten highest
individual scores in accuracy were achieved by FN learners. Although an advantage does appear to exist for learners who begin studying Russian formally, the formal condition experiment did not provide a mechanism to test for any statistical significance of the effect of sequencing. Additional research will be required to examine the exact relationships between the sequencing of conditions and learner accuracy.

As seen in the results for fluency and nativelikeness, more data for participants with over 14 years of experience will be required to definitively determine long-term effects of experience on the further development of accuracy for the NF and FN conditions. However, due to the nature of fossilization, which has been described as the automatization of inaccurate patterns in morphosyntax (Hulstijn, 1989), it is reasonable to expect little change in accuracy for learners following the 14-year point, since after the passage of this amount of time immersed in the L2, the automatization of such inaccurate grammatical usage would very likely have already occurred. Naturalistic learners in this data actually exhibited a slight increase in average accuracy (.5 percent) among learners with between 14 and 19 years of L2 Russian experience.

**Context-related results for observed fossilization rate by amount of experience**

The relationships between formal experience and accuracy noted above are also reflected in the data for rate of potential fossilization. The fossilization rate observed among the different learning conditions was the only measure that did not demonstrate any evidence of a ‘plateau effect’. Depending on the condition, the occurrence of such potentially fossilized forms in relation to an increase in the amount of overall experience appears to be either increasing, decreasing, or has potentially stabilized. In the data for naturalistic learners average fossilization rate increases for learners with increasing
amounts of experience, and for the first two experience groups the rate increases by .6, whereas FR appears to decrease considerably for the same experience categories for both NF and FN learners. The average rate is also nearly as high for NF learners as it is for NC learners initially, but then demonstrates a decrease by .6 points, or 35%. Average performance in the NC condition evinces a 27% increase for the same experience groups. In contrast, average FR for the FN context demonstrates an 11% decrease across the same experience groups. The mean variances for fossilization rate also indicate sharp differences across the three contexts with a high variance demonstrated by the naturalistic data (6.0), and relatively lower variances for both the NF (.5) and FN (.2) conditions. Overall average FR was nearly two times lower for the FN context than it was for NF learners, with a similar difference between NF and NC condition learners. The observed trends in FR among the learning contexts are shown in Figure 5.4.

![Figure 5.4 Average FR by amount of experience](image.png)

**Figure 5.4 Average FR by amount of experience**

Figure 5.4 demonstrates distinct differences between learning conditions in terms of the observed rate of potential fossilization. Three apparent trends in the data are of significance: 1. Naturalistic condition learners demonstrated far higher rates of
fossilization for each experience category; 2. While the two combined conditions demonstrated an overall decrease in FR with increasing amounts of experience, fossilization for naturalistic learners has continued to increase over time; 3. Among the two combined learning conditions, FN learners demonstrated low rates of fossilization (below 1.0) from the minimum experience category, whereas NF learners showed rates similar to naturalistic learners with the same amount of experience. It also appears that this high initial rate may be potentially limited in its effect when initial naturalistic learning is followed by formal learning conditions, as is the case for NF learners.

These differences in observed FR across the three conditions appear to provide evidence indicating that learners who include formal training as part of their overall strategy for learning L2 Russian increase their potential to avert any barrier to native-like proficiency that may exist in terms of accuracy and potential fossilization. Although more data is needed for the NF and FN conditions to better determine the effects of formal learning conditions on preventing fossilization in the long term, the lower rates for those learners who have included such training as part of their strategy support this conclusion. Additionally, it appears that native-like levels in terms of both fossilization and accuracy are indeed possible when learning is approached with initial formal training followed by more naturalistic experiences. While both FN and NF contexts seem to afford learners with the tools to more closely approach native-like consistency in terms of morphosyntactic accuracy, FN learners appear better suited to overcome accuracy-related limitations and avoid fossilization than learners in the NF condition.

The above conclusion is further supported by two additional bits of evidence found in this study. The first consists of the individual data related earlier in this chapter: six of
the ten learners with the lowest rate of potential fossilization and highest levels of accuracy were learners who followed formal training experiences with naturalistic learning. The second type of evidence is provided by an analysis related to the impact on tested aspects of proficiency of the amount or extent of formal training as a percentage of overall learning experiences for the two combined learning contexts. The isolation of formal experience in this way provides an even clearer account of its potential importance, especially in the areas of accuracy and the occurrence of potentially fossilized forms.

**Combined conditions: the impact of amount of formal training on accuracy and FR**

One of the key observations in this study is the apparent existence of distinct differences in attainment between the two combined (NF and FN) conditions in terms of the effect and sequencing of formal training in relation to naturalistic experiences. The isolation of formal training as an independent variable for these two contexts produced results that indicate several interesting relationships between the amount or extent of such training and the attainment of major aspects of proficiency. The first observation in the data is what appears to be no considerable correspondence between the amount of formal experience and fluency for either condition. For the FN context average fluency increases with an increasing percentage of formal experience, but then drops off by half a point for learners with more than 14% of their experience consisting of formal training. Fluency actually decreases with increasing amounts of formal training for the NF context. Similarly, the results for the native-likeness score show distinct differences between the two conditions: for NF learners there does not appear to be any consistent relationship,
while the FN context seems to demonstrate a general increase in average NL scores with greater percentages of formal experience.

The results for accuracy, however, appear to indicate a more positive correlation between this measure and the extent of formal training for both conditions: average accuracy increases up to the 14-22 percentage group for NF learners and increases as well for FN participants through the second level of formal experience, though it drops slightly in both conditions with learners who had the highest percentage of experience comprised of formal training. It is important to point out, however, that those participants with the amount of formal training at over 18 percent of the total also had lower amounts of overall experience, with average amount of overall experience for FN learners in this category about seven years, while the average for NF learners was 7.8 years. In addition to the above findings, an apparent relationship between the percentage of formal experience and proficiency in terms of the rate of fossilization was also found.

The percentage of total experience involving formal instruction seems to indicate a negative correlation with the rate of fossilization: as the overall percentage of formal experience increases, the rate of fossilization decreases. For NF learners with over four years of experience average FR decreases by 18% with an increase in the percentage of formal experience up to more than 22% of total experience. The decrease does not appear to be consistent for this condition, however, and FR actually increases for learners with between eight and fourteen percent formal experience. A more consistent and considerable decrease in fossilization holds for FN learners: average FR decreases by 46% between learners with from two to eight percent formal training and those with between eight and fourteen percent. The rate decreases an additional 16% for learners
with between 14 and 22% formal experience. The patterns for both accuracy and FR appear to demonstrate a stronger relationship between the amount of formal instruction and accuracy-related measures for FN learners.

Some of the inconsistencies found between accuracy-related aspects of proficiency and the amount of formal instruction for the NF condition may be related to the effects of sequencing. For example, the increase in the rate of fossilization that occurs prior to a later decrease with an increase in formal instruction may relate to the observation made by some L2 or FL Russian language teachers that learners who follow naturalistic learning with formal training often experience an increase in accuracy-related errors after initial formal training, which is then followed by a consequent decrease in error rates, whereas FN learners do not appear to exhibit such difficulties (A. Gavrilyuk, D. Wheeler, personal communication, October 4, 2012).

**FN and NF condition accuracy and FR by amount of formal training**

The additional categorization of NF and FN data according to the total amount of formal experience (rather than as a percentage) produced some additional results that are pertinent to this discussion. The data indicated that for the NF condition participants with the greatest overall amount of formal experience (approximately 1.6 years) and the highest percentage of formal training, achieved the highest average in accuracy and the lowest average rate of fossilization, while those NF learners with a lower average of overall formal experience (approximately .64 years) demonstrated a much lower average accuracy and a considerably higher rate of fossilization. In contrast, FN learners with both a high percentage of formal experience and a higher average of total formal training (1.9 years) demonstrated an average accuracy of 6.7 and had an average FR of .94.
Interestingly, those FN learners with both a lower average amount of total formal training (.78) as well as a lower percentage of overall experience consisting of formal experience also showed an average accuracy of 6.7 and an average FR of .93 FR.

It appears that there is a greater correlation between increasing amounts of formal training and both higher accuracy and lower rates of fossilization for the NF condition than there is for the FN condition. FN learners appear to show a greater consistency in both measures with varying amounts of formal experience. The results for an apparent relationship between formal training and accuracy in particular were confirmed by means of logistical regression, which found a significant predictive relationship (p. = .033) between the amount of formal training as a percentage of overall experience and accuracy for the FN condition, though not for the NF context. In relation to fossilization, the moderated regression test for interactions also indicated that while they do not affect the rate independently, the percentage of experience consisting of formal training and working memory work together to affect fossilization rate for FN condition learners (see section 5.3).

The amount of formal training as a proportion of experience also demonstrates what appears to be a negative correlation with the rate of observed fossilization for both of the combined conditions: as the overall percentage of experience consisting of formal training increases, the rate of fossilization decreases. This decrease is more considerable for the FN condition than it is for NF learners. Last of all, the results appear to demonstrate differences in the effect of formal experience on accuracy and FR between the two conditions related to learning condition sequencing. The results indicating significant correlations between formal learning experiences and the rate of observed
fossilization can be interpreted as at least partial confirmation of this study’s third hypothesis: a significant negative correlation was found between formal learning and FR. An examination of the overall significance in the effect of learning condition sequencing upon observed fossilization will require further experimentation.

**Some observations related to the above findings: accuracy and fossilization**

The overall findings in this study related to differences in attainment across learning conditions by amount of overall experience, and results concerning the effect of the amount of formal training on accuracy-related aspects of proficiency, motivate several key conclusions about the impact of formal learning experience. One of the most important observations that can be made is that the findings for accuracy, the combined native-likeness score and rate of potential fossilization discussed above appear to provide support for the contention that formal types of learning are essential if learners are to be successful in acquiring native-like grammatical accuracy. In terms of accuracy, the results provide further support of previous research that indicates distinct advantages for formal study over more naturalistic learning or even implicit styles of language instruction, with the most important consequence appearing to be the potential curtailment of the processes that cause fossilization in morphosyntax. The findings of a number of studies have demonstrated the importance of formal types of learning for the long-term development of grammatical accuracy in a number of different FL or L2 language contexts, including L2 Russian.

Harley (1989) found that without formal instruction English-speaking French L2 learners consistently failed to acquire the French past preterite and imperfect tenses, even when exposed to massive amounts of implicit or incidental input with feedback provided
in the implicit condition. A number of ESL studies have demonstrated the importance of instruction for the effective acquisition of English morphosyntax (e.g. Ammar and Lightbrown, 2005; Doughty, 1991; Eckman, Bell and Neslon, 1988; Gass, 1992; Hamilton, 1994). Research has shown that in the acquisition of relative clause structures (the basis of a large part of the accuracy data elicited in this study), those who have experienced formal classroom training outperform learners who have not. Examples of such research include studies of L2 Italian (e.g. Croteau, 1995) and Japanese (e.g. Yabuki-Soh, 2007), among others. The accuracy results in the present study reflect results found by Dunn (2007) in a study of the impact of instruction on adult L2 acquisition of Russian relative clauses. She found that learners made significant gains in the mastery (and application to other contexts) of relative clause structures for which they had received training than with structures for which no training had been received. Dunn concluded that learners who are instructed in a particular area will make greater progress in that specific area (see also Croteau, 1995).

A number of other researchers have emphasized the importance of instruction in terms of accuracy. Ellis (2002, 2008) has concluded that there are distinct advantages to classroom instruction in the acquisition of the great number of grammatical forms necessary for effective communication. Schmidt (1994) strongly suggests that true second language learning does not occur without the kind of explicit focus on form that is provided in formal instructional contexts. A number of key aspects of language learning that are difficult to address naturalistically can be addressed in formal classroom conditions, including corrective feedback (Lyster, 2004), online types of planning (Yuan and Ellis, 2003), and the ability to encourage the noticing of certain forms that in
naturalistic conditions often go unnoticed (Ellis, 2008). Research by Loewen (2002) demonstrated that even short amounts of time devoted to a focus on grammatical forms can produce improved results in post-instruction tests, whether they are given immediately following or after passage of a considerable amount of time following instruction. Formal training has also been cited as an important ingredient in the prevention of fossilization.

Doughty (2003) states that differences in attainment that represent a failure to acquire native-like accuracy, and consequently higher degrees of fossilization, are largely attributable to a lack of instruction; which has been implicated in long-term failure to develop native levels of proficiency. A number of researchers have similarly cited a lack of adequate instruction as the cause of fossilization, including Schmidt (1983) and Seliger (1975). Other potential causes include elements that are essential components of instructed learning contexts and not naturalistic conditions, including corrective feedback (Lightbrown and Spada, 1999; Tomasello and Herron, 1988; Valette, 1991). Without adequate amounts of feedback, many learners inevitably repeat inaccurate forms in speech, which would further reinforce errors that could be potentially eliminated with such feedback.

According to Hulstijn (1989), fossilization is directly attributed to the automatization of erroneous grammatical knowledge. Such automatization of errors would be a natural characteristic of more naturalistic or incidental types of learning, in which there is little if any corrective feedback or time to process grammatical information analytically, while the repetition of the errors themselves would tend to be ‘forced’ since the immersion environment demands immediate production in order to survive. It is
reasonable to conclude that if an individual learner’s production is characterized by grammatical inaccuracies in key areas such repetition would naturally lead to an automatization of the forms with fossilization as the result. Based on previous aptitude-related research, it would also follow that such fossilization would occur even more rapidly and extensively in the case of learners who possess lower levels of key cognitive resources, such as inductive analytic ability or potentially working memory.

Unlike most previous research, the present study helps to move the discussion beyond questions concerning the importance of formal instruction in proficiency to a discussion of other potential variables, such as the timing of formal and naturalistic experiences. The results related to the combined condition experiment speak to the issue of the sequencing of formal and naturalistic learning, and indicate distinct differences in attainment between the two combined (NF and FN) conditions. While the value of formal instruction appears to have been confirmed in this study, the findings also indicate that learners fair far better in accuracy-related abilities when L2 Russian is initially approached with formal instruction. The advantages of this sequence are not only seen in the higher accuracy scores and lower average FR for given amounts of experience for FN learners, but also in the more consistent development of these elements of proficiency in the long term.

**Additional observations: fluency**

Results that demonstrate on the one hand a lack of any correspondence between the extent of formal experience and learner fluency, while the results for accuracy appear to evince a direct correspondence with formal learning, further substantiate previous research that shows that the traditional nature of formal learning, with its tendency to
focus on form, is more conducive to the development of accuracy than to the native-like rapidity or flow of speech. The traditional college classroom is rarely structured on the premise of providing ample opportunities for the student to speak, and it has been said that 80% of the speaking in such contexts originates with the instructor (Ellis, 2005). The common emphasis on explanations of grammatical rules in the student’s L1, listen and repeat exercises for the production of grammatically accurate sentence patterns, and certain types of corrective feedback, certainly benefit the development of accuracy, but do not contribute much to the development of fluency.

Some researchers have suggested that true fluency can only develop in the context of actual communication (e.g. Prabhu, 1987; Long, 1996). Since Krashen (1981) and the rise of implicit or communicative language teaching, a number of researchers have continued to argue that if fluency is to be developed effectively in the context of formal language programs, then they must incorporate more frequent opportunities for spontaneous communication activities. According to DeKeyser (1998), in order to develop L2 fluency the learning context has to afford learners with the opportunity to focus on pragmatic meaning, which requires communicative types of activity. Long’s interaction hypothesis (1983; 1996) maintains that the process of negotiating meaning is crucial to overall proficiency, and especially fluency. The emphasis on the acquisition of pragmatic meaning is based on the idea that meaning is the key means by which the learner develops his or her linguistic knowledge and thus full acquisition (see Johnson and Swain, 1997). This theoretical perspective has been the driving force behind the development of a number of language immersion programs throughout the world (Ellis, 2005).
Traditional formal training unfortunately does not often afford many opportunities for interaction in the target language, and some of the newer developments in FL pedagogy, such as task-based learning and the oral-situational approach, were developed in order to promote more communicative opportunities in the formal classroom (see discussion in Ellis, 2005). In light of the research, the findings in this study for both fluency and accuracy in terms of formal experience appear to support previous research related to the effects of formal instruction on these aspects of proficiency. This study also motivates an additional conclusion pertaining to learner fluency: in the long term formal training does not appear to have any adverse effect on its development, and both FN and NF condition learners demonstrated equivalent levels of fluency for the amount of overall experience as their naturalistic counterparts. Results also appear to indicate that formal training may in fact even work to enable the development of fluency earlier than when learning is approached without initial formal instruction.

Although the results related to the type and amount of learning experience demonstrate the importance of learning context, it is also clear that something else is at work besides learning conditions: the formal condition-related results appear to indicate the tendency of higher WM participants to exhibit greater skills in measures of accuracy, including lower rates of fossilization, while it seems that the impact of lower WM resources can be offset by increasing amounts of formal instruction for those in the combined condition groups. Such patterns indicate an apparently intricate relationship between learning conditions and aptitude in the form of working memory, which is a conclusion that fits well with the results that were found for the WM-learning condition interaction experiment: WM does not appear to operate in isolation, with learning
conditions apparently having a moderating effect in the learner’s utilization of WM resources in acquisition. These intriguing results provide support for the adequacy of Robinson’s (2002b) framework in describing the relationships between cognitive abilities in learner aptitude and learning environment in L2 proficiency.
CHAPTER VI: CONCLUSION

The results that have been found in this study motivate both general and specific conclusions related to the relationships between working memory capacity, the main types of learning conditions or strategies learners have utilized, and dependent variables related to the key aspects of proficiency in adult L2 Russian. Four general conclusions can be made. The first relates to the observation that the overall amount of experience is not nearly as important in learner proficiency as the type of experience. A second conclusion concerns evident limitations or a ‘plateau effect’ in relation to accuracy, with an apparent barrier to more advanced levels of fluency and overall proficiency (native-likeness) that exists for most learners, although formal training appears to limit the impact of any such barrier on accuracy-related abilities. A third general conclusion involves the comparatively different impact of the most common learning strategies in the overall attainment of specific sub-domains of the IL grammar, including accuracy and the observed rate of fossilization. Lastly, working memory appears to play a differential role for specific characteristics of proficiency depending on the nature of the learning condition examined: the role of working memory does not appear to be restricted to an isolated one-to-one relationship with particular aspects of proficiency, but is greatly dependent on the nature of the learning context as well.

This study demonstrates the importance of learning context in the attainment of some of the major characteristics of native-like proficiency. The aspects of the IL grammar that are most keenly affected by types of learning are a) morphological
accuracy, and b) the complexity of sentence structure or syntactic elements of native-like speech. The comparison of the effects of the type of learning experience on these subdomains of the IL grammar produced results indicating that formal instruction plays a key role in the proper development of accuracy/complexity. The results also appear to show that the amount of formal instruction becomes crucial only when initial language learning experiences are comprised of naturalistic exposure, whereas learners who first study Russian under formal instruction do not appear to show much variation in accuracy in respect to the amount of instruction received, though it is apparent that a minimum amount of formal experience is essential. The study demonstrated a significant correlation between the extent of formal training as a percentage of overall experience and fossilization rates for FN learners, while no significant correlations were found between accuracy-related proficiency measures, including fossilization, and either the amount of naturalistic experience or overall experience in general in the combined learning conditions. It is therefore apparent that the type of learning context is of greater importance than the amount of experience. The amelioration of limitations on accuracy attainment shown for learners who include healthy amounts of formal training in their language learning further points out the importance attached to the type of experience rather than the duration of exposure.

Related to the above conclusion is the observation of a ‘plateau effect’ in respect to accuracy, an overall increase in fossilization rate with greater experience, and an evident barrier to native-like levels of fluency and overall proficiency that was observed for most learners. The barrier to native-like fluency appears to affect learners who have acquired L2 Russian under any of the targeted learning conditions. Contrary to expectations, the
results also demonstrate no advantage for the attainment of fluency for naturalistic learning. The long-term proper development of native-like rapidity of speech without undue pausing or hesitation does not appear to be greatly affected by the nature of initial learning experience. The only potential advantage in the attainment of fluency appears to be correlated with a learner’s working memory capacity, although any correlation between the two is apparently only significant in terms of an interaction between WM and naturalistic experience, while for this context a near significance was found for a direct predictive relationship between the two variables. The effect of the plateau or barrier to advanced proficiency seen for accuracy-related measures is another matter, however.

For learners in the combined experience categories, the amount of formal instruction received appears to limit the impact of any such barrier on accuracy and fossilization, and the three types of learning experience examined in this study indicate differential effects on these two related measures. The results for accuracy and fossilization appear to demonstrate that the amount of formal instruction received has an impact on attained accuracy with consequences for the development of fossilized forms over time. Compared with participants in the combined learning conditions, learners who did not include formal instructional experience as part of their overall strategy demonstrated a barrier to native-like accuracy and a considerable increase in the incidence of potentially fossilized forms with similar amounts of experience. It is apparent that adult L2 Russian learners who include formal instruction as a major part of their learning are more likely to develop native-like accuracy and prevent fossilization of inaccurate characteristics of the IL grammar. A relationship between experience in
formal instructional contexts and accuracy-related proficiency was further underscored by the results for logistic regression which produced two significant findings: 1. the amount of formal learning predicts the rate of fossilization within the FN learning context, with a significant negative correlation seen between the two; and 2. in the FN context formal learning appears to interact with working memory resources in the learner’s attainment of accuracy. These results for the FN context underscore another finding related to the order or sequencing of naturalistic and formal types of instruction.

The results for accuracy and fossilization further indicate that the sequencing of formal instruction with naturalistic experiences also has a direct bearing on the attainment of accuracy and the occurrence of fossilized forms with equivalent amounts of formal experience. While large amounts of formal instruction appear to ameliorate or even potentially limit accuracy-related deficiencies in the IL grammar for those who follow initial naturalistic learning with intensive formal training, the data show a distinct advantage for those learners who begin their L2 Russian learning experience with formal types of language learning. Average accuracy was found to be higher and the average fossilization rate lower for FN learners than it was for NF learners with similar amounts of formal instruction. Furthermore, the amount of formal experience did not appear to be as important a factor for FN learners as long as they had a minimum of about five-six months (one month equals 80 contact hours) of such experience prior to complete immersion. Such results point out the apparent advantages in beginning L2 Russian acquisition in a formal context, with subsequent immersion experience.

A final general conclusion motivated by the results of this study is that working memory appears to play a differential role in the attainment of overall proficiency and its
specific elements, depending on the nature of the learning condition examined: the WM-related regression results indicate that the role of working memory is not limited to specific relationships with major characteristics of proficiency, such as accuracy, fossilization or fluency, but is greatly dependent on the type of learning context. The logistical regression experiment demonstrated a correlation between WMC and accuracy-related aspects of the IL grammar among naturalistic learners, with significance found for both accuracy and rate of fossilization, but with no such relationship found for either of the combined learning conditions. However, the moderated (interactive) regression experiment did find significance for a relationship between WM and accuracy for the FN context, while an interaction also occurs between WM and overall proficiency or native-likeness for these learners. Additionally, WM was also found to interact with naturalistic learning in the attainment of fluency for naturalistic learners.

When compared with the results that indicate a predictive relationship between formal experience and accuracy for FN learners, and the absence of significant correlations between WM and proficiency for NF condition learners, these results provide a strong indication that there is a differential effect of working memory on specific aspects of the IL grammar between various learning conditions. The role of WM also appears to be moderated by other independent variables such as the nature of different learning experiences. It is therefore likely that rather than functioning independently in the development of aspects of proficiency, working memory ‘interacts’ with various learning contexts differently for the attainment of different aspects of the IL grammar.
Specific conclusions: The three hypotheses revisited

This study was designed to address several research questions related to the relationships between working memory, common L2 Russian learning conditions and the attainment of specific elements of proficiency. The three hypotheses articulated in this study directly relate to three general questions: 1. To what extent does WM predict specific aspects of proficiency in L2 Russian, including accuracy, fluency, and potential fossilization for the three targeted learning conditions? 2. To what degree does WM interact with these learning contexts in producing attained levels of these three aspects of proficiency (accuracy, fluency and fossilization rate)? 3. What are the effects, if any, of the different learning conditions and their different sequences on the major aspects of proficiency and the incidence of potentially fossilized forms?

The first hypothesis tested relates to the first research question, and predicted that WM would independently demonstrate stronger positive correlations with attained levels of proficiency for strictly naturalistic learning than for conditions that rely heavily on formal instruction. This hypothesis appears to have been confirmed by results indicating that although working memory shows significant predictive validity for two related aspects of proficiency under naturalistic learning experiences (accuracy and fossilization rate), it does not act as a predictor of such proficiency characteristics in either of the two combined conditions. Results indicated no significant predictive relationship between WM and any of the tested aspects of proficiency for either the NF or the FN conditions. Conversely, for the FN condition the amount of formal instruction was shown to be predictive of fossilization rather than working memory.
The second hypothesis predicted differential effects in the interaction of working memory with the different learning conditions in relation to both overall proficiency (as measured by the native-likeness score) and specific characteristics of the IL grammar, such as fluency or accuracy-related measures. This hypothesis was confirmed in the analysis. Three significant interactions between working memory and learning conditions were found. The first interaction was between WM and the strictly naturalistic context for fluency, while no such interactions were found for either one of the combined conditions. Working memory also demonstrated two significant interactions with the amount of formal experience for the FN condition, but not for NF condition learners: in the FN context significant interactions were found between these two fixed variables for both accuracy and general proficiency or native-likeness. These results demonstrate an apparent differential impact in the relationships between WM and types of learning experience for both different aspects of the IL grammar, and overall level of proficiency.

The third and final hypothesis relates to the effect of formal instructional experience on the presence of potentially fossilized forms in the IL grammar. It was predicted that the amount of formal learning would demonstrate a significant negative correlation with the presence of potentially fossilized forms, and that the sequencing of naturalistic and formal learning contexts would have a significant effect on fossilization. The first part of this hypothesis was apparently confirmed based on several observations: the rate of fossilization for learners within the two combined conditions is considerably lower than that demonstrated by learners in the naturalistic context, and a negative correlation is seen between the amount of formal training and the occurrence of such forms. The second part of this hypothesis was neither confirmed nor disconfirmed. Although the average
rate of fossilization is considerably lower for FN condition learners both in terms of overall average FR and for equivalent amounts of formal experience, the significance of the correlation between the FN sequence and lower rates of fossilization will need to be confirmed via additional research.

**Study limitations and recommendations**

This study had a number of general limitations in relation to its design. The first and most apparent limitation relates to the difficulties associated with the need to more accurately isolate specific learning conditions and their duration with greater numbers of participants for each condition. The comparison of the three learning contexts and their effect on major aspects of proficiency for highly experienced L2 acquirers was limited due to the low number of participants with over 14 years of experience in the two combined learning conditions. A follow-up study with greater numbers of such learners for these contexts would act to further clarify the study’s results for these experience categories. A second limitation relates to the failure to provide data for formal instruction in a more isolated way, such as including learners whose L2 experience has been limited to formal instruction. Third, results would also be further clarified by more narrowly defining the amount of experience for each learning context: a method needs to be implemented that limits the participant pool to specific amounts of past learning under the different conditions. In order to confirm these results future research may therefore require a more narrowly defined research program that emphasizes a longitudinal design for given learning conditions. A fourth, more specific, limitation relates to the design of the combined condition experiment.
The second part of the third hypothesis predicted a significant effect for the sequencing of the two main conditions on accuracy-related aspects of the IL grammar. However, the experiment failed to provide a clear mechanism to test this aspect of the hypothesis. This drawback in experimental design could be eliminated by providing a more concrete statistical analysis by isolating the two condition sequences in the data and testing for predictive relationships with accuracy measures. For such a procedure to be effective a greater number of participants would also be necessary with a more narrowly defined subject pool in terms of the amount of experience for the NF and FN contexts.

**Final Conclusion and recommendations for future research**

In spite of its limitations, this study provides insights into two important aspects of second language proficiency. The first is that a learner’s working memory plays a vital role in the acquisition of the major elements of proficiency in terms of their development and the level of success learners achieve. The results of this study also appear to confirm the interactionist model proposed by Robinson (2002b), in that this important cognitive ID variable does not appear to function independently: the role working memory plays appears to be highly dependent on the nature and proportions of naturalistic and formal conditions of learning. The indication of significant interactions between working memory and different learning contexts in the level of proficiency acquired by learners of L2 Russian motivates additional research designed to discover both the extent of these interactions and the specific elements of the conditions that ‘activate’ working memory resources. For example, the interaction of working memory with naturalistic learning in the acquisition of fluency should be more closely analyzed to determine which aspects of the context are implicated in the relationship, i.e. lexical versus semantic processing, etc.
Such interactions also motivate research into the interactions between learning contexts and other cognitive IDs, such as analytic or inductive language ability. Furthermore, the confirmation of such relationships will require research for L2s other than Russian.

The second insight this study provides is largely pedagogical and relates to the impact of different learning strategies on the major aspects of proficiency. In agreement with a number of past studies (e.g. Doughty, 2003; Loewen, 2002; Lyster, 2004; Schmidt, 1983), there appears to be a strong correlation between the extent of formal learning experiences and accuracy-related aspects of proficiency. This study’s findings related to fossilization in particular shed some light on this elusive and vital aspect of inter-learner variability in L2 attainment. Additional research related to the relationships between formal instruction and lower fossilization rates, and the tendency of morpho-syntactic elements to fossilize with continued naturalistic exposure could potentially uncover the exact causes of this phenomenon.

It is again evident that the design of the current study was limited in respect to a more close examination of accuracy in general and fossilization in particular. Future research will need to more accurately isolate the conditions under which fossilization occurs, with an emphasis on identifying the potential effects of naturalistic learning conditions on its occurrence. A longitudinal research program would potentially be ideally suited to discover these relationships. Data is also needed with respect to any possible relationships between fossilization and implicit types of language learning, which were not targeted in this study. Studies designed to discover the potential causes of the evident barrier to native-like accuracy and the consequent occurrence of fossilized forms in speech are of paramount importance to SLA research and could have long-term
future consequences for the improvement of both pedagogical practices and learner strategies.
REFERENCES


APPENDIX A: LANGUAGE BACKGROUND QUESTIONNAIRE

LANGUAGE BACKGROUND QUESTIONNAIRE

Participant Code: ___________

Please provide complete answers to the following questions about past experiences learning Russian. The information is important to ensure accurate results for the data that is collected.

1. When did you begin to study or learn Russian (month and year)? ________________________________
2. Why did you choose to learn Russian? ________________________________
3. How old were you when you began your study of Russian? ________________________________
4. Are you currently actively studying the language? □ Yes □ No If yes, how many hours/week: ______

5. On the chart below, please list all strategies you have used to learn Russian in chronological order. Include the location, dates, time spent, and the approximate number of hours spent actively learning each week. Choose from the following methods: a. classroom instruction with grammar explanations; b. naturally through interaction (no grammar); c. intensive course with grammar instruction; d. conversational practice with a tutor or group (no grammar); e. grammar tutoring with a native speaker; f. other (please describe)

<table>
<thead>
<tr>
<th>Method Used</th>
<th>Location &amp; dates</th>
<th>Duration</th>
<th>Hours per week</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Classroom w/explanation □</td>
<td>__________________</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>b. Naturally w/ interaction □</td>
<td>__________________</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>c. Intensive instruction □</td>
<td>__________________</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>d. Conversation partner □</td>
<td>__________________</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>e. Grammar tutoring □</td>
<td>__________________</td>
<td>______</td>
<td>______</td>
</tr>
<tr>
<td>f. Other □</td>
<td>Please explain:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. Important Note: “Native-like” or “Proficient” is defined as a native-like ability in three major areas: grammatical accuracy (endings or inflections), complex sentence structure and word order (ability to form native-like complex sentences), and fluency or native-like flow/rapidity of speech.

Based on the above definition, please characterize as closely as possible your current level of speaking proficiency in Russian:

□ Poor □ Fair □ Good □ Very good □ Native-like

4. Please characterize your current level of speaking proficiency in Russian:

□ Poor □ Fair □ Good □ Very good □ Proficient

5. If you checked “proficient” above, at what age did you acquire proficiency? ____________________

6. Have you ever had your proficiency tested? If so, describe the test and the rating you received:

_____________________________________________________________

7. What does “proficiency” or “fluency” mean to you? __________________________________________

8. What do you think is more important? □ Accuracy □ Just communicating □ Both

9. Please list any other foreign languages you have studied or learned and rate your proficiency:

__________________________________________________________ □ Poor □ Fair □ Good □ Very good □ Proficient

10. Have you ever taken a language aptitude test before (the MLAT or PLAB, etc.)? □ Yes □ No

11. Please choose the statement that best reflects your attitude about taking an experimental test.

□ I try hard on all tests □ It depends on my mood □ I just want to get it over with

12. Can you promise that you will try hard on every question during testing? □ Yes □ No

Thank you for taking the time to fill out this questionnaire!
APPENDIX B: PROFICIENCY COMPONENT TEST MEASURES

Part I: General Proficiency Interview (Levels 1+, 2 and 4)

Level 1+  Situation: Вы сейчас на вокзале, и хотите купить билет на поезд в Москву. Я здесь работаю и продаю билеты. Купите у меня столько билетов, сколько нужно и отвечайте на вопросы. Понимаете? Давайте!

Куда вы хотите поехать?
Когда вы хотите поехать в Москву?
Во сколько вы хотите приехать в Москву?
Как долго вы будете в Москве?
Почему вы едете в Москву?

Опишите, пожалуйста, эту фотографию. Расскажите подробно о том, что вы видите.

1+ Situation: Right now you are at the train station and you want to buy tickets for the train to Moscow. And I work here selling tickets. Buy the tickets you need and answer my questions. Do you understand? Let’s go!

Where do you want to go?
When do you want to leave?
At what time would you like to arrive in Moscow?
How long will you be in Moscow?
Why are you traveling to Moscow?
How many tickets do you need?
Do you want an open car or a sleeper?
Do you need a return ticket?
Can I see your passport, please?

Please describe this photograph in detail. What do you see?

Descriptors for level 1+:
Does student have the skills to communicate in basic social situations beyond giving rudimentary biographical info? Can he/she initiate an exchange?

Can student speak with basic grammatical accuracy?
Does student show hesitation in his/her responses? If he/she fails to comprehend all the questions and cannot answer after instructor paraphrases, score as a level 1.

Can student provide a detailed description of something he/she sees?

Level 2: Ситуация: вы стоите на улице и видите вашего русского друга, который хочет знать о вашей учёбе. Подойдите к нему и расскажите ему о своей учёбе.

A как ты обычно проводишь своё свободное время?


Situation: ваши друзья стоят на улице и болтают. А вы хотите сказать им о вечеринке, которая будет у вас дома. Что вы скажете, чтобы узнать о том, о чём они разговаривают? (Russian instructor begins to ask the student questions about the party this weekend)

Каким образом вы будете готовиться к вечеру с друзьями, чтобы они могли расслабиться и получить удовольствие?

Level 2: Situation: You are standing on the street and you see your Russian friend who wants to find out about your studies. You walk up to him and tell him all about it.

And how do you spend your free time?

What is happening right now in Russia/Ukraine/Kazakhstan? How do you like life there? What can you tell me about the culture there? Who is the President? Tell me about the town where you live.

Situation: Your friends are standing around and talking. And you want to tell them about a party that you are having at your house. What will you say in order to find out about what they’re talking about? (Instructor asks about the party this weekend)

In what way will you prepare for the party with your friends, so that they will be able to relax and enjoy themselves?
Descriptors for level 2:

Can student manage an everyday conversation and talk about his work, current events and other subjects of a personal nature? During this exchange, instructor must interrupt student and ask for clarification.

Does student control basic grammatical constructions and word use well? If student fails to get this far, score a 1+.

Can student describe personal and special fields of interest in detail? Are there weaknesses, though student is able to participate with ease? At this level students should be able to elaborate in their answers. If he fails to understand or properly respond to your last question, score 2-.

Level 4

You want to work in the former Soviet Union, perhaps in the Russian Federation. What kind of visa is needed for this? Please elaborate.

Some Americans believe that before going to Russia it is necessary to study the Russian language, and others feel that you can operate there by using translators. What do you think about this? Back up your idea and clarify your answer with examples.

Situation: You have been invited to a gathering at a university and unexpectedly are asked to give a talk on the topic: What does American culture have in common with Russian culture? What are the differences? Can Russia ever truly coexist with America?) Explain.

Descriptors for level 4:

Is student able to create the atmosphere necessary for formal or informal interactions related to his profession? If communication breaks down in his/her answer, or after clarification student fails to understand the question, score a 3+
Can the student express opinions related to personal experience but on
topics which may or may not have been explored in the past? If he
fails to comprehend the question, score a 3+. Score a 4- if answer is
adequate but he fails to complete the last question.

Can student communicate effectively in many different settings, such
as debates, lectures, conferences, etc.? If student fails to comprehend
this question after clarification, score a 4-. If he fails to adequately
address the question using circumlocution and other means, score a 4-
as well.

PART II: FLUENCY MONOLOGUE

Tell me, please, about what you are interested in. What is your favorite hobby or
subject? You can speak about any topic at all, for example a field of study or your
everyday life here. Tell me everything you can about it as well as why this subject is so
interesting to you.

PART III: ACCURACY ELICITATION TASK

1. Relative pronoun ‘kotoryj’ (‘who, which’)

Situation a): Преподаватель говорит помощнику о студентах:

Преподаватель: «Вот у этой студентки нет учебника.»

Вопрос: О какой студентке говорит преподаватель?
Ответ: Он говорит о студентке, у которой нет учебника.

Situation a): A teacher is talking to his aid about the students:

Teacher: “Now that (female) student does not have a textbook.”

Question: ‘About which student is the teacher talking?’

Answer: ‘He is talking about the student (LOC/SG/F) with (by or near) whom
(GEN/SG/F) is not textbook (GEN/SG/M).’

‘He is talking about the student who does not have a textbook.’

Situation б): Валерий говорит об американцах:

Валерий: «Вот американцы. Я вчера ними поговорил.»
Вопрос: О каких американцах говорит Валерий?
Ответ: Он говорит об американцах, с которыми он вчера поговорил.

Situation b): Valerij is talking about (some) Americans:
   Valerij: “Here are the Americans. I spoke with them yesterday.”

Question: ‘About which Americans is Valerij talking?’
Answer: ‘He is talking about the Americans (LOC/PL) with whom (INSTR/PL) he spoke yesterday.’
   He is talking about the Americans that he spoke with yesterday.

2. Subjunctive Conjunction ‘chtoby’ (‘that be’):

Слушайте ситуацию и ответьте на вопрос ‘Listen to the situation and answer the question’:

Ситуация а): Лариса приглашает Машу на дачу:
   «Маша, приезжай к нам на дачу!»

Вопрос: Что хочет Лариса?
Ответ: Она хочет, чтобы Маша приехала к ним на дачу.

Situation a): Larisa is inviting Masha to his family’s cottage:
   “Masha, come visit us out at the dacha!”

Question: ‘What does Larisa want?’
Answer: ‘She wants so that Masha come (PAST F) to them to cottage.’
   ‘She wants her to come visit them at the cottage.’

Ситуация б): Ивану холодно и он поэтому обращается к друзьям и говорит:
   «Ребята, закройте окно, пожалуйста!»

Вопрос: Что хочет Иван от друзей?
Ответ: Он хочет, чтобы они закрыли окно.

Situation b): Ivan is cold and he therefore turns to his friends and says:
   ‘Guys, please close the window!’

Question: ‘What does Ivan want of his friends?’
Answer: ‘He wants so that they close (PAST PL) the window.’
   ‘He wants them to close the window.’

3. Conjunction ‘to, chto’ (‘that, what’ (‘that, which’):

Слушайте ситуацию и ответьте на вопрос (listen to the situation and answer the question):
Ситуация а): Говорит Василий Андрею:
«Андрей, знаешь чем Лариса интересуется? Я хочу рассказать тебе об этом.»

Вопрос: О чём хочет рассказать Василий Андрею?
Ответ: Он хочет рассказать ему о том, чем интересуется Лариса.

Situation a): Vasilij is talking to Andrej:
“Andrej, you know what Larisa is interested in? I want to tell you about it.”

Question: ‘About what does Vasilij want to tell Andrej?
Answer: ‘He wants to tell him about that (LOC/SG), what (INSTR/SG) Larisa is interested.’
‘He wants to tell him about what Larisa is interested in.’


Вопрос: О чём думает Лариса?
Ответ: Она думает о том, чего она боится.

Situation b): Larisa is thinking: Oh! I am very afraid of that.

Question: ‘About what is Larisa thinking?’
Answer: ‘She is thinking about that (LOC/SG), what (GEN/SG) she is afraid.’
‘She is thinking about what she is afraid of.’
APPENDIX C: WM BY COMPONENTS OF PROFICIENCY

Table C.1 Naturalistic Learners: WMC/Experience by Accuracy Regression

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>37.234</td>
<td>2</td>
<td>18.617</td>
<td>6.094</td>
<td>.021</td>
</tr>
<tr>
<td>Intercept</td>
<td>4.864</td>
<td>1</td>
<td>4.864</td>
<td>1.592</td>
<td>.239</td>
</tr>
<tr>
<td>Experience</td>
<td>1.155</td>
<td>1</td>
<td>1.155</td>
<td>.378</td>
<td>.554</td>
</tr>
<tr>
<td>WMC</td>
<td>37.065</td>
<td>1</td>
<td>37.065</td>
<td>12.133</td>
<td>.007</td>
</tr>
<tr>
<td>Error</td>
<td>27.495</td>
<td>9</td>
<td>3.055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>303.250</td>
<td>12</td>
<td>27.495</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>64.729</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .575 (Adjusted R Squared = .481)

Table C.1 above depicts the results of the logistic regression for WM and overall experience for accuracy for naturalistic condition learners. The level of significance is presented in boldface type, with significance seen in the relationship for working memory capacity (WMC), but not for amount of experience. Similar results are seen in relation to the rate of potential fossilization in Table C.2.

Table C.2 Naturalistic Learners: WMC/Experience by Fossilization Rate Regression

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>56.407</td>
<td>4</td>
<td>14.102</td>
<td>16.178</td>
<td>.005</td>
</tr>
<tr>
<td>Intercept</td>
<td>93.025</td>
<td>1</td>
<td>93.025</td>
<td>106.721</td>
<td>.000</td>
</tr>
<tr>
<td>WMC</td>
<td>53.857</td>
<td>2</td>
<td>26.929</td>
<td>30.893</td>
<td>.002</td>
</tr>
<tr>
<td>EXPERIENCE</td>
<td>3.101</td>
<td>2</td>
<td>1.551</td>
<td>1.779</td>
<td>.261</td>
</tr>
<tr>
<td>Error</td>
<td>4.358</td>
<td>5</td>
<td>.872</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>208.990</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>60.765</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .928 (Adjusted R Squared = .871)
Table D.1 Naturalistic Condition: WMC by Naturalistic Experience Moderated Regression for Fluency

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>59.667*</td>
<td>5</td>
<td>11.933</td>
<td>3.703</td>
<td>.071</td>
</tr>
<tr>
<td>Intercept</td>
<td>327.758</td>
<td>1</td>
<td>327.758</td>
<td>101.718</td>
<td>.000</td>
</tr>
<tr>
<td>EXPERIENCE</td>
<td>9.549</td>
<td>2</td>
<td>4.775</td>
<td>1.482</td>
<td>.300</td>
</tr>
<tr>
<td>WMC</td>
<td>29.622</td>
<td>2</td>
<td>14.811</td>
<td>4.597</td>
<td>.062</td>
</tr>
<tr>
<td>EXPERIENCE * WMC</td>
<td>20.485</td>
<td>1</td>
<td>20.485</td>
<td>6.357</td>
<td>.045</td>
</tr>
<tr>
<td>Error</td>
<td>19.333</td>
<td>6</td>
<td>3.222</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>586.000</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>79.000</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .755 (Adjusted R Squared = .551)

Table D.1 demonstrates significance for the interaction between WM and naturalistic experience for fluency among naturalistic learners. Naturalistic learning appears to ‘moderate’ the effect of WM resources to produce fluency.

Table D.2 provides results for the moderated regression experiment for learner accuracy under the FN condition. The results appear to show that the role of a learner’s WM resources is somehow moderated by the nature of formal learning.

Table D.2 FN Learners: WMC by percent of Formal Training Moderated Regression for Accuracy

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>42.103*</td>
<td>7</td>
<td>6.015</td>
<td>13.880</td>
<td>.005</td>
</tr>
<tr>
<td>Intercept</td>
<td>363.103</td>
<td>1</td>
<td>363.103</td>
<td>837.929</td>
<td>.000</td>
</tr>
<tr>
<td>WMC * FPERCENT</td>
<td>32.444</td>
<td>2</td>
<td>16.222</td>
<td>37.436</td>
<td>.001</td>
</tr>
<tr>
<td>FPERCENT</td>
<td>11.095</td>
<td>4</td>
<td>2.774</td>
<td>6.401</td>
<td>.033</td>
</tr>
<tr>
<td>WMC</td>
<td>.190</td>
<td>1</td>
<td>.190</td>
<td>.440</td>
<td>.537</td>
</tr>
<tr>
<td>Error</td>
<td>2.167</td>
<td>5</td>
<td>.433</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>626.500</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>44.269</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .951 (Adjusted R Squared = .883)
The Tukey post-hoc analysis for the formal experience moderated regression is provided in Table D.3, which shows a fairly significant overall interactive relationship, although the significance in the interaction between WM and formal experience does not occur between all levels of formal experience and WM.

Table D.3 FN learners: Interaction of WM and Accuracy Parameter Estimates

<table>
<thead>
<tr>
<th>Dependent Variable:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>8.000</td>
<td>.329</td>
<td>24.306</td>
<td>.000</td>
<td>7.154</td>
<td>8.846</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[WMC=2.00] * [FPERCENT=1.00]</td>
<td>-4.000</td>
<td>.736</td>
<td>-5.435</td>
<td>.003</td>
<td>-5.892</td>
<td>-2.108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[WMC=2.00] * [FPERCENT=2.00]</td>
<td>-1.333</td>
<td>.503</td>
<td>-2.652</td>
<td>.045</td>
<td>-2.626</td>
<td>-.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[WMC=2.00] * [FPERCENT=3.00]</td>
<td>-1.354E-15</td>
<td>.736</td>
<td>.000</td>
<td>1.000</td>
<td>-1.892</td>
<td>1.892</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[WMC=2.00] * [FPERCENT=4.00]</td>
<td>1.000</td>
<td>.736</td>
<td>1.359</td>
<td>.232</td>
<td>-0.892</td>
<td>2.892</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[WMC=2.00] * [FPERCENT=5.00]</td>
<td>-5.000</td>
<td>.736</td>
<td>-6.794</td>
<td>.001</td>
<td>-6.892</td>
<td>-3.108</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[WMC=3.00] * [FPERCENT=1.00]</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[WMC=3.00] * [FPERCENT=2.00]</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[WMC=3.00] * [FPERCENT=3.00]</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[WMC=3.00] * [FPERCENT=4.00]</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[WMC=3.00] * [FPERCENT=5.00]</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. This parameter is set to zero because it is redundant.
Table D.4 FN Learners: WMC by percent of Formal Training Moderated Regression for Native-likeness

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>13.400*</td>
<td>7</td>
<td>1.914</td>
<td>4.336</td>
<td>.063</td>
</tr>
<tr>
<td>Intercept</td>
<td>374.325</td>
<td>1</td>
<td>374.325</td>
<td>847.849</td>
<td>.000</td>
</tr>
<tr>
<td>WMC * FPERCENT</td>
<td>10.674</td>
<td>2</td>
<td>5.337</td>
<td>12.088</td>
<td>.012</td>
</tr>
<tr>
<td>FPERCENT</td>
<td>3.736</td>
<td>4</td>
<td>.934</td>
<td>2.115</td>
<td>.216</td>
</tr>
<tr>
<td>WMC</td>
<td>.014</td>
<td>1</td>
<td>.014</td>
<td>.033</td>
<td>.864</td>
</tr>
<tr>
<td>Error</td>
<td>2.208</td>
<td>5</td>
<td>.442</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>553.220</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>15.608</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. R Squared = .859 (Adjusted R Squared = .661)

Though not part of the original design, the moderated regression for WM and formal experience demonstrated significance in the interaction for the general proficiency or native-likeness score for formal/naturalistic learners, as shown in Table D.4. The relationship indicates that while WM alone does not predict overall proficiency in this condition, there is an interaction between WM and formal experience that impacts overall proficiency.
Table E.1 FN learners: WMC/formal experience by Fossilization Rate

<table>
<thead>
<tr>
<th>Source</th>
<th>Type III Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected Model</td>
<td>.996&lt;sup&gt;a&lt;/sup&gt;</td>
<td>2</td>
<td>.498</td>
<td>4.777</td>
<td>.057</td>
</tr>
<tr>
<td>Intercept</td>
<td>.005</td>
<td>1</td>
<td>.005</td>
<td>.045</td>
<td>.840</td>
</tr>
<tr>
<td>FormalPercent</td>
<td>.939</td>
<td>1</td>
<td>.939</td>
<td>9.006</td>
<td>.024</td>
</tr>
<tr>
<td>WMC</td>
<td>.314</td>
<td>1</td>
<td>.314</td>
<td>3.011</td>
<td>.133</td>
</tr>
<tr>
<td>Error</td>
<td>.626</td>
<td>6</td>
<td>.104</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9.840</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>1.622</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> R Squared = .614 (Adjusted R Squared = .486)

Table E.1 shows results for the regression experiment related to the amount of formal experience as a predictor of proficiency for the two combined learning conditions. The results of the experiment found significance for the relationship between formal experience and the rate of potentially fossilized forms for the formal/naturalistic condition, although not for the NF context. These results indicate a significant predictive relationship between the extent of formal training and lower rates of potential fossilization for FN learners.