LEVEL OF INTAKE, DEPTH OF PROCESSING, 
AND TYPE OF LINGUISTIC ITEM IN L2 DEVELOPMENT

A Dissertation 
Submitted to the Faculty of the 
Graduate School of Arts and Sciences 
of Georgetown University 
in partial fulfillment of the requirements for the degree of 
Doctor of Philosophy 
in Spanish

by

Anne M. Calderón, M.A.

Washington, D.C. 
May 27, 2014
LEVEL OF INTAKE, DEPTH OF PROCESSING,
AND TYPE OF LINGUISTIC ITEM IN L2 DEVELOPMENT

Anne M. Calderón, M.S.

Mentor: Ronald P. Leow, Ph.D.

ABSTRACT

L2 learners must employ selective attention minimally to isolate linguistic features in input to be able to take them in. Based on previous theoretical postulations and empirical research, Leow (forthcoming) postulates in his Model of the L2 learning process three levels of intake, namely, noticed intake, detected intake, and attended intake while underscoring the important role for depth of processing. This study sought to test the tenets of his Model’s early stages of the L2 learning process by addressing the potential existence of different levels of intake and the role of depth of processing during these early stages. The study also addressed whether type of linguistic item (grammatical versus lexical) plays a role. To explicate the roles of level of intake, depth of processing, type of linguistic item, and reactivity in adult L2 learner’s subsequent intake, the present study employed eye-tracking and concurrent verbal reports: 96 beginning learners of Spanish read a text and then completed production, recognition, and comprehension assessments in a pretest/posttest design. Results revealed no reactivity and that different levels of intake do appear to exist, that depth of processing not only may play a role in subsequent processing of intake but also appears to facilitate the deeper processing needed for incorporation of intake into the developing system, as postulated by Leow’s Model of the L2 learning process (forthcoming). Finally, the results also revealed differences in processing type of linguistic item.
Acknowledgements

After six years and three babies, I have finally accomplished my life-long goal of writing a dissertation and getting my Ph.D. This would have been an impossible task without the endless support and help of so many colleagues, family, and friends.

First I want to thank my dissertation mentor, Ron Leow. Ron, you have given me so many opportunities to excel and grow intellectually. I have tried my best to use them all to the fullest. You have also been extremely supportive and understanding regarding my family life and the responsibilities I have at home. I can’t think of anyone with whom I would have rather worked over the past six years.

I am also very grateful to those who served on my dissertation committee: Aline Godfroid, Melissa Bowles, and Cristina Sanz. Whether near or far, all of you were incredibly helpful whenever I needed guidance or had questions. Aline, thank you so much for tirelessly answering all my eye-tracking and statistical questions via email. You went above and beyond to help me in every way. Melissa, you helped me so much by facilitating my data collection in Champaign. And Cristina, thank you for your thoughtful feedback on my chapters even when you were busy with many other grad students. I am also much indebted to Dr. Chen, who tirelessly assisted me with the experimental design and every aspect of the statistical analyses with unlimited patience. I would also like to extend a sincere thanks to José Ignacio Hualde for introducing me to linguistic research and supporting my early academic efforts.

Many thanks also to the Department of Spanish and Portuguese at Georgetown University and the Department of Spanish, Italian and Portuguese at the University of Illinois, Urbana-Champaign, both of which facilitated my data collection process.
Two eye-tracking labs graciously allowed me to use their facilities and equipment to test participants: the Cognitive Neuropsychology Laboratory at Georgetown University and the Second Language Acquisition and Bilingualism Lab at the University of Illinois. Heidi Getz, Kelli Sullivan, Sarah Ferguson Snider, and Myeonghyeon Kim, you were all a great help with scheduling and use of the equipment. I am indebted to Marcus Johnson from SR Research for his unending patience and help in developing my eye-tracking task.

I was so fortunate to work with wonderful research assistants who were always there when I most needed them: Adam Barton, Cindy Borges, Allison Caras, Caylah Green, Young Sun Lee, Erin Shores, and Katherine Thompson. Thank you for taking my data collection as seriously as I did! Also, a big thanks to Janire Zalbidea for being my second coder.

I could have never completed this dissertation without the love and support from my family and friends. Ernesto, thank you for your endless patience with my studies over the past six years. You have always been so encouraging and proud of me. And to think that it was originally your idea that I go back to school! You always knew how important this was to me and I love you for it. Ernesto papá y Yolanda, ¡qué bendición es ser tu hija! Gracias por tu apoyo incondicional y todos los sacrificios que hicieron ustedes para que yo pudiera graduarme. And my peanuts, Siena, Luca, y Liam: gracias por darle a mami un tiempito para que estudiara. Yo sé que eso ha sido un gran sacrificio para ustedes también porque no he podido estar siempre en casa con ustedes. Los amo con todo el corazón y pronto verán la recompensa: ¡mami en casa!

Thank you to the people who kept me grounded during this process and always understood when my time and energy available to maintain friendships was very limited:
Yamilette Colón-Cales, Jessie Cox, Cara Gibson, Nube Maurizaca, and Nicole Rumeau.

And to my dear sister Katie, thank you for just being you. You are always the first one to brag about me and defend me, and I love you for it!

Most importantly, I want to thank my parents, Hal and Barbara Thinglum. Through your examples and the time you spent with me, you showed me the value of hard work and the importance of education. From helping me study for spelling tests in the 2\textsuperscript{nd} grade to making sure I had every physical comfort I need during college to be able to excel, you two have encouraged me every step of the way. You have been my cheering team in achieving all my goals. I love you so much. This is for you.
# Table of Contents

Chapter One: Statement of the Problem and Definitions of Terms ................................................. 1  
Statement of the Problem .................................................................................................................. 1
Definitions of Terms ....................................................................................................................... 11

Chapter Two: Review of the Literature ............................................................................................. 16  
Frameworks in Non-SLA Fields for the Roles of Attention and Awareness in Learning .......... 16
  Attention ........................................................................................................................................ 16
  Measuring Attention ....................................................................................................................... 17
  Empirical Studies .......................................................................................................................... 18

Theoretical Underpinnings in the SLA Field for the Roles of Attention and Awareness in Learning ................................................................. 22
  McLaughlin (1987) ......................................................................................................................... 23
  Gass (1988 and elsewhere) ........................................................................................................... 25
  Schmidt’s (1990 and elsewhere) noticing hypothesis ................................................................. 28
  Tomlin and Villa (1994) ................................................................................................................ 28
  Robinson (1995) .......................................................................................................................... 30
  Truscott and Sharwood Smith (2011) .......................................................................................... 33
  Summary ........................................................................................................................................ 34

Intake ............................................................................................................................................. 35
  Assessment Tasks Employed to Measure Intake ........................................................................ 38
Empirical studies. ........................................................................................................................................ 41
High level of intake..................................................................................................................................... 41
Medium level of intake. ............................................................................................................................. 48
Low level of intake. .................................................................................................................................... 50
Summary..................................................................................................................................................... 52
Variables postulated to be associated with intake. .................................................................................... 53
Depth of processing during input processing............................................................................................... 53
Theoretical accounts .................................................................................................................................... 54
Operationalizing depth of processing............................................................................................................ 56
Empirical accounts ....................................................................................................................................... 57
Awareness in input processing....................................................................................................................... 62
Empirical research on awareness at the level of noticing. ............................................................................. 63
Leow’s Model of the L2 Learning Process. ................................................................................................. 66
Input processing stage................................................................................................................................... 67
Intake processing stage. ............................................................................................................................... 69
Internal system.............................................................................................................................................. 71
Type of Linguistic Item................................................................................................................................. 73
Empirical studies. ....................................................................................................................................... 74
Type of linguistic item in corrective feedback studies. .................................................................................. 75
Type of linguistic item in attentional/processing/comprehension studies. .................................................... 79
Summary................................................................................................................................................. 89

Methodological Approaches to Operationalizing and Measuring Attention. ......................... 91

Concurrent vs. retrospective verbal reports. ............................................................................. 92

Empirical studies using concurrent verbal reports............................................................... 94

The issue of reactivity. ................................................................................................................. 95

Coding noticing, awareness, and similar constructs. ........................................................... 105

Eye-tracking as a measure of processes related to the attentional strand. ......................... 107

Conclusions for the Role of Noticing in SLA .......................................................................... 127

Rationale and Research Questions ......................................................................................... 128

Rationale for the present study. ............................................................................................. 128

Research questions. .................................................................................................................. 129

Chapter Three: Research Design and Methodology .............................................................. 131

Pilot Study ........................................................................................................................................ 131

Pilot methodology..................................................................................................................... 131

Pilot procedure.......................................................................................................................... 133

Pilot coding.................................................................................................................................... 133

Pilot results and discussion....................................................................................................... 133

Pilot modifications...................................................................................................................... 135

Current Study ............................................................................................................................ 137

Method.......................................................................................................................................... 137
Participants ....................................................................................................................... 137

The target grammatical form and target lexical items. ................................................. 139

The target grammatical form ....................................................................................... 140

The target lexical items. ............................................................................................... 142

Rationale for the selection of the target lexical items. .............................................. 142

The control grammatical and lexical items. ................................................................. 143

The control grammatical form ..................................................................................... 143

The control lexical items. .............................................................................................. 144

Experimental conditions .............................................................................................. 144

Level of intake conditions. ............................................................................................ 144

Determination of final group membership. ................................................................. 153

Latin square design ...................................................................................................... 155

Materials ......................................................................................................................... 157

Experimental reading passage ...................................................................................... 157

Pretest and immediate posttest assessments. ............................................................... 160

The controlled written production pretest and immediate posttest. ......................... 160

The items pretest. ........................................................................................................... 161

The verb stem meaning pretest .................................................................................... 162

Immediate recognition posttest. .................................................................................. 163

Comprehension posttest. ............................................................................................... 164
Questionnaires .......................................................................................................................... 165
Debriefing questionnaire ........................................................................................................... 165
Language contact profile ........................................................................................................ 166
Practice tasks .......................................................................................................................... 166
Practice TA protocol task ......................................................................................................... 166
Practice eye-tracking task ........................................................................................................ 166
Procedure .................................................................................................................................. 167
Coding and Scoring ................................................................................................................... 168
Pre- and posttest measures ....................................................................................................... 168
Criteria for correct answers ..................................................................................................... 170
Criteria for inclusion of data in statistical analysis ................................................................... 170
Assignment into grammatical and lexical intake groups ............................................................ 172
Depth of processing think-aloud protocols .............................................................................. 175
Fixation measures ..................................................................................................................... 178

Chapter Four: Results ............................................................................................................. 179
Statistical Notes ....................................................................................................................... 179

Why percentages are used for lexical items ............................................................................. 179
Why subtracted fixation values are used for eye-tracking measures ........................................ 180
Tests of Validity ....................................................................................................................... 181

Text order .................................................................................................................................. 181
Were posttest scores higher for target words than for control words? .............................................. 181

Grammatical intake groups. .................................................................................................................. 182

Lexical intake groups. ......................................................................................................................... 184

Summary. ........................................................................................................................................... 185

Were posttest scores significantly higher than pretest scores? ......................................................... 186

Grammatical intake groups. .................................................................................................................. 186

Lexical intake groups. .......................................................................................................................... 188

Summary. ........................................................................................................................................... 189

Results of verb stem meaning pretest. ............................................................................................... 189

Summary ............................................................................................................................................. 190

Research Questions ............................................................................................................................. 191

Research question one. ......................................................................................................................... 191

Does thinking aloud while performing a reading task have any ....................................................... 192
effect on beginning Spanish readers’ a) recognition, b), production, and c) 
comprehension of a grammatical form when compared to readers not thinking 
aloud? .................................................................................................................................................. 192

Does thinking aloud while performing a reading task have any ....................................................... 194
effect on beginning Spanish readers’ a) recognition, b), production, and c) 
comprehension of lexical items when compared to readers not thinking aloud? ...................... 194

Research question two. ....................................................................................................................... 195
Does level of intake have an effect on adult L2 learner’s recognition of a grammatical form? .............................................................. 195

Does level of intake have an effect on adult L2 learner’s recognition of a lexical form? .............................................................. 197

Research question three. .............................................................................................................. 198

Does level of intake have an effect on controlled written production of a grammatical form? .............................................................. 198

Does level of intake have an effect on controlled written production of a lexical form? .............................................................. 199

Research question four. .............................................................................................................. 200

Are there any significant relationships between fixation measures and recognition of a grammatical form? .............................................................. 200

Are these any significant relationships between fixation measures and recognition of a lexical form? .............................................................. 203

Research question five. .............................................................................................................. 208

Is there a relationship between depth of processing, as measured by form-meaning connections in think-aloud protocols during a reading task, and subsequent recognition and written production of a grammatical form embedded in the reading task? .............................................................. 208

Is there a relationship between depth of processing, as measured by form-meaning connections in think-aloud protocols during a reading task, and subsequent recognition and written production of lexical items embedded in the reading task? .... 210
Research question six. ........................................................................................................ 211

Is there a relationship between depth of processing, as measured by form-meaning
connections in think-aloud protocols during a reading task, and comprehension of
a reading passage, as measured by a multiple-choice comprehension test? ............ 212

Research question seven. ........................................................................................................ 213

Are there any differences between recognition of a) a grammatical form and b)
lexical items in beginning Spanish learners from a written text as measured by
recognition tests? ........................................................................................................ 214

Overall Summary .................................................................................................................. 217

Chapter Five: Discussion and Conclusion ........................................................................ 218

Discussion ............................................................................................................................ 218

Effect of Thinking Aloud on Recognition, Production, and Comprehension. ............ 218

Effect of Level of Intake on Recognition ............................................................................. 219

Effect of Level of Intake on Production ............................................................................. 222

Relationship between Eye Fixation Time Measures and Recognition ....................... 225

Relationship between Depth of Processing and a) Recognition and b) Production .... 227

Relationship between Depth of Processing and Comprehension .................................... 230

Difference between Recognition of Grammatical Form vs. Lexical Item .................. 232

Limitations and Directions for Future Research ............................................................... 235

Conclusions and Implications ......................................................................................... 237

Appendix A: Language Background Questionnaire ......................................................... 246
Appendix B: Experimental Reading Passages ................................................................. 249
Appendix C: Reading Instructions by Condition ............................................................. 254
Appendix D: Items Pretest .............................................................................................. 257
Appendix E: Recognition Posttest .................................................................................. 265
Appendix F: Comprehension Posttest ........................................................................... 272
Appendix G: Comprehension Posttest Answer Sheet .................................................... 293
Appendix H: Debriefing Questionnaire ........................................................................ 294
Appendix I: Practice Think-aloud Protocol Task .......................................................... 296
Appendix J: Practice Eye-tracking Task ........................................................................ 297
References ..................................................................................................................... 298
Tables

Table 1: Summary of Past Studies Comparing Two Grammatical Items or Two Lexical Items ..........81
Table 2: Participant Characteristics, Overall and by Intake Group .............................................. 138
Table 3: Participant Self-Rated Proficiency, Overall and by Intake Group ....................................... 139
Table 4: Preterit Conjugation of –ar Verb (estudiar) ....................................................................... 140
Table 5: Intake Characteristics .......................................................................................................... 147
Table 6: Experimental Condition Characteristics ............................................................................. 148
Table 7: Latin Square Design ............................................................................................................. 155
Table 8: Methods of Experimental and Control Conditions .............................................................. 168
Table 9: Time Spent (in Seconds) on Reading Passages .................................................................... 173
Table 10: Criteria for Grammatical and Lexical Depths of Processing .............................................. 177
Table 11: Summary Results for Paired-Samples T-Tests for Instances of Target vs. Control
Grammatical Forms .............................................................................................................................. 183
Table 12: Summary Results for Paired-Samples T-Tests for Instances of Target vs. Control
Lexical Forms ....................................................................................................................................... 185
Table 13: Summary Results for Paired-Samples T-Tests for Differences in Grammatical Scores
between Pretests and Posttests ............................................................................................................... 187
Table 14: Summary Results for Paired-Samples T-Tests for Differences in Lexical Scores
between Pretests and Posttests ............................................................................................................... 188
Table 15: Descriptive Statistics for the First Administration of the Verb Stem Meaning Pretest² .......... 190
Table 16: Descriptive Statistics for RQ1 .................................................................................................. 193
Table 17: ANOVA results of RQ1 ........................................................................................................... 193
Table 18: Descriptive Statistics for RQ2 ................................................................................................ 196
Table 19: ANOVA results of RQ2 ........................................................................................................... 196
Table 20: Homogenous subsets from Scheffé post hoc on Grammatical Level of Intake and Recognition Posttest .......................................................... 196

Table 21: Homogenous subsets from Scheffé post hoc on Lexical Level of Intake and Recognition Posttest .......................................................... 198

Table 22: Descriptive Statistics for RQ3 .............................................................................................................................. 199

Table 23: ANOVA results of RQ3 ............................................................................................................................................ 199

Table 24: Descriptive Statistics for Fixation Measures and Recognition of a Grammatical Form .................. 201

Table 25: Correlations between Fixation Measures and Recognition of a Grammatical Form .................. 202

Table 26: Descriptive Statistics for Fixation Measures and Recognition of a Lexical Form ..................... 205

Table 27: Results of Generalized Linear Mixed model for Lexical Noticed Intake Group ..................... 205

Table 28: Correlations between Depth of Processing and a) Recognition and b) Production of a Grammatical Form and Lexical Items ...................................................................................... 209

Table 29: Relationships between Depth of Processing and Comprehension of a Grammatical Form and Lexical Items .......................................................................................................................... 212

Table 30: Summary Results for Paired-Samples T-Tests for Difference between Grammatical and Lexical Recognition .................................................................................................................. 215
Chapter One: Statement of the Problem and Definitions of Terms

Statement of the Problem

A great majority of theorists and researchers in the fields of SLA and cognitive science concur that attention is a necessary condition for learning (e.g., Carr & Curran, 1994; Dulany, Carlson & Dewey, 1984, 1985; Nissen & Bullemer, 1987; Reber, 1967, 1976, 1989, 1993; Robinson, 1995a; Schmidt, 1990, 1993, 1994, 1995, 2001; Shiffrin & Schneider, 1977; Tomlin & Villa, 1994). Nonetheless, there has been no consensus regarding the mechanisms that make up the human attentional system, thus giving rise to several theoretical accounts of attention. Among these are Broadbent’s (1958) model of a selective attention filter under voluntary control, Deutsch and Deutsch’s (1963) late selection model positing parallel processing of stimuli, and Wicken’s (1989) multiple pools of attentional resources.

The role of awareness in SLA and cognitive science, however, is more controversial. Some theoretical accounts (Robinson, 1995; Schmidt, 1990, 1993, 1994, 1995, 2001) argue that noticing is necessary for successful SLA and critically includes awareness, which itself requires attention. This view makes attention and awareness difficult to separate given that they are essentially “two sides of the same coin” (Schmidt, 2001, p. 5). Other theoretical accounts do not make a strong case for awareness; Tomlin and Villa (1994) claim that awareness is not required in order to make input available for intake. In their view, awareness may enhance input processing, but it is not required. Robinson (1995) reconciles the differing views of Schmidt (1990) and Tomlin and Villa (1994) by placing noticing farther along the acquisitional processing timeline than detection; he also concurs with Schmidt that a lack of awareness precludes learning.

Empirical studies in both cognitive science and cognitive psychology (e.g., Cohen, Ivry,
& Keele, 1990; Reber, 1967, 1976, 1989, 1993) and SLA (e.g., Leung & Williams, 2011, 2012; Williams, 2004, 2005) have reportedly evidenced learning without awareness, however others (e.g., Curran & Keele, 1993; Dulany, Carlson, & Dewey, 1984; Nissen & Bullemer, 1987) have shown that the only learning able to occur without awareness is non-significant in terms of the task of learning an L2.

The attention and awareness strand of SLA (e.g., Leow, 1997a, 1997b, 1998a, 1998b, 2000, 2001a, 2001b; Rosa & Leow, 2004a, 2004b; Rosa & O’Neill, 1999) has gained momentum and depended heavily on concurrent verbal reports as the main methodology to measure learners’ depth of processing and awareness of certain features of the L2 input. Considering that these concurrent verbal reports are able to show the learner’s cognitive processes while interacting with the L2 (Bowles, 2010), the verbalizations have been used to evidence different levels of several cognitive concepts such as awareness (e.g., Leow, 1997, 2000; Rosa & Leow, 2004b; Rosa & O’Neill, 1999; Sachs & Suh, 2007) and depth of processing (e.g., Leow, Hsieh, & Moreno, 2008; Morgan-Short, Heil, Botero-Moriarty, & Ebert, 2012, defined in this study as the amount of attention, cognitive effort, or time spent processing or elaborating on the target item in the input, and/or noticing induced by different types of experimental tasks or conditions (cf. e.g., Robinson, 2003; Hulstijn, 2001; Laufer & Hulstijn, 2001; Leow, Hsieh, & Moreno, 2008)

Given that current SLA theoretical underpinnings (e.g., McLaughlin, 1985; Gass, 1988; Robinson, 1995; Schmidt, 1990, 1993, 1994, 1995, 2001; Tomlin & Villa, 1994; VanPatten, 1996; Leow, forthcoming) postulated to account for the early stages of the learning processes in SLA all posit a role for attention in the L2 learning process, it seems logical that attention should also play a role in intake. Although there is no consensus regarding the definition of intake, most current definitions of intake do agree that it is an essential component of SLA and state either
explicitly or implicitly that it is a subset of the input (e.g., Corder, 1967; Faerch & Kasper, 1980; Gass & Selinker, 2008; Leow, forthcoming; Tomlin & Villa, 1994). These theories and definitions diverge, however, both in their classification of intake as a product (Chaudron, 1985; Faerch & Kasper, 1980; VanPatten, 2004) versus a process (Gass, 1998; Slobin, 1985) and in their views of intake as a unilateral versus a multi-level construct (cf. Leow, forthcoming). While Schmidt’s (1993) and VanPatten’s (2004) intake is more of a final product—more specifically, in Schmidt’s case, a product of the process of the conscious noticing of input—Gass’ (1988) definition of intake describes the process of assimilating linguistic material. Furthermore, while some researchers (Corder, 1967; Leow, 1993; VanPatten, 2004) describe intake as a unilateral phenomenon of only one level or type, others (Chaudron, 1985; Faerch & Kasper, 1980; Gass, 1988; Gass & Selinker, 2008; Leow, forthcoming; Slobin, 1985) view it as having types or stages. Chaudron assigned three stages of information processing to intake, Faerch and Kasper and Gass distinguished between two types of intake, Slobin proposed two processes, and Leow distinguishes between intake as a product with three levels (noticed, detected, and attended) and intake processing as a process. The current study tested the input processing stage of Leow’s model by investigating intake as a multi-leveled product and considered intake to be a product that is “that part of the input that has been attended to by the second language learners while processing the input. Intake represents stored linguistic data which may be used for immediate recognition and does not imply language acquisition” (Leow, 1993, 334).

Crucially, results from empirical studies also appear to reflect different levels of intake. These levels are characterized by several features, such as depth of processing (i.e., elaboration and high degree of consciousness (Craik, 2002)), cognitive effort (i.e., the amount of attention
that a learner pays to L2 input (e.g. McLaughlin, 1987), and awareness (i.e., "A particular state of mind in which an individual has undergone a specific subjective experience of some cognitive content or external stimulus" (Tomlin & Villa, 1994, 193)). A high level of intake, or noticed intake, most resembles the product of Schmidt’s (1993) construct of noticing that purports that attention controls access to awareness and is responsible for the noticing (attention plus a low level of awareness) that is necessary for intake to take place. Studies such as Leow (1997, 1998b), Rosa and Leow (2004), and Rosa and O’Neill (1999) serve as support for noticed intake for grammatical items, and other studies such as Martínez-Fernández (2008), Godfroid, Housen, and Boers (2010), and Godfroid, Boers, and Housen (2013) present findings that support the concept of noticed intake for lexical items. Each of these studies includes a recognition test to measure intake and a low level of awareness that includes attention and some depth of processing. Furthermore, all the studies collected concurrent data. Noticed intake, therefore, is accompanied by a low level of awareness on the part of the learner and a somewhat higher depth of processing and cognitive effort in relation to attended and detected intake. Overall, however, depth of processing is still thought to be low at the noticed intake level.

A medium level of intake, or detected intake, mirrors Tomlin and Villa’s (1994) concept of detection and is accompanied by focal attention, unawareness, low to nonexistent levels of cognitive effort, and low to nonexistent depth of processing compared to those of noticed intake. Critically, detected intake differs from noticed intake in that detected intake occurs in the absence of awareness while noticed intake includes awareness. Empirical evidence for detected intake appears to come from several implicit learning studies (Chan & Leung, 2012; Chen, Guo, Tang, Zhu, Yang, & Dienes, 2011; Leung & Williams, 2011, 2012; Williams, 2004, 2005)

A low level of intake, or attended intake premised on peripheral attention paid, is
exemplified in Godfroid, Housen, and Boers (2010), who found that some participants recognized lexical items on a post-test even though there was no increase in viewing behavior for that target word during reading. This suggests the possibility of a critical boundary in terms of viewing activity. Furthermore, Godfroid et al. (2010) defined “noticing” as being characterized by “feeble memory traces” and “(very) partial word knowledge” (p. 176); this actually seems to be addressing a finer-grained version of intake that is in line with Chaudron’s (1985) notion of the initial stages of perception of input. This ‘attended intake’ appears to be a very low-level intake characterized by peripheral or extremely low selective attention; however, whether this attended lexical item receiving the fleeting attention was processed further, with or without awareness, was not addressed in this study. In addition, attended intake is characterized by no cognitive effort, no depth of processing, and no awareness. It differs from detected intake in that attended intake has absolutely no cognitive effort or depth of processing, while detected intake may have very low levels of cognitive effort and/or depth of processing.

Based on the findings from these aforementioned studies, in addition to theoretical support (Chaudron, 1985; Faerch & Kasper, 1980; Slobin, 1985) for different stages of intake, Leow’s Model of the L2 learning process (forthcoming) should be empirically tested regarding this possible distinction between these levels of intake. In order to operationalize these levels of intake, more sensitive measurements need to be employed to obtain a more fine-grained description of the processes occurring, especially at the lower end of the intake spectrum. Eye-tracking technology appears to be a promising way to operationalize the three levels of intake.

Prior research (Carpenter & Just, 1976; Rayner, 1998) has confirmed that eye movements are directly related to underlying cognitive processes. For example, Carpenter and Just’s theoretical account of the eye fixation sequence and duration during cognitive tasks reported that
the eyes commonly fixate on the external referent as its corresponding internal representation is processed. Eye-tracking is a useful measure of the allocation of attention that fits in well with the selective attention feature of current computational models (Rehder & Hoffman, 2005a). Furthermore, eye-tracking data from readers can conceivably show the online ease or difficulty of processing in addition to highlighting “the processes that underlie incremental integration of words and phrases in developing sentence representation” (Witzel, Witzel, & Forster, 2012, p. 3). A definite advantage of eye-tracking is that it does little to interfere with natural reading; it even allows participants to employ many of the reading strategies that they typically use. There is little consensus regarding the most appropriate eye-tracking measure, however previous research shows the merit of examining many measures in one study so as to reveal a more thorough examination of the reading process (Rayner, Sereno, et al., 1989; Schmauder, 1992). Previous empirical studies in the field of cognitive science (Rehder & Hoffman, 2005a, 2005b; Theeuwes & Van der Stigchel, 2009) have successfully used eye-tracking to measure attention, however, Godfroid et al. (2010) was the first to employ eye-tracking in SLA to measure what they called ‘noticing’ of lexical items in L2 learners. Pilot results showed that participants sometimes recognized words on a multiple-choice recognition post-test even though the eye-tracking data showed no marked increase in viewing for that target word during the reading; the authors suggested the possibility of a critical boundary in terms of viewing activity, namely, input below this boundary would be ‘perceived’ and input above it would be ‘noticed.’ Also, Godfroid et al. (2013) reported a significant positive relationship between fixation duration on a novel word and post-test recognition of that word and Godfroid and Schmidtke (2013) found that total fixation time on a target and recollection predicted word recognition.

Furthermore, both theoretical accounts and recent studies hint at the existence of the
concept of depth of processing, a variable posited in Leow’s Model to play an important role in subsequent processing of learner intake. Craik and Lockhart (1972) first proposed this concept in the field of cognitive psychology, describing it in terms of deep processing versus shallow processing. More recently, Gardiner and Richardson-Klavehn (2000) related depth of processing to other concepts such as amount of attention, degree of elaboration, and level of awareness. In order to fill in the gap of a working definition of depth of processing, the current dissertation defines depth of processing as amount of attention, cognitive effort, or time spent processing or elaborating on the target item in the input, and/or noticing induced by different types of experimental tasks or conditions (cf. e.g., Robinson, 2003; Hsieh, Moreno, & Leow, forthcoming; Hulstijn, 2001; Laufer & Hulstijn, 2001; Leow, Hsieh, & Moreno, 2008). The handful of researchers that have empirically addressed the concept of depth of processing within an attentional framework have found that depth of processing is facilitative in L2 comprehension (Morgan-Short et al., 2012), intake (Hsieh, Moreno, & Leow, forthcoming; Shook, 1994), oral and written production (Hsieh, Moreno, & Leow, forthcoming), and language-related events during a writing task (Qi & Lapkin, 2001). Furthermore, Leow, Hsieh, and Moreno (2008) reported that participants who reported having processed deeper the ten instances of each of four target forms (a lexical item, a feminine article, a masculine article, and a verbal morpheme) did not have lower comprehension scores than those who had processed more superficially. On the contrary, Gass, Svetics, and Lemelin (2003) reported that except in the area of syntax, learning occurred regardless of “focused attention.” Although the preliminary evidence overall suggests that depth of processing has a facilitative effect on several aspects of the second language in the early stages, several methodological limitations need to be addressed in future research. More specifically, Leow et al. (2008) suggested that this strand of research needs to employ both
assessment tasks and target items that are more likely to facilitate deeper processing of form at
the level of form-meaning connections. They also proposed that future studies should include
recognition as an independent variable in studies of processing so that another stage of the
acquisitional process may be addressed. The relationship between depth of processing and
comprehension also needs to be further investigated, especially given the results of Leow et al.
(2008) and Morgan-Short et al. (2012). Furthermore, no previously published study has
employed separate criteria for depth of processing of grammatical forms versus lexical items.

The few studies addressing depth of processing have either operationalized it by means of
experimental condition (Craik & Tulving, 1975; Gass et al., 2003; Shook, 1994) or through data
from concurrent verbal reports (Hsieh, Moreno, & Leow, forthcoming; Leow et al., 2008;
Morgan-Short et al., 2012; Qi & Lapkin, 2001). Given that participant verbalizations from
concurrent verbal reports are able to show the learner’s cognitive processes while he or she is
interacting with the L2 (Bowles, 2010) and that the verbalizations have then been used to
evidence different levels of several cognitive concepts such as levels of awareness (Leow, 1997a,
2001a, 2001b; Rosa, 1999; Rosa & Leow, 2004a, 2004b; Rosa & O’Neill, 1999; Sachs & Suh,
2007), levels of processing (e.g., Hsieh, Moreno, & Leow, forthcoming; Leow et al., 2008;
Morgan-Short et al., 2012), it is logical that future studies on depth of processing should include
concurrent verbal reports as part of the methodology. Any discussion of think-aloud protocols
needs to consider the possibility of reactivity. Bowles’ (2010) meta-analysis showed that
overall, thinking aloud while performing a verbal task has a small effect on post-task
performance: positive reactivity was found on receptive learning, negative reactivity was
documented from written production, and there appears to be a small, positive effect for
comprehension tasks. Furthermore, several previous studies have shown no reactivity (e.g.,
Bowles & Leow, 2005; Leow & Morgan-Short, 2004; Sachs & Suh, 2007; Sanz et al., 2009, Experiment 1; Stafford et al., 2012). However, since other studies have shown reactivity (Goo, 2010; Morgan-Short et al., 2012; Sachs & Polio, 2007; Sanz et al., 2009, Experiment 2), this study employed a silent condition in order to address the issue of reactivity.

The influence of type of linguistic item on different aspects of L2 learning has been considered in both corrective feedback studies (e.g., Ellis, 2007; Mackey, Gass, & McDonough, 2000) and studies in the attentional/processing/comprehension strands (e.g., Leow, 1993, 1995; Martínez-Fernández, 2008; Shook, 1994; Uggen 2012); however, only a handful have contrasted a grammatical form with lexical items. VanPatten (1990) and Greenslade, Bouden, and Sanz (1999) compared grammatical forms and a lexical item, and both found that conscious attention to a grammatical form negatively affected comprehension, but that attention to the lexical item did not have this same effect. On the other hand, Leow et al. (2008) and Morgan-Short et al. (2012) reported conflicting findings. However, no study yet has compared a grammatical form with lexical items in a recognition or written production task.

In summary, as evidenced by the review of the related literature, there are several areas worthy of further investigation within the attentional strand of SLA. More specifically, the tenets of Leow’s Model of the L2 learning process (forthcoming) postulated to occur during the early stages (the input-to-intake and intake processing stages) of L2 learning process, namely, levels of intake and the role of depth of processing, need to be empirically tested. To measure online processing, the need to employ the most current technologies is also of importance. In addition, how type of linguistic item affects learner intake and how recognition might be different for a grammatical form versus lexical items still need to be investigated together with the issue of reactivity. In order to address these issues, the current study operationalizes level of
intake via eye-tracking measures (specifically, first fixation duration, gaze duration, second pass time, and total fixation time); depth of processing is operationalized through think-aloud protocols and a silent group is included in the research design to address the issue of reactivity.

The current thesis is divided into several chapters. This introductory chapter provides a statement of the problem and definitions of terms. The second chapter offers a comprehensive review of relevant literature, including both theoretical and empirical accounts. It closes with the rationale for the present investigation and the research questions posited. The third chapter explains the pilot study and the current study, including details regarding research design, information on the participants, materials, procedure, and methodology. The fourth chapter is dedicated to an in-depth description of the results obtained from the statistical analyses of data obtained on a pilot test. The fifth chapter includes a discussion of the qualitative findings that takes into account the research questions, theories, and empirical evidence, and then closes with limitations and an overall conclusion.
Definitions of Terms

Alertness: An overall readiness to deal with incoming stimuli (Tomlin & Villa, 1994).

Apperception: The learner’s recognition of the mismatch between what the learner knows about the L2 and what speakers of that language produce, which is required for availability for eventual integration into the learner’s developing system (Gass, 1997).

Attention: The concept of attention in SLA has three different uses and it can be used to describe (a) “the process involved in “selecting” the information to be processed and stored in memory”; (b) “the capacity for processing information”; (c) the mental “effort” involved in processing information” (Robinson, 1995b, pp. 287-288). The present study defines attention as a selection mechanism that allows the learner to store certain information in working memory that may allow further processing.

Awareness: "A particular state of mind in which an individual has undergone a specific subjective experience of some cognitive content or external stimulus" (Tomlin & Villa, 1994, 193).

Cognitive effort: The amount of attention that a learner pays to L2 input (e.g., Kahneman & Treisman, 1984; Lachman et al., 1979; McLaughlin, 1987).

Consciousness: An online awareness of subjective experiences (Schmidt, 1990).

Depth of processing: Amount of attention, cognitive effort, or time spent processing or elaborating on the target item in the input, and/or noticing induced by different types of experimental tasks or conditions (cf. e.g., Robinson, 2003; Hulstijn, 2001; Laufer & Hulstijn, 2001; Leow, Hsieh, & Moreno, 2008).

Detection: The cognitive registration of some stimuli that can interfere with the processing of other information. Detection alone is the key for further processing and learning to take place
Eye-tracking: The technique of measuring either the point of gaze of the eye or the motion of the eye in relation to the head used in psychology-based fields to infer cognitive processing.

First fixation duration: The duration of the first eye fixation in an interest area when that area is first encountered during forward reading (in English and Spanish, this is left-to-right reading). Rayner and Pollatsek (1987) proposed that very fast cognitive operation will affect first fixation time.

First pass reading time: Sum of the fixation durations in a region before leaving that region in any direction. Synonymous with 'gaze duration.' (Witzel et al., 2011)

Fixation: The time span during which the eyes are relatively still.

Fixation number: Total number of eye fixations on each target, including the first fixation on the target and any regressions back to the target (Smith & Renaud, 2013).

Gaze duration: Sum of the fixation durations in a region before leaving that region in any direction. Synonymous with 'First pass reading time' (Witzel et al., 2011).

Intake: “That part of the input that has been attended to by the second language learners while processing the input. Intake represents stored linguistic data which may be used for immediate recognition and does not imply language acquisition” (Leow, 1993, 334). The current study investigates intake as a product.

Invisible boundary paradigm: A technique used in eye-tracking that involves manipulation of the characteristics of an upcoming word. An invisible interest area is created around a target word and once the reader crosses the boundary into the interest area, then the word either becomes visible or is masked for the duration of the fixation on the interest area (Rayner, 1975).

Level of intake: Refers to noticed, detected, or attended intake, as operationalized in this study.
and based on Leow’s (forthcoming) types of phases of intake. Compare with “phase/type of intake.”

**Microsaccades:** Most likely due to less-than-perfect control of the oculomotor system, the eyes occasionally make small, slow movements; microsaccades quickly bring the eyes back to where they were. The fixations resulting from these microsaccades are most commonly lumped together as one single fixation if they are on adjacent characters (Rayner, 1998).

**Noticing:** Schmidt (1990) proposes that attention controls access to awareness and is responsible for noticing, which he defines as attention plus a low level of awareness.

**Noticing hypothesis:** “The necessary and sufficient condition for the conversion of input into intake” (Schmidt, 1993, p. 209).

**Number of visits to the target area:** The number of times readers’ line of sight enters the area of interest (Winke, 2013).

**Nystagmus:** The small, constant tremor of the eyes that occurs even during fixations.

**Orientation:** Attentional resources being specifically directed to some sensory information (Tomlin & Villa, 1994).

**Overall dwell time:** The sum of all fixations on a word, including first run dwell time and all regressions (Smith, 2012).

**Perceptual span:** The size of the perceptual span is the region of effective vision.

**Phase/Type of intake:** Refers to Leow’s (forthcoming) postulated types or phases of attended intake, detected intake, and noticed intake, upon which the current study’s concept of “levels of intake” is based.

**Re-reading time:** The amount of time obtained from subtracting first-pass reading of the area of interest from the total time spent fixating on the area (Winke, 2013).
**Right-bounded reading time:** The sum of the fixation durations in a region before moving out of that region to the right (Witzel et al., 2011). Does not include regressive fixations outside of the region.

**Saccade:** The rapid movement the eyes make when moving from one location to the next, for example, when searching for an object or reading. It is during a saccade that the reader is sensitive to visual input. Saccadic suppression refers to the fact that we do not perceive information when the eye moves (Matin, 1974).

**Saccade latency:** The time it takes to program a saccade.

**Salience:** Perceptual salience is “how easy it is to hear or perceive a given structure” (Goldschneider & DeKeyser, 2001, p. 22); Slobin (1971) and Brown (1973) stated that perceptual salience can depend on phonetic substance, stress, pitch, word position, and utterance position.

**Second pass time:** The sum of fixations made the second time that the eyes re-enter an interest area after they had left. Second pass time typically represents rereading episodes that show the reanalysis of the interest area on the part of the reader.

**Think-aloud protocols (“think alouds”, “TAs”, "concurrent verbal reports"):** Audio recordings performed in an experimental setting in which a participant thinks their thoughts aloud while completing a task. Used to measure processing constructs such as awareness, depth of processing, and cognitive effort, among others. Protocols may be metacognitive, that is, provide an explanation for what was just thought aloud, or non-metacognitive.

**Total fixation time:** The sum of all fixations over all passes. Total fixation time is a significant predictor of the probability of posttest recognition (Godfroid et al., 2013).
**Total reading time:** The sum of all the fixation durations in a region (Witzel et al., 2011).

Synonymous with total fixation time.
Chapter Two: Review of the Literature

Frameworks in Non-SLA Fields for the Roles of Attention and Awareness in Learning

Attention.

Cognitive psychology views of attention throughout the years have maintained that attention is a limited resource (e.g., Broadbent, 1958; Deutsch & Deutsch, 1963; Kahneman, 1973; Wickens, 1989). Furthermore, most well-supported positions maintain that attention to stimuli, whether linguistic or non-linguistic, is essential for long-term storage and that serial learning cannot occur without attention (e.g., Carr & Curran, 1994; Dulany, Carlson & Dewey, 1984, 1985; Nissen & Bullemer, 1987; Reber, 1967). Regardless, other theories and limited empirical evidence exist that suggest otherwise (Cohen, Ivry, & Keele, 1990; Curran & Keele, 1993).

Classical psychology held that humans are limited capacity processors (e.g., Broadbent, 1958; Kahneman, 1973) and that there is a limited supply of attention. This was taken to mean that attention must be selective. There was disagreement, however, regarding the timing of selectiveness. Earlier models supported filter theories of attention in which only some stimuli from the total set of available stimuli are processed (Broadbent, 1958). Broadbent’s (1958) model posits a selective attention filter under voluntary control. This model considers attention to be an executive process that directs the serial passage of information between separate short-term and long-term memory stores. On the other hand, later research, such as the ‘late selection’ model (Deutsch & Deutsch, 1963), proposed that all stimuli are processed in parallel; the decision of what gets attended to occurs later. While both early and later models agree that the human processor is a passive recipient of information, more recent models assign the human processor a more active role due to the important role they give to effort in the allocation of
attention (Kahneman, 1973). According to Wickens (1984), there is not just one but multiple pools of attentional resources; task difficulty depends on whether these attentional resources are drawn from the same pool or from different pools. The case of resources being drawn from the same pool results in great difficulty, or interference (Wickens, 1984). Learners need to exert more cognitive effort and thus attention may be focused on only select parts of the input. If resources are drawn from different pools, the task is likely to be simpler because less cognitive effort is required and attention may be available to attend to other items in the input simultaneously. If attention is considered to be processing capacity, the amount of attention deployed is dependent on not only the difficulty of the task but also the amount of controlled or automatic processing involved (Lachman, Lachman, & Butterfield, 1979).

**Measuring Attention.**

As mentioned, many attempts at investigating attention in the field of cognitive science and cognitive psychology have focused on selective attention. One way of measuring attention within the cognitive science field is with dichotic listening tasks, typically used in studies of selective attention (e.g., Moray, 1959; Norman, 1969). Dichotic listening tasks require participants to attend to one stimulus while ignoring another; this is usually done by having participants wear headphones that present a different aural stimulus in each ear. The accompanying task is then designed to heighten attention to the stimulus in one ear while drawing attention away from the stimulus in the other ear. An assessment is made after the exposure to determine how much of the unattended stimulus was detected by the participant.

A main disadvantage of dichotic listening tasks that have been used to measure selective attention is that studies using this methodology cannot rule out the possibility that participants may have paid some attention to the suppressed stimulus. Schmidt (1995) points out that even if
instructions tell participants to ignore one stimulus and shadow the other, participants may momentarily switch attention.

Divided attention tasks, such as Nissen and Bullemer’s (1987) task of divided attention, are a more recent technique used to measure attention. This type of task has participants perform two tasks simultaneously, one of which requires attention and thus makes it very difficult to attend to the second task. This is done in order to ensure that attention is not oriented to the suppressed stimulus.

Another approach to measuring attention is Posner’s cueing paradigm. It assumes that attention is moved to and then engaged at a cued location; attention has to be disengaged from the cued location if a target appears at an uncued location.

Empirical Studies.

Attention has been widely-studied outside the field of SLA; experimental design has moved from dichotic tasks to divided attention tasks and cued attention in recent years. The following is a review of some of the most-cited attentional studies outside the field of SLA together with a critical summary.

Eich (1984) carried out an experiment designed to explore the possibility of dissociation between memory and what he called “awareness of memory” for unattended events. Basing his method on Jacoby and Witherspoon (1982), he tested 16 college students fluent in English on their ability to discriminate between old and new homophones in a dichotic, selective attention task. Participants were presented with old homophones and semantically-similar words in one ear, while in the other ear they were presented with an essay that they had to orally shadow. They were told they would be tested for comprehension and retention of the essay. A recognition test included old and new homophones and there was also a spelling test. Results
showed that the prior unattended presentation of homophones had a reliable influence on its subsequent interpretation. Furthermore, the spelling results suggest that unattended information may undergo some deep, semantic analysis. Eich concludes that these results may signal a dissociation between memory and awareness of memory for unattended events.

Nissen and Bullemer (1987) carried out a series of four experiments, three of which are presented here, that is considered to be strong evidence for the important role of attention in learning. In the first experiment, participants had to press a button corresponding to the position of a light on the screen as soon as it appeared; the order of the lights was either random or a set sequence, depending on the condition. Data on serial reaction time showed that the repeating group achieved significant improvements (decreases in reaction time), thus suggesting that they learned the sequence of lights and were able to predict the location of the next light. In the second experiment, participants were tested on their ability to predict the location of the next light after a certain number of trials. Conditions included a repeating-sequence group, a repeating-sequence and dual-task group, and a random-sequence dual-task group. Results showed no evidence of learning in the repeating dual-task group as neither response time nor prediction indicated learning in the repeating dual-task group. It was concluded that dividing attention had a significant negative effect on performance, thereby suggesting indirectly that attention may be necessary for learning. In the third experiment, performance of the repeated dual-task group on a subsequent repeating single-task was compared with that of a group with no previous experience. Since the results showed that the more experienced group did not score higher than the group with no previous experience, the authors concluded that the experience group had not learned anything from exposure during the previous experiments. Overall, this series of experiments supports the idea that no serial learning takes place without attention.
Basing their methodology on Nissen and Bullemer (1987), Cohen, Ivry, and Keele (1990) investigated the role of attention and sequence structure in learning a structured sequence of actions. They employed dual versus single-task and random versus repeating-sequence paradigms to determine whether the nature of the sequence affected learning in a dual task. Results showed that structured-sequence learning was possible in dual tasks as long as structures were linear. On the contrary, hierarchal structures, like the ones used in current syntactic theory and thus most relevant to language learning, could not be learned when attention was divided. Overall, Cohen et al. concluded that non-attentional learning is possible.

Curran and Keele (1993) used experimental paradigms developed by Nissen and Bullemer (1987) in an attempt to determine whether variation in attentional availability changes the pattern of results. One of the experiments they conducted included an intentional group who received an explicit description of the light sequence and practiced under single-task conditions before moving on to dual-task conditions and also an incidental group with no explicit description and only dual-task conditions. Learning was assessed under distraction. Although the incidental group was outperformed by the intentional group on the single task, both groups showed the same, significant amount of learning under dual-task conditions. This suggests that learning without attention is possible. Furthermore, due to the apparent inability of participants to use their own declarative knowledge in non-attentional (dual-task) contexts, it seems that type of knowledge acquired under non-attentional conditions is different from the knowledge acquired under attentional conditionals. Overall, these results seem to confirm Cohen et al.’s findings that non-attention learning is possible.

Based on an experimental paradigm developed by Reber (1967), Whittlesea and Dorken (1993) assessed learning of miniature artificial grammars. In one experiment, there were three
experimental conditions. While participants in the incidental repetition condition performed a distraction task, participants in the incidental analysis condition had to perform an analysis of a letter from letter strings presented to them. The memorize condition had participants memorize letter strings. Participants did not know that they would be tested on the data. Acquisition in all conditions was measured via a grammaticality judgment of strings. Results revealed that the incidental repetition condition showed no sensitivity to the target string. In the incidental analysis condition, only limited deep-structural sensitivity can occur incidentally in some, not all, novel stimuli. Results suggest that attention is necessary for learning even linear structures, thus contradicting findings by Cohen et al. (1990) and Curran and Keele (1993). These contradictions may be due to differences in the stimulus materials: while Cohen et al. (1990) and Curran and Keele (1993) used lights, Whittlesea and Dorken (1993) used letter strings.

In summary, the reviewed studies do not seem to provide a clear explanation for the role of attention in learning in non-SLA fields. Results from Nissen and Bullemer (1987) support the idea that no serial learning can occur without attention. On the contrary, results from Cohen et al. (1990) and Curran and Keele (1993) suggest that non-attentional learning is possible. Eich (1984) hints at the possibility of a dissociation between memory and awareness of memory for unattended events. Finally, Whittlesea and Dorken (1993) conclude that attention is necessary for learning even linear structures.

The lack of an unequivocal answer regarding the necessity of attention for learning may partially be due to the methodological limitations seemingly inherent in studying attention. As Schmidt (1995) points out, it is currently impossible to show a complete lack of attention: studies can only create experimental conditions of *less* attention rather than *more* attention. Because of this, the very concept of learning without attention is problematic on a methodological and
theoretical level (Al-Hejin, 2005). If the argument is maintained that learning necessarily requires detection, and detection is part of attention, then learning in the absence of attention is theoretically impossible (Truscott, 1998).

The conclusions from these studies should be interpreted with caution because even Curran and Keele (1993) admit that “When we refer to one form of learning as nonattentional, we do not wish to imply that no attention whatsoever is used on the primary task. Undoubtedly, subjects must in some sense attend to a visual stimulus to make a response” (p. 190). Instead of making a case for learning without attention, the results from the studies reviewed should be interpreted as hinting at the possibility of very low levels of attention that, when activated, can help learners show improved performance on simple tasks that involve, for example, sequences of lights. In other words, results are not easily transferable to more complex types of learning, such as learning a second language.

Theoretical Underpinnings in the SLA Field for the Roles of Attention and Awareness in Learning

SLA is driven by what learners pay attention to and notice (Schmidt, 2001). Attentional devices are what “tune in” some information and “tune out” other information; this is one similarity that language processing shares with other types of processing (Gass et al., 2003). Although the fields of cognitive science, cognitive psychology, and SLA generally agree that attention is crucial for L2 learning to occur (e.g., Robinson, 1995a; Schmidt, 2001; Shiffrin & Schneider, 1977; Tomlin & Villa, 1994), there is less consensus regarding the role of awareness. While some empirical studies support a dissociation between learning and awareness (e.g., Cohen et al., 1990; Reber, 1967), others have argued that the only learning able to occur without awareness is non-significant in terms of the task of learning an L2 (e.g., Curran & Keele, 1993;
It is important to note that the terms “awareness” and “consciousness” have been used interchangeably in the SLA literature. Theoretical approaches to consciousness have related it to attention and awareness, among other constructs (Schmidt, 1990). Consciousness is an online awareness of subjective experiences (Schmidt, 1990) and relies on the postulate of humans as limited capacity processors of information who are limited in what they can attend to at one point in time (McLaughlin et al., 1983, Schmidt, 1990).

There are six major published SLA theoretical frameworks postulated to account for the early stages of the learning processes in SLA: McLaughlin (1987), Gass (1988), Schmidt (1990), Tomlin and Villa (1994), Robinson (1995), VanPatten (1996). Furthermore, Truscott and Sharwood Smith (2011) recently proposed the interdisciplinary MOGUL framework, which can be applied to the early stages of L2 learning. They all address the early stage of learning and each posits a somewhat different role for attention in the L2 learning process. While Gass, and VanPatten provide a fuller theoretical account of all stages of learning, McLaughlin, Schmidt, Tomlin and Villa, and Robinson provide partial accounts.

**McLaughlin (1987).**

Psycholinguistic perspectives of SLA, largely based on findings from cognitive psychology, maintain that the adult learner is a limited capacity processor (McLaughlin, 1987). In other words, learners are limited in what they can attend to at one point in time. Additionally, there are limits on what they can process on the basis of expectations and previous knowledge (McLaughlin, 1987; McLaughlin, Rossman, & McLeod, 1983). Cognitive effort dictates the amount of attention that a learner pays to L2 input (e.g., Kahneman & Treisman, 1984; Lachman et al., 1979; McLaughlin, 1987, McLaughlin et al., 1983). As long as attention is viewed as
processing capacity, it is flexible and can be adjusted according to the difficulty of the task (Kahneman, 1973) and the degree of controlled versus automatic processing (Lachman et al., 1979).

The postulate of humans as limited capacity processors of information (McLaughlin et al., 1983; Schmidt, 1990) is associated with the notion of consciousness in information processing theories. First, consciousness as a limited capacity memory system appears in several models that attempt to define input processing in terms of a group of storage structures (e.g., Kihlstrom, 1984). More specifically, Kihlstrom suggests that consciousness, focal awareness and short term store are essentially the same constructs. Furthermore, multi-store models of memory propose that permanent storage requires processing in short term memory. Since information in short term memory is lost if not further encoded, if consciousness is the same as the subjective experience of short term store, then storage without conscious awareness is impossible.

Within the attentional strand of SLA, a central question is how L2 learners allocate attention to L2 input. Central to this issue is McLaughlin’s (1987) description of the adult L2 learner as a limited capacity processor. As previously explained, this psycholinguistic notion hypothesizes that learners are limited in what they can attend to during input processing and attention is in competition to be allocated to certain parts of the input (McLaughlin, 1987; McLaughlin, Rossman, & McLeod, 1983).

The distinction between controlled and automatic processes plays a key role in Cognitive theory. Controlled processes are more characteristic of the learner’s information processing mechanism during the early stages of SLA. Requiring great amounts of cognitive effort, these controlled processes are tightly-limited by capacity and are voluntarily controlled by the learner.
They are generally used with new and/or inconsistent information that the learner encounters. While controlled processes may be conscious, automatic processes presumably do not require consciousness. Automated processes also require minimal cognitive effort and occur rapidly. Because of this, the amount of attention that the learner pays to the L2 data depends greatly on the amount of cognitive effort required by the processing of the input. Because Cognitive theory views conscious attention as processing capacity, the difficulty of the task and the amount of controlled or automatic processing involved influences the amount of attention deployed.

McLaughlin (1987) makes several important postulations regarding the input-to-intake phenomenon. First, learners do not attend to all possible incoming input. However, the input that is attended to is selected by the learner. Second, processing new and/or inconsistent input requires more cognitive effort; activating old information requires only minimal attention, minimal cognitive effort, and potentially no awareness. Furthermore, task difficulty and the amount of controlled or automatic processing involved may require differing amounts of cognitive effort or depth or levels of processing.

Gass (1988 and elsewhere).

Gass (1988) proposes a framework for second language studies that provides a global and comprehensive perspective showing how various components interrelate. The learner’s conversion of input to output involves five levels: 1) apperceived input, 2) comprehended input, 3) intake, 4) integration, and 5) output.

The first step of this conversion process is apperceived input, or the process of understanding how observed characteristics relate to prior knowledge (Gass, 1988; Gass & Selinker, 2008). In other words, this is an internal cognitive act in which noticed material is related to past experiences. Apperceived input is a small piece of linguistic datum that is noticed
by the learner because of a particular feature. It is also a priming device in that it prepares the learner for the possibility of upcoming analysis. Gass and Selinker (2008) explain that several factors serve as input filters. First, both extremes of frequency can serve as filters: very frequent features of input are likely to be noticed, as are very infrequent features when dealing with learners of advanced proficiency. Second, affect influences apperception. For a learner who feels socially or psychologically distant from the L2 community, input may not be available. Prior knowledge also plays a role in determining whether data is apperceived. Helping to decide whether input is meaningful, prior knowledge needs to be integrated with new knowledge to result in learning. Finally, attention can also be an input filter: this is what allows learners to notice the mismatch between native-like input and their own knowledge of the L2.

In Gass’ (1997) view, attention is what allows learners to notice the gap between what they know about the second language and what speakers of that language produce; noticing this gap is required for availability of target-like material for eventual integration into the learner’s developing system. According to Gass, the concept of apperception involves the learner’s recognition of this mismatch. Determined to a large extent by selective attention, apperception also involves the process of understanding that relates observed qualities of an object to past experiences. As Gass and Selinker state, “without selective attention, grammar development does not take place” (2008, 307). Gass’ (1997) concept of apperception is closest to what Tomlin and Villa (1994) call “detection” and Schmidt (1990) defines as noticing.

Intake comprises the assimilation of linguistic material and describes the mental activity that mediates between input and grammars (Gass & Selinker, 2008). It is the component where psycholinguistic processing takes place. More specifically, Gass and Selinker’s intake involves information being matched up against prior knowledge; processing occurs during intake and
takes into account the existing internalized grammatical rules. This seems to constitute Gass and Selinker’s use of ‘mental activity’, along with some other major processes that take place in the intake component, such as hypothesis formation, hypothesis testing, hypothesis rejection, hypothesis modification, and hypothesis confirmation.

Gass and Selinker (2008) described the processes that the L2 learner undergoes to form, test, reject, and modify hypotheses, all of which are some of the major processes that occur in the intake component. First, hypothesis formation can occur when new information is added. The L2 learner is able to make conclusions based on this new information by attending to the form, apperceiving it (relating it to past experiences), and understanding it in terms of its meaning and structure (Gass & Selinker, 2008). Prior knowledge and L1 knowledge both contribute to these steps. Prior knowledge is necessary for apperception and comprehension; it also plays a role in the eventual intake because the analysis matches up with what the learner already knew. The learner’s L1 knowledge is also crucial in that it helps facilitate a conclusion. Hypothesis testing occurs when the L2 learner tests the new hypothesis against a reasonable assumption based on the similarity between the L1 and L2 (Gass & Selinker, 2008). If the hypothesis is disconfirmed and erroneous, it may be modified at a later point in time if the learner processes intake that contradicts the original hypothesis and a stronger hypothesis tested against an even more reasonable assumption can be made. In this case, the original hypothesis is rejected and is no longer relevant in the learner’s developing grammar.

The fourth stage of Gass’ (1997) model is integration. Integration involves the two possible outcomes of language intake: development of the learner’s grammar and/or storage. Output is the final stage in Gass’ (1997) model. Two occurrences at the output stage, hypothesis testing and syntactic rather than semantic analyses of language, can cause a feedback loop back
Schmidt’s (1990 and elsewhere) noticing hypothesis.

Schmidt’s (1990, 1993, 1994, 1995) framework of attention in L2 learning is embodied in his noticing hypothesis (1990), which proposes that attention controls access to awareness and is responsible for the noticing that Schmidt defines as attention plus awareness. Learners need to consciously notice a form in the L2 before more processing can occur. He defines noticing as “the necessary and sufficient condition for the conversion of input into intake” (Schmidt, 1993, p. 209). In his view, focal attention and awareness are the same. Although Schmidt asserts that noticing can be operationalized as the availability for a self-report either concurrently or immediately after exposure to the input, he also warns that a lack of self-report does not necessarily imply a lack of awareness.

Schmidt (1990) proposes two levels of awareness: awareness at the level of noticing and awareness at the level of understanding. This distinction accounts for the difference between intake/item learning and restructuring/system learning. Awareness at the level of understanding is related to the ability to analyze, compare, and test hypotheses about linguistic input.

Schmidt (1990) has related attention to consciousness, which he defines as an online awareness of subjective experiences. Conscious processing is partially subject to deliberate control (Schmidt, 1990). Unconscious processes, on the contrary, are attributed to specialized systems that operate in parallel. Although this type of process is not under voluntary control, it is fast, efficient, and accurate. Furthermore, it is responsible for most details of cognitive processing.

Tomlin and Villa (1994).

According to Tomlin and Villa (1994), a central question in SLA is how select parts of
the general input that the learner is exposed to are chosen as intake for acquisition. They explain that it is generally understood that the learner uses attention for one of two things: to aid in the comprehension of the meaning of an utterance (attention to meaning) or to aid in the psycholinguistic processing (attention to form). Grammar learning is enhanced when the learner attends to form.

They also point out that the course-grained notion of attention employed in SLA up to that point is based on four main concepts: “attention is a limited-capacity system, attention is the process of selecting critical information for further processing, attention represents effortful processing that can be contrasted with more automatic and less effortful processing, and attention is a matter of the control of information and actions” (Tomlin & Villa, 1994, 189).

Tomlin and Villa specifically address the role of attention in SLA through their functionally-based model of input processing based on Posner and Petersen’s (1990) description of the human attention system. According to both sets of researchers, attention subsumes three components that have neurological correlates, namely, alertness, orientation, and detection.

Alertness is an overall readiness to deal with incoming stimuli. Tomlin and Villa claim that the level of alertness of an individual can be measured and manipulated and therefore can be shown to have different effects on processing. As alertness increases, so does the rate at which information is selected for further processing, but accuracy may suffer as a result. Tomlin and Villa highlight two important aspects of alertness as being: 1) the speed of information selection is a function of alertness, and 2) rapid selection may result in a lack of sufficient information processed in order to allow for an accurate response. Alertness has been operationalized by means of presenting targets with or without a warning that comes in the form of an auditory or visual cue that a target is about to appear. While the lack of a warning creates a decrease in
alertness, the presence of a warning creates increased alertness.

Orientation, modulated by alertness, is defined as attentional resources being specifically directed to some sensory information at the exclusion of other. According to Tomlin and Villa, “the key idea of orientation is that the specific aligning of attention (“orienting”) on a stimulus has facilitative or inhibitory consequences for further processing depending on whether information occurs as expected or not as expected” (p. 191). Orienting attention generally facilitates detection; stimuli that do not receive attentional orientation are inhibited so that their detection requires more effort than normal. The concept of orientation is important in the field of SLA because attention to form or function would benefit if it were possible to determine to which the individual’s attention was oriented at the moment of acquisition. Orientation is typically operationalized by using tasks that make a person expect more than one outcome while cuing attention to be either congruent or incongruent with the expectation.

Detection is the cognitive registration of some stimuli and most closely resembles Schmidt’s (1990) definition of noticing. Notably, detection can interfere with the processing of other information. According to Tomlin and Villa, detection alone is key for further processing and learning to take place. In their view, awareness may enhance input processing, but it is not required. Awareness is what distinguishes noticing from simple detection. This is in direct conflict with Schmidt, who has postulated that conscious noticing is important for learning.


Robinson (1995) is able to reconcile the differing views of Schmidt (1990) and Tomlin and Villa (1994) when he postulates that noticing includes detection plus rehearsal in short term memory. He posits that a certain threshold in short-term memory must be reached before an activation can become a part of awareness. By placing noticing farther along the acquisitional
processing timeline than detection, Robinson concurs with Schmidt that lack of awareness precludes intake and, as a consequence, potential learning.

According to Robinson, attention describes three aspects: 1) the process of selecting which information will be processed and stored, 2) someone’s capacity for processing information, and 3) the mental effort that processing entails. The model implies the operation of a central executive attentional mechanism that allocates attention to fulfill task demands. Information processing requires conscious attention to form in the input.

Robinson proposed a model that complements Schmidt’s claim and opposes Krashen’s (1981) dual-system hypothesis: he argues that the learner’s performance on implicit and explicit learning and memory tasks is due to differences in “consciously regulated processing demands of training tasks (p. 283),” not by the activation of systems accessed both consciously and unconsciously.

*VanPatten (1996, 2004).*

VanPatten (1996, 2004) developed a model of adult second language acquisition and use that describes the three distinguishable sets of processes involved in acquisition (see Figure 1).

Figure 1: VanPatten’s (1996) SLA Framework

<table>
<thead>
<tr>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>input</td>
<td>intake</td>
<td>developing system</td>
</tr>
</tbody>
</table>

I = input processing  
II = accommodation, restructuring  
III = access, control, monitoring

Input processing (I) occurs during the conversion from input to intake, while the second set of processes (II) accounts for accommodation of intake and the restructuring of the developing linguistic system (McLaughlin, 1990). The third set includes those that promote
certain aspects of language production such as monitoring, accessing, and control (Schmidt, 1992; Terrell, 1991).

The first set of processes (I) captures input processing, the initial stage during which learners make form-meaning connections and the process by which they interpret the role of nouns in relation to verbs. It is premised on the idea that “comprehension and processing for natives cannot and is not the same process as that for beginning non-natives who must not only comprehend but also come to discover linguistic data in what they comprehend (VanPatten 2004, 21).” VanPatten’s model aims to capture not only the conditions under which a learner may make the connection between a form in the input and its meaning but also the processes they initially bring to the act of acquisition. According to VanPatten (2004), processing occurs when the learner makes a connection, be it partial or complete, between form and meaning in real time. In other words, the learner notes a form and, at the same time, determines its meaning or function. While learners may choose to pay attention to input in order to comprehend, processing “basically happens to the learner (VanPatten, 2004, 9).” VanPatten (2004) supports Schmidt’s (1990) affirmation that noticing is any conscious registration of a form without the necessity of having meaning attached to it. He furthermore maintains that the L2 learner must notice forms in the input for acquisition to occur.

VanPatten refers to two widespread claims in SLA: learners are focused on extracting meaning from the input (e.g., Faerch & Kasper, 1986; Krashen, 1982) and learners must notice items in the input for acquisition to eventually happen (Schmidt, 1990). Given these claims, it follows that learners are pushed to look for meaning in the input. VanPatten’s (2004) Primacy of Meaning Principle posits that learners process input for meaning before they process it for form. However, considering the limited resources of the L2 learner for processing input, certain
elements of form in the input will be passed over for processing.

*Truscott and Sharwood Smith (2011).*

Truscott and Sharwood Smith (2011) aim to further detail the key SLA concepts of input, intake, and consciousness by means of their interdisciplinary framework, Modular Growth and Use of Language (MOGUL). A processing-oriented framework, it is motivated by what Truscott and Sharwood Smith refer to as the fundamental problem of Schmidt’s Noticing Hypothesis (1990): the lack of a solid account of what consciousness is. This framework attempts to integrate language acquisition accounts with proposals of how language and cognition interact. It includes a modularized core language system that contains two subsystems: the phonological module and the syntactic module. The interface system is what determines which phonological structures and syntactic structures are linked up. This requires the auditory-acoustic systems for the phonological module, the articulatory systems for speech production, and the conceptual system responsible for conceptual structures (semantics and pragmatics). Perceptual output structure is the output of the modality-specific processing systems. Each module shares the same basic internal structure of an information store of structures (representations) and a computational (processor) system. The interface processors serve to match up structures in adjacent modules, such as the phonological representation of a word with its syntactic function (e.g., /hot/ with (Adj)).

The concept of consciousness within the MOGUL framework is explicitly defined. While the objects at which the consciousness is aimed are representations, these representations can only become conscious if its activation level crosses a threshold. People can only be aware of linguistic information in the form of sound (auditory structures) or written form (visual structures), however indirect awareness of extramodular linguistic knowledge is possible.
According to the MOGUL framework, language learning can be explained using the MOGUL framework via several representations. First, input is defined as the representation(s) available to the processor (learner). The activation level associated with the representation needs to cross its threshold in order to become conscious. In the case of language acquisition, input is a perceptual representation (an auditory or visual structure) of spoken or written language. Intake consists of any information in an input representation that can be used by the processor for a certain purpose at that moment. However, this information must be extracted from the perceptual output structure representation because this representation is nonlinguistic. Therefore, whether or not an input representation is actually intake depends on the module’s ability to perform this extraction, which in turn depends partially on the current state of the information store. The final output is represented by perceptual output structures, each dealing with input from one of the senses. Awareness of the input is considered to be awareness of a perceptual representation, while noticing is awareness of a follow-up perceptual representation that contains one part of the original representation. Notably, Truscott and Sharwood Smith posit that within the MOGUL framework, noticing is important in acquiring metalinguistic knowledge but should not play a direct role in language development.

**Summary.**

In summary, the models of the early stages of the learning process in SLA proposed by Gass (1998 and elsewhere), Schmidt (1990, 2001), Tomlin and Villa (1994), Robinson (1995), VanPatten (1996) are all similar in that they agree that attention facilitates L2 learning, however some views (Schmidt, 1990; Robinson, 1995) posit that awareness is necessary for L2 learning. Truscott and Sharwood Smith’s (2011) MOGUL framework deems noticing important but not necessary. However, all models do appear to concur that there is a progression from input >
intake > output. In other words, the learning process begins with exposure to the L2. Some of the input is taken in, and then some portion of that intake is available for further processing with the potential to be incorporated. The final stage of the L2 learning process is output, or the L2 production of the learner (cf. Leow, forthcoming, for a more detailed description of the stages of the learning process in SLA). These six models, however, also have several differences at the early stages, especially surrounding intake, depth of processing, and the role of awareness. These differences are discussed below.

**Intake**

As discussed above, major theoretical viewpoints of the L2 learning process (e.g., Corder, 1967; Faerch & Kasper, 1980; Gass & Selinker, 2008; Tomlin & Villa, 1994; Truscott & Sharwood Smith, 2011) postulate a preliminary stage in which a subset of the input is taken in, namely, what is referred to as intake. At the same time, although there is a general agreement that intake is an essential component of SLA, there is no consensus regarding its definition.

The term “intake” was first coined by Corder (1967). In an attempt to distinguish input from intake, he defined intake as “what goes in and not what is available to go in (p. 165).” Such early attempts to distinguish input from intake were motivated by a desire to show the inappropriateness of the learner as a passive agent in learning. On the contrary, it is the learner who controls the intake (Chaudron, 1985).

A main question concerns exactly what it is about the learner’s language acquisition mechanism that controls intake. Corder (1967) mentions two different aspects that independently influence intake: internalized language (representations of rules and structures acquired by the learner) and the mechanisms for perception and learning (procedures, processes, and other psychological variables that together comprise the learner’s cognitive apparatus).
Faerch and Kasper (1980) maintain that intake is the subset of the input that is assimilated into the interlanguage and to which the interlanguage accommodates. They also distinguish between intake for learning and intake for communication.

Chaudron (1985) defines intake as “the mediating process between the target language available to learners as input and the learners’ internalized set of L2 rules and strategies for second language development” (Chaudron 1985: 1). The learner plays an important role as an active agent in L2 learning because it is he who controls the intake. In Chaudron’s view, intake should not be considered a single event or process. He discusses intake as consisting of three stages of information processing, moving from preliminary intake to final intake. First there is a perception of the input, following by the recoding and encoding of communicated message into long-term memory. Last, there is a series of stages in which the learner integrates and incorporates the linguistic information into the developing grammar. This process-oriented description exists as a continuum from preliminary intake to final intake.

Slobin (1985) also distinguished between intake in a way similar to Chaudron (1985). He proposed two types of processes: 1) those used for converting input into stored data for use in language construction and 2) those used for organizing stored data into linguistic systems. Most empirical efforts have addressed the first type of process, which is also the one that corresponds to Leow’s (1993) definition of intake as “that part of the input that has been attended to by the second language learners while processing the input. Intake represents stored linguistic data which may be used for immediate recognition and does not imply language acquisition.” (p. 334).

Gass (1988) defines intake as “the process of assimilating linguistic material.” (p. 206). Intake is a process of mental activity, mediating between input and grammars, and leading to
grammar formation. According to Gass’ framework, integration is a result of the process of intake and has as possible outcomes storage and development of one’s L2 grammar. Intake differs from apperception and comprehension because the latter two do not necessarily lead to grammar formation. According to Gass and Selinker, psycholinguistic processing takes place during what they call the “intake component” (p. 303). In other words, it is during the process of intake that information is matched up against prior knowledge and processing takes place in light of the existing rules of the learner’s internalized grammar rules.

VanPatten (2004) explains that intake is the part of input that has already been processed in working memory and is available for further processing. Importantly, intake is not just filtered data; it also includes data that has been processed incorrectly and therefore comprises both correct and incorrect form-meaning/function connections. Crucially, VanPatten seems to conflate input processing and intake processing because he defines input processing as making form/meaning connections, which may be argued to involve further additional processing when compared to mere intake.

Truscott and Sharwood Smith (2011) define intake as an abstract notion and one that includes any information in an input representation that can be used by the processor (learner) for a certain purpose at that time. They elaborate by explaining that an input representation needs to contain linguistic information to qualify as linguistic intake. Since the MOGUL framework is interdisciplinary and does not focus only on language acquisition, linguistic information is never explicitly present in perceptual output structures. Instead, the module must be able to extract the linguistic information from the perceptual output structure, which is dependent upon the readiness of the module to deal with particular information. In summary, two requirements must be met in order for the information contained in the perceptual output
structure to constitute intake: first, the module must be able to deal with the representation and second, the module must be able to extract the information from the representation. Because of this, the preferred option of Truscott and Sharwood Smith is to abandon the concept of intake and instead concentrate on determining ways that the input representation can and cannot be used by certain modules in certain contexts.

An important difference between these definitions is that of intake as a product versus intake as a process (cf. Leow, forthcoming, for a description of the learning process as both a product and a process). While Schmidt’s (1995) and VanPatten’s (1996) intake is a product—more specifically, in Schmidt’s case, a product of the process of the conscious noticing of input—Gass’ definition of intake describes the process of assimilating linguistic material. The part of Gass’ framework that most closely resembles intake as a product is comprehended input. VanPatten and Gass also have converging opinions regarding the stage in L2 learning that precedes the taking in of specific input information. According to VanPatten, the taking in of input is the first step of L2 learning; this is what happens during input processing. Gass’ framework, on the other hand, places apperception prior to the taking in of specific input.

Intake is considered by some (Corder, 1967; Gass, 1988; Gass & Selinker, 2008; Leow, 1993; VanPatten, 2004) to be a unilateral phenomenon of only one level or type, while others (Chaudron, 1985; Faerch & Kasper, 1980; Slobin, 1985) describe it as consisting of types or stages. Among those who posit various types of stages, Chaudron proposed that intake consists of three stages of information processing, Faerch and Kasper distinguished between two types of intake, and Slobin identified two processes.

**Assessment Tasks Employed to Measure Intake**

Two types of assessments have traditionally been used to measure intake: the multiple-
choice sentence completion task and the word identification task. Many studies (e.g., Leow 1993, 1995, 1997, 1998b; Rosa & Leow, 2004; Shook, 1994) have used the multiple-choice sentence completion task to measure intake. Each item on this task consists of a sentence taken directly from the input, however either the target structure/item or the part of the sentence containing the target structure/item is missing. The target appears as one of a number of possible answers, along with other, incorrect options. The learner, therefore, needs to choose the correct completion of the sentence according to how they heard/read it in the input. This task is commonly presented as both a pretest and posttest.

The word identification task is another way to measure intake and is a modified technique from L1 reading research to determine whether vocabulary is acquired indirectly through reading (see Nagy, Herman, & Anderson, 1987, and Nagy, Anderson, & Herman, 1987). It typically consists of a list of all target structures/items with many distracters mixed in. Respondents read the list and then mark which items appeared in the input to which they were just exposed. According to Lee (1998, p. 38), it is assumed that if “learners noticed, detected, or otherwise cognitively registered a form, then they would identify it in the recognition task.”

All efforts to measure intake must take into account that certain variables may influence interactions between the input and the learner’s capacity for processing it (Chaudron, 1985). Some of these possibly confounding variables include time on task, time between experimental exposure and intake assessment, production ability, and grammar knowledge. Controlling for time on task is crucial because learners who have more time available to complete a task also have more opportunities to use higher-level, less automatized knowledge sources, thus confounding the analysis of the early stages of learning (Chaudron, p. 9). Learners should not be able to return to any previously-completed item for the same reason. Therefore, a relatively
short time limit should be enforced for completion of the intake measurement. Furthermore, the time elapsed between exposure to the experimental input and the administration of the intake measure should be as short as possible to avoid memory decay. Ideally, the intake measurement should immediately follow the presentation of the input. In addition, learners should be able to report input that they took in, regardless of their production abilities. Intake measures should move away from production and focus on the more passive form of selection of the target among distractors. Last of all, intake measurements should be designed so that grammar knowledge cannot interfere with the learner’s indication of what has been taken in.

Multiple-choice sentence completion tasks have several weaknesses in relation to the aforementioned variables. First, the multiple-choice sentence recognition task, although controlling for production ability, risks conflating intake with grammar knowledge. Even if a grammar pretest is administered beforehand and the learner demonstrates no knowledge of the target grammatical structure, it is possible that some prior knowledge of the structure is activated during the intake measurement. Furthermore, a learner may be able to eliminate certain distracter options on a multiple-choice choice question because they know the grammar form of the distracter. The word identification task, however, is a strong alternative because it is not as prone to confounding intake with other variables. There is no need for production, little risk of existing grammar knowledge influencing performance, and time on task and time elapsed between treatment and assessment can be easily controlled.

The preceding information highlights the importance of excluding variables that may influence the interaction between the input and the learner’s capacity for processing the input. Given that prior grammatical knowledge can influence results of multiple-choice sentence completion tasks, researchers should rely on the word identification task as a measurement for
intake in future studies.

**Empirical studies.**

The theoretical accounts of intake vary greatly in their views of intake. There has been no consensus regarding how many stages of intake there are, nor has there been an ultimate decision regarding whether intake refers to a process or a product. Indeed, a careful review of the concept of intake appears to indicate the potential existence of levels of intake and some empirical research on intake also seems to hint at the potential existence. The intake discussed in the empirical studies to follow seems to fall into three categories: a high level of intake, a medium level of intake, and a low level of intake. Therefore, the studies will be presented in that order.

**High level of intake.**

A high level of intake, most resembles Schmidt’s (1993) construct of noticing that purports that attention controls access to awareness and is responsible for the noticing (attention plus a low level of awareness\(^1\)) that is important for intake to take place. Several empirical studies lend support to the concept of a high level of intake. These studies are similar in that the participants show a low level of awareness of the target, employ selective attention, demonstrate low cognitive effort, and only shallowly process the target. In addition, all of these studies employ concurrent measures. Since this dissertation investigates both lexical and grammatical targets, empirical support for a high level of intake will be divided accordingly: Leow (1997,

\(^1\) The definition used for awareness is the same as the one used in the Definitions of Terms, or "A particular state of mind in which an individual has undergone a specific subjective experience of some cognitive content or external stimulus" (Tomlin & Villa, 1994, 193).
Rosa and O’Neill (1999), Rosa and Leow (2004), all lend support to the concept of a high level of intake for grammatical items, while Martínez-Fernández (2008), Godfroid et al. (2010) and Godfroid et al. (2013) constitute the empirical support for lexical items.

In a study supporting a high level of intake for grammatical items and investigating the role of awareness in L2 behavior in relationship to Schmidt’s (1990, 1993, 1994, 1995) noticing hypothesis, Leow (1997) used concurrent verbal reports to first establish that noticing did occur; he then addressed the role of levels of awareness in the L2 learners’ intake and written production of a grammatical form (impersonal imperative) in the L2. The participant group included 28 adult beginning learners of Spanish who completed a problem-solving task in the form of a crossword puzzle. A multiple-choice recognition assessment task measured intake and a written production task measured production. Three levels of awareness were identified in the concurrent verbal reports, one of which corresponds to a low level of awareness, that is, noticing, ([+cognitive change, -meta-awareness, -morphological rule formation] when participants did not comment on their subjective experience or verbalize a rule). Results showed that while participants demonstrating a higher level of awareness performed significantly better than those showing a lower level of awareness (noticing,) even at this low level, positive effects were found.

Leow (1998b) also provided empirical support for a high level of intake for grammatical items when he investigated the immediate and delayed effects of Tomlin and Villa’s (1994) fine-grained analysis of attention (more specifically, attentional functions: alertness, orientation, and detection) on the intake and written production of adult beginning L2 learners as measured by performance on a problem-solving task examining a morphological form (irregular third-person singular and plural forms of stem-changing verbs in the preterit). The pool of subjects comprised
83 students enrolled in first-year Spanish at the university level. The target form was the irregular third-person singular and plural preterit forms of stem-changing –ir verbs in Spanish; the experimental task was a crossword puzzle. Detection was operationalized as the cognitive registration of a new, targeted grammatical form as revealed by the learner’s performance on a think-aloud protocol. Four versions of the crossword were used to isolate the effects of alertness, orientation, and detection. The crossword addressing alertness was not manipulated and did not require subjects to know the irregular forms in order for successful completion. The second condition addressed the effects of orientation and had instructions that told participants to note that some verbs forms were irregular, in addition to a choice of a correct and incorrect answer with the participant’s attention directed to the incorrect form. The third condition had the irregular verb form instruction in addition to clues in the crossword. The fourth condition measured the effects of detection with orientation and was the same as the third condition except for the special note about irregular verb forms in the instructions. A multiple-choice recognition assessment task measured intake and a fill-in-the-blank test measured written production ability. Results showed that learners demonstrating superior cognitive registration (groups 3 and 4) performed significantly better than those not demonstrating an equal amount of cognitive registration on both the recognition task and the written production task. As Leow pointed out, this strongly indicates that further cognitive processing can occur at higher cognitive levels only when a morphological form is detected. Additionally, TA protocols reveal instances of further processing of the detected forms only in groups 3 and 4; these further processing qualified this as a high level of intake. Leow (1998) concluded that detection is important for L2 morphological development and permits detected items to be further processed. Interestingly, due to the type of task (a crossword puzzle) employed in this study, this high level of intake is arguably more in
line with Schmidt’s notion of noticing, that is, attention accompanied by a low level of awareness, than Tomlin and Villa’s notion of detection.

In another study providing empirical support of a high level of intake for grammatical items, Rosa and O’Neill (1999) investigated implicit and explicit conditions and the level of awareness raised while processing input. Participants included 67 L1 English undergrad students enrolled in fourth-semester Spanish. A multiple-choice jigsaw puzzle comprised the experimental task and a pretest/posttest multiple-choice recognition assessment task measured their intake of the target structure, namely, contrary-to-fact sentences in the past. Concurrent verbal reports were recorded during the experimental task and instances of awareness were coded as awareness at the level of noticing, awareness at the level of understanding, and as “no verbal report.” The two levels of awareness follow Schmidt’s (e.g., 1990) noticing hypothesis. Similar to Leow’s (1997, 1998b) findings, Rosa and O’Neill reported positive effects of noticing on intake.

Likewise, Rosa and Leow (2004) also provided support for a high level of intake for grammatical items: they investigated whether different levels of awareness would influence learners’ ability to recognize contrary-to-fact statements in the past. One hundred L1 English college students enrolled in fourth-semester Spanish completed think-alouds during the experimental treatment of multiple-choice jigsaw puzzles. The think-alouds were coded for two levels of awareness following Rosa and O’Neill (1999) and based on Schmidt’s (e.g., 1990) noticing hypothesis. Awareness was also measured off-line with a post-exposure questionnaire. A multiple-choice recognition task measured intake and a controlled writing task measured production. Results corroborated findings of previous studies (Leow, 1997, 1998b; Rosa & O’Neill, 1999) that intake at the level of noticing is available for use on subsequent recognition
tasks after exposure.

Turning to the empirical support for a high level of intake for lexical items, Martínez-Fernández (2008) aimed to determine whether tasks with different degrees of involvement load induce different levels of awareness in light of the involvement load hypothesis (Laufer & Hulstijn, 2001). Participants included 45 L1 English college students enrolled in fourth-semester Spanish; they were randomly assigned to either a control group, a single gloss group, a fill-in group, or a multiple-choice group. There were eight target lexical items: four concrete and four abstract nouns. A pretest assured that participants had no prior knowledge of the lexical items. Participants performed think-alouds while reading the experimental text in which the target lexical items were glossed, provided as a fill-in, provided via multiple-choice, or not altered from the bolded lettering. These conditions differed in the degree of involvement load due to the presence or absence of ‘need,’ ‘search’, and ‘evaluation.’ Posttest assessments included a written recall protocol and a series of English-to-Spanish translation tests at the word and sentence level; there were also multiple-choice translation tests at the word level. The word form production test had participants choose the correct target word from four options when provided with the L1 translations. This task is most similar to the multiple-choice recognition test that measures intake. Verbal reports were coded for awareness by classifying as noticing (ability for verbal report) attempts to read the glosses out loud or comment about the targeted lexical items. Verbal reports in the think-aloud protocols were coded for two types of noticing, defined by the ability for verbal report. ‘Noticing of one word aspect’ included noticing of meaning only or form only while ‘noticing of two word aspects’ referred to noticing both the meaning and the form. Martínez-Fernández’ operationalizations of awareness provide support for a high level of intake. Referencing Schmidt’s (1995) distinction between low and high levels of awareness,
Martínez-Fernández likens her level of ‘noticing of meaning only’ to Schmidt’s level of noticing. This level involves cognitive registration of a form, attention to the form, and deeper processing of the word’s meaning. Critically this level includes a low level of awareness. Although this study did not specifically investigate the relationship between intake and attention or awareness, the fact that the researcher coded awareness in concurrent data to reflect a low level of awareness based on Schmidt’s (e.g., 1990) noticing hypothesis and also measured intake (via the word form production test) provides support for a high level of intake for lexical items.

Empirical support for a high level of intake for lexical items also comes from Godfroid et al. (2010), as discussed previously. Their study explored what they defined as learners’ noticing of pseudo words in written input via online eye-tracking and stimulated recalls. Eye-tracking data showed longer first run dwell times and regressions made on target words; according to Rayner (1998), this indicates that the reader has encountered difficulty in text comprehension. Godfroid et al. affirm that this also means that the reader was cognitively engaged with the target word, which they then interpret as a noticing event. This demonstrated comprehension difficulty and cognitive engagement with the target words arguably also imply a low level of awareness: if a reader is having enough comprehension difficulty to regress to a previous word and become cognitively engaged with the text, a low level of awareness seems to logically accompany this.

Additional empirical support for a high level of intake of lexical items is found in Godfroid et al. (2013). As detailed in a previous section, Godfroid et al. (2013) aimed to assess whether learners' fixation times on novel words are positively associated with subsequent recognition of those words. Because attention was operationalized as a quantitative variable reflected in eye fixation times during reading, this study also supports a high level of intake. Although Godfroid
et al. (2013) clearly stated that eye-tracking cannot be used to measure awareness, this study still contains the other two characteristics essential to a high level of intake: a concept of attention based on Schmidt’s (1990) noticing hypothesis and a multiple-choice recognition assessment task. It is important to note, however, that Godfroid et al. premised their study on Schmidt’s weak version (1994, 1995, 2001) of the noticing hypothesis. Regardless, this study can still provide albeit cautious support for a high level of intake.

The studies described above focusing on grammatical items (Leow, 1997, 1998b; Rosa & O’Neill, 1999; Rosa & Leow, 2004) and lexical items (Martínez-Fernández, 2008, Godfroid et al., 2010; Godfroid et al., 2013) all provide empirical support for a high level of intake for both grammatical and lexical items. Each one includes some form of a recognition test to measure intake and a low level of awareness that includes attention and some depth of processing. Crucially, this low level of awareness documented in these studies is based on Schmidt’s noticing hypothesis, which gives important roles to both attention and awareness.

In summary, this high level of intake most resembles Schmidt’s (1993) construct of noticing that purports that attention controls access to awareness and is responsible for the noticing (attention plus a low level of awareness) that is important for intake to take place. Therefore, L2 input needs to be noticed in order for it to be usable to the learner. Since noticing requires awareness, the learner must have a conscious apprehension and awareness of the part of input noticed.

A high level of intake, therefore, is accompanied by a low level of awareness on the part of the learner and low levels of depth of processing and cognitive effort when compared to lower levels of intake (cf. below). It is also characterized by selective attention. Due to the presence of relatively more depth of processing, more cognitive effort, and selective attention, empirical
evidence appears to indicate that this level of intake has the most potential to remain stored in working memory and made available for further processing. This further processing may result in incorporation into the L2 learner’s grammar system or lexicon (e.g., Godfroid et al., 2010; Godfroid et al., 2013; Leow, 1997, 1998b; Martínez-Fernández, 2008; Rosa & Leow, 2004; Rosa & O’Neill, 1999).

**Medium level of intake.**

Several empirical studies have claimed to evidence the role of learning without awareness in SLA (e.g., Chan & Leung, 2012; Chen, Guo, Tang, Zhu, Yang, & Dienes, 2011; Leung & Williams, 2011, 2012; Williams, 2004, 2005). Although these studies did not provide any mention of a theoretical underpinning, they all have documented what appears to be minimally a medium level of intake mirroring Tomlin and Villa’s (1994) concept of detection. It seems that this type of intake might have been accompanied by focal attention, unawareness, very low levels of cognitive effort, and very low depths of processing.

In his 2005 study (a methodologically improved version of his 2004 study), Williams conducted two additional experiments to determine whether participants from a variety of language and linguistics-related backgrounds could learn miniature noun class systems in the absence of awareness. The four determiners carried both distance values (near/far) and animacy (animate/inanimate); participants were only explicitly informed of the distance function. The training phase consisted of presentation of a novel determiner and an English noun embedded in a sentence; they listened, indicated distance, repeated the phrase, and were told to create a mental image of the situation. As a distraction from animacy, participants were told that the focus of the study was memorization of the presented phrases. A multiple-choice interpretation assessment task with either trained or untrained (new) noun phrases instructed participants to select the “more familiar, better, or more appropriate” option (Williams, 2005,
Afterward, participants were asked about the criteria they used to make their choices and classified as aware or unaware based on references made to animacy. Results showed that many unaware participants performed significantly above chance in their ability to select the correct noun phrase on the multiple-choice interpretation assessment task, thus supporting the possibility of a medium level of intake.

A series of studies employing the basic research design of Williams (2005) but addressing additional variables all reported similar results.

In their three experiments, Chen et al. (2011) summarized that people not showing awareness can learn to use the appropriate determiner based on noun animacy. Participants in Leung and Williams (2011, 2012) demonstrated learning (as operationalized by Leung & Williams) even when awareness was not present. Results from Chan and Leung (2013) revealed that unaware participants had accuracy scores significantly above chance on the pronunciation judgment task. The results from all of the studies may serve as evidence for the possibility that the participants of these studies were performing at a medium level of intake level. Each study judged a portion of its participants to be unaware (each according to its own operationalization, see below for further detail), however, there was still some cognitive registration of the target forms that allowed these unaware participants to perform above chance on subsequent assessment tasks after exposure.

However, the above studies serving as evidence for the possibility of a medium level of intake do have some limitations that need to be taken into account; the most relevant of these limitations revolve around the construct of awareness\(^2\). All studies discussed above employed questionable operationalizations of awareness, namely, offline measures, given that none used concurrent measures.

\(^2\) A controversial methodological characteristic present in all studies in the implicit learning strand is that assessments included only two options. With binary options, any participant at any time had a 50% chance of responding accurately.
Furthermore, none of these studies took into account that a lack of evidence of offline verbalized report of awareness at the level of understanding does not necessary mean that a participant was unaware; it simply means that no lower level of awareness was documented. Therefore, some of the participants classified as unaware might in fact have been aware, but the awareness measurement was not valid enough to document it.

In addition, it is important to note that there have been other studies addressing the role of awareness or lack thereof in subsequent intake that have reported conflicting findings (Faretta-Stutenberg & Morgan-Short, 2011; Hama & Leow, 2010; Leow, 2000). Furthermore, some experiments from the implicit strand did not find evidence for learning without awareness (Experiment 2 of Williams, 2004; Experiment 2 of Leung & Williams, 2012; Chen et al., 2011).

Regardless of limitations and some conflicting findings, the results from the cited studies do appear to provide empirical support for the notion of a medium level of intake.

**Low level of intake.**

While high and medium levels of intake are theoretically driven and, to a certain extent, empirically supported, there also appears to be a level of intake that may be below these two, namely, intake that is peripherally attended to, and referred to here as a low level of intake. This level of intake is characterized by a lack of awareness, extremely low depth of processing, extremely low levels of cognitive effort, and only peripheral attention. Although only at a lexical level, eye-tracking results appear to serve as evidence for a low level of intake (cf. Godfroid et al., 2010). These results showed that some participants recognized words on a posttest even though there was no increase in viewing for that target word during the reading. This suggests the possibility of a critical boundary in terms of viewing activity, albeit with reference to lexical items in the input. Furthermore, Godfroid et al. (2010) defined “noticing” as
being characterized by “feeble memory traces” and “(very) partial word knowledge” (p. 176); this actually seems to be addressing a finer-grained version of intake that is in line with Chaudron’s (1985) notion of the initial stages of perception of input.

Results from Godfroid et al. (2010) necessitate consideration of the type of knowledge demonstrated on a posttest measuring intake. The posttest in Godfroid et al. measured contextually cued retrieval of written form or, in other words, mainly receptive orthographic knowledge or recognition, depending upon one’s assumption of the origin of the retrieval, namely, from working memory or from the internal system. According to N. Ellis (1994), implicit learning of the orthographic form of a word is indeed possible. Since N. Ellis reported that some participants were able to learn the orthographic form of a word implicitly, without the eye-tracking data showing any marked increases in viewing activity for the corresponding target area, it may be that those participants were implicitly learning only the word form, not the word meaning. It is the word meaning, claims Ellis, which necessarily requires explicit processing.

The intake referred to in Godfroid et al. (2010) is characterized by at least minimal attention, and accompanying this minimal attention appears to be a lack of both cognitive effort and depth of processing. The existence of a low level of intake appears to establish that some form, in this case, a lexical one, can be fleetingly attended to, enter working memory, and is capable of allowing the learner to subsequently identify it on a post-exposure word recognition task. However, whether the word has been learned, that is, internalized in terms of word-meaning connection, remains an empirical question. In addition, whether this low level of intake is similar to attention paid to grammatical items also needs to be empirically investigated. This apparent stage of intake is most similar to Chaudron’s (1985) concept of the initial stages of perception of input. As previously mentioned, Chaudron discusses intake as consisting of three
stages of information processing that move from preliminary intake to final intake. The concept of preliminary intake, or the stage of perception of the input, aligns best with a low level of intake.

**Summary.**

A high level of intake most resembles Schmidt’s (1993) construct of noticing that purports that attention controls access to awareness and is responsible for the noticing (attention plus a low level of awareness) that is crucial for intake to take place. While Leow (1997, 1998b), Rosa and O’Neill (1999), Rosa and Leow (2004), and Smith (2012) provide support for a high level of intake of grammatical items, Martínez-Fernández (2008), Godfroid et al. (2010), and Godfroid et al. (2013) all provide support for a high level of intake of lexical items. All of these studies include a recognition test or similar intake test to measure intake, in addition to demonstration of a low level awareness. Also included is selective attention, a low depth of processing and low level of cognitive effort. Crucially, this low level of awareness documented in these studies is based on Schmidt’s (1990) noticing hypothesis, which gives important roles to both attention and awareness. Due to the presence of relative depth of processing and cognitive effort, this level of intake may have the most potential to remain stored in working memory and made available for further processing.

A medium level of intake mirrors Tomlin and Villa’s (1994) concept of detection and is accompanied by focal attention, unawareness, very low levels of cognitive effort, and very low depths of processing. Empirical evidence for a medium level of intake may come from the implicit strand of research in SLA (cf. Williams, 2004, 2005, Chen et al., 2011, Leung & Williams, 2011, 2012; Chan & Leung, 2013).

A low level of intake is most similar to Chaudron’s (1985) concept of the initial stages of
perception of input. Therefore, a low level of intake is proposed to be characterized by a lack of awareness, extremely low cognitive effort, extremely low depth of processing, and only minimal or peripheral attention. The existence of a low level of intake is substantiated by the idea that something can be fleetingly attended to, enter working memory, and then quickly disappear. Empirical support for a low level of intake may come from Godfroid et al. (2010).

**Variables postulated to be associated with intake.**

The previous discussion of intake introduced several variables that different theorists and empirical findings have suggested to be associated with the concept of intake. Depth of processing or amount of cognitive effort, and a low level of awareness all appear to have close relationships with intake. These two variables are discussed below.

**Depth of processing during input processing.**

Considering the various definitions and theories of intake, it appears that depth of processing may be related. Although these definitions do differ in some ways, as discussed in the previous section, most of them do make reference to some level of processing. First of all, Faerch and Kasper’s (1980) definition of intake assumes that further processing is necessary for intake to be assimilated into the interlanguage. Leow’s (1993) definition of intake states that intake includes the part of the input that has been attended to while processing the input; Chaudron’s (1985) concept of intake consists of three stages of information processing. According to Gass and Selinker, intake is the component where psycholinguistic processing takes place. Furthermore, Schmidt (1990) proposes that although learners need to consciously notice before more processing can occur, unconscious processes are responsible for most details of cognitive processing.
Many theorists highlight the intrinsic relationship between cognitive effort and attention (e.g. Kahneman & Treisman, 1984; Lachman et al., 1979; McLaughlin, 1987, McLaughlin et al., 1983; Robinson, 1995; Tomlin & Villa, 1994). Robinson (1995) described attention to be, among other things, the mental effort that processing entails. Cognitive effort dictates the amount of attention that a learner pays to L2 input (e.g., Kahneman & Treisman, 1984; Lachman et al., 1979; McLaughlin, 1987, McLaughlin et al., 1983). Furthermore, McLaughlin’s (1985) Cognitive theory states that the amount of attention that the learner pays to the L2 data depends greatly on amount of cognitive effort required by the processing of the input. Tomlin and Villa (1994) explain that attention represents effortful processing. Given that these examples serve to relate cognitive effort to attention, it is important to also note the relationship between attention and intake. Schmidt’s (1990) noticing hypothesis proposes that attention is responsible for noticing, which Schmidt defines as “the necessary and sufficient condition for the conversion of input into intake” (Schmidt, 1993, p. 209). Therefore, it appears logical that cognitive effort may be related to intake. Furthermore, McLaughlin (1987) discusses the possible link between depth of processing and cognitive effort when he purported that task difficulty and the amount of controlled or automatic processing involved may require differing amounts of cognitive effort or depth or levels of processing.

Theoretical accounts.

Craik and Lockhart (1972) first mentioned the concept ‘levels of processing’ in the field of cognitive psychology. They suggested that remembered information depends not only on having attended to it during its occurrence and having rehearsed it afterward, but also on how deeply it is processed. They distinguished between conceptual or semantic processing (i.e., deep
processing) and perceptual processing (i.e., shallow processing). Craik (2002) and Craik and Lockhart posited that memory for items is dependent upon the depth of processing during encoding. This prediction indicates that depth of processing and comprehension may be positively related, if in fact memory for items is considered to be comprehension posttest scores. They also state that shallow processing is less likely to lead to memory recall than is deep processing.

Depth of processing has recently been related to other concepts such as amount of attention and level of awareness during processing (e.g., Gardiner & Richardson-Klavehn, 2000; Leow, 2012). Several frameworks accounting for the initial stages of L2 learning also assign an important role to depth of processing. In Tomlin and Villa’s (1994) view, awareness can result in enhanced processing. Schmidt (2001) concurs with Tomlin and Villa through his concept of detection which enables further processing of a stimulus at higher levels. Also, Robinson’s model accounts for data-based and conceptually-driven processing that occur after initial intake. Furthermore, Gass purports that what she calls ‘level of analysis’ affects intake: the level of analysis of the input that a learner reaches determines whether a particular instance of comprehended input will become intake. She distinguishes between global comprehension versus a more linguistic focus. More recently, Craik (2002) has related depth of processing to elaboration and high degree of consciousness.

Cognitive effort arguably plays a role in depth of processing. Given that L2 learners are limited capacity processors (McLaughlin, 1987), L2 learners who direct their attention to a specific aspect of the input are limited in what they can attend to at that point in time. Depth of processing can be envisioned as cognitive effort with the possibility for making a form/meaning connection at higher depths of processing. Intake is also reliant on depth of processing.
Leow’s (1993) definition of intake as “that part of the input that has been attended to by the second language learners while processing the input” (p. 334), the process of the learner attending to the intake intrinsically involves the ability to process the selected part of the input, even at an extremely low level.

Operationalizing depth of processing.

The few studies that have incorporated the notion of depth of processing have either operationalized it indirectly by means of experimental condition and task instructions or directly addressed it through data from concurrent verbal reports.

The studies indirectly operationalizing depth of processing by means of task instructions have used techniques such as asking a question designed to induce processing of a target word to one of three levels of analysis immediately prior to visual exposure to the word (Craik & Tulving, 1975). Similarly, Shook (1994) operationalized depth of processing through his concept of “attention drawn” to target items versus “no attention drawn.” Gass, Svetics, and Lemelin (2003) used experimental conditions termed ‘+focused attention’ and ‘-focused attention’; the ‘+focused attention’ condition attempted to draw attention to the target item through underlining and instructions.

Studies have also used concurrent verbal reports to collect data on depth of processing. Qi and Lapkin (2001) coded concurrent verbal reports for instances of “quality of noticing,” which they defined as language-related problems that the learners themselves were able to resolve. Leow et al. (2008) and Morgan-Short et al. (2012) also looked to data from concurrent verbal reports to operationalize depth of processing. They defined three levels of processing: Level 1 participants had simply circled the target form, while Level 2 participants provided some additional report of attending to the target form, and Level 3 participants either interpreted or
translated the target form.

Critically, no previously published study that investigated depth of processing has employed separate criteria for the processing of a grammatical form versus a lexical item.

*Empirical accounts.*

A handful of studies have investigated the concept of depth of processing, but only the most recent studies have referred to it in these terms.

Craik and Tulving (1975) conducted a series of experiments to explore levels of processing and memory. In each experiment there were between 12 and 36 participants who were visually exposed to a word, before which a question was asked to induce them to process the word to one of three levels of analysis. Questions required an analysis of the word’s physical structure (lower level of processing), phonemic level (medium), or semantic category (higher). Participants took a retention test (free recall, cued recall, or recognition) of the words viewed. Results showed that when deeper level questions were asked about a word, subsequent retention of the target words was enhanced. Overall results hint that depth of processing is not a continuum. Crucially, “structural analyses do not shade into semantic analyses” (p. 290).

Shook (1994) investigated whether L2 learners can process grammatical information presented via written input as intake in light of three variables: explicit attention to grammatical item, type of grammatical item, and L2 experience. Three conditions operationalized explicit attention: no attention drawn to the items (unenhanced, no special instructions), attention drawn to the items (textual enhancement but no special instructions), and attention drawn to the items plus forming grammatical rules\(^3\) (textual enhancement and special instructions). Grammatical items were the more-meaningful present perfect tense and the less-meaningful relative pronoun.

\(^3\) As discussed in the next section, this operationalization of attention actually mirrors the concept of levels of processing, and is therefore included as a study of depth of processing.
Participants were 125 1\textsuperscript{st} and 2\textsuperscript{nd} year adult learners of Spanish who were exposed to two reading passages. Assessments included production and recognition tasks for each grammatical item. Shook reported that participants whose attention was drawn to the grammatical items gained more information about the items than those whose attention was not, but type of attention was not significant. It was explained that learner’s detection of information about the grammatical items was enhanced by explicit instructions. Notably, no concurrent data was collected; Shook instead used the three experimental conditions to operationalize the attention condition and then assumed that participants had performed according to their condition.

In a case study of two L2 English learners, Qi and Lapkin (2001) investigated what learners notice when writing and what effects this had on their text. During three sessions, participants wrote a text (Stage 1), compared it with a reformulated version (Stage 2), and revised it without access to the reformulation (Stage 3). Participants did think-alouds and were video-recorded during all tasks. Data on LRE (language-related events) were collected. Results revealed that the higher proficiency participant resolved more Stage 1 LREs and provided more reasons for reformations at Stage 2. Qi and Lapkin claim that noticing without understanding or for no clear reason (noticing without providing a reason or demonstrating an understanding of the nature of the gap between the interlanguage and the target language) does not have the same impact as noticing with understanding. Thus, while promoting noticing is important, improving the quality of noticing may be even more important.

Gass, Svetics, and Lemelin (2003) investigated the extent to which focused attention affects learning and the relationship between focused attention and proficiency. Participants were 34 English speakers enrolled in a 1\textsuperscript{st}, 2\textsuperscript{nd}, or 3\textsuperscript{rd} year Italian class. There were two experimental conditionals ([+focused attention], [-focused attention]), three linguistic areas, and
four groups into which participants were randomly placed. The procedure consisted of a lexicon pretest, experimental materials, and an immediate posttest. Experimental materials for the +focused attention group drew attention to the target item through underlining and instructions. Tasks included reading a story and then either +focused or –focused written tasks. Results showed that except for the linguistic area of syntax, learning occurred regardless of focused attention. The greatest effect of the [+focused attention] condition was in syntax, thus suggesting that focused attention is better used in more complex linguistic areas. The authors conclude that focused attention does play a role in learner development, at least in the short term.

Leow, Hsieh, and Moreno (2008) examined simultaneous attention to form and meaning to determine whether limited attentional resources during input processing compete to be allocated to either form or meaning. Participants were 72 2nd semester students of Spanish as an L2 who were randomly assigned to read for meaning only or read for meaning plus circle instances of a lexical item (sol), a feminine article (la), a masculine article (lo), or a verbal morpheme (-n). Participants read the passage, followed their condition’s directions, and recorded think alouds. A multiple-choice assessment tested comprehension. After examining the data, the researchers noticed instances of depth of processing and hypothesized, based on VanPatten’s (2004) Primacy of Meaning Principle, that participants who had reported not only circling but also deeper processing of the target forms would obtain lower comprehension scores in comparison with those who had not. They defined three levels of processing: Level 1 participants had simply circled the target form, while Level 2 participants provided some additional report of attending to the target form, and Level 3 participants either interpreted or translated the target form. The data, however, do not show a clear trend in support of this hypothesis. Regarding the effect of attention condition on L2 reading comprehension, the
quantitative analysis showed no significant comprehension differences among the five groups. Leow et al. summarize by stating that the relatively low level of processing of the target forms as found in all experimental groups may explain the non-significant difference in comprehension between the experimental conditions. Low levels of processing seemingly did not create any cognitive overload while processing for meaning. Regardless, these results must be interpreted with caution due to the absence of statistical analyses and low cell numbers.

Morgan-Short et al. (2012) replicated Leow et al. (2008) in hopes of obtaining more conclusive findings regarding the possible relationship between comprehension and depth of processing. Modifications included increasing the number of participants to 308 and adding a silent (non-think-aloud group); otherwise, all materials and procedures mirrored those of Leow et al. (2008). Results showed a small but reliable positive relationship between level of processing and comprehension.

The preliminary evidence suggests that depth of processing has a facilitative effect on several aspects of the second language learning in the early stages, however further research needs to examine this more closely. Morgan-Short et al. (2012) reported a small but reliable positive relationship between level of processing and comprehension; results from Craik and Tulving (1975) showed enhanced retention of target words with deeper levels of processing. Qi and Lapkin (2001) reported facilitative effects for what they called “quality of noticing” on learning; findings from Shook (1994) show that higher “attentional level” resulted in more information gained. Leow et al. (2008) reported that participants who reported having processed the target form deeper did not have lower comprehension scores than those who had not. On the contrary, Gass et al. (2003) reported that except in the area of syntax, learning occurred regardless of “focused attention.” Gass et al. suggests that “focused attention” may be better
used in more complex linguistic areas.

It is evident that the relationship between depth of processing and comprehension is a ripe area of investigation. For example, VanPatten’s primacy of meaning principle (1994, 1996) logically predicts that processing at deeper levels may interfere with written comprehension because the deeper processing of form conflicts with his first principle that learners process input for meaning before processing it for form. Regardless, Leow et al. (2008) found that participants who processed the target form deeper did not have lower comprehension than those who did not. Further complicating the matter, Morgan-Short et al. (2012) did find a positive relationship between level of processing and comprehension. Additional studies need to further address this possible relationship between depth of processing and comprehension.

One of the most serious limitations to previous studies regards the terminology and basic theoretical underpinnings surrounding attention and processing. Specifically, one can pay attention to an item in the input but this does not always translate to it being further processed. Several researchers have conflated the terms "(level of) attention” with “(level of) processing:” Shook (1994) referred to “attentional levels,” Gass et al. (2003) used the term “focused attention,” and Qi and Lapkin (2001) talked about “quality of noticing.” Furthermore, VanPatten (2004) uses the term “attention to form and meaning” when he is actually describing processing for form and meaning simultaneously. It appears that all of these researchers are actually referring to processing instead of mere attention. In addition to clarifying terminology and concepts, future studies should include concurrent data to be able to address levels of processing. Gass et al. (2003), for example, did not employ concurrent verbal reports; more decisive conclusions could have been made had they used this measure.

Another limitation of previous studies involves the depth of processing revealed in the
data. Simultaneous attention to form and meaning may not have been adequately operationalized in previous studies due to the aforementioned conflation of “attention” and “processing.” As Leow et al. (2008) explain, “the previous operationalization of attention to form might not be robust enough to address the issue of processing that form in relation to its connection with its meaning or function” (Leow et al, 2008, p. 686). Therefore, it appears that future studies should avoid techniques to measure attention such as writing a check mark on a piece of paper when hearing the target form. Leow et al. (2008) suggested that this strand of research needs to employ both assessment tasks and target items that are more likely to facilitate deeper processing of form at the level of form-meaning connections. Furthermore, they asserted that future research should include intake as a dependent variable in studies of processing so that another stage of the acquisitional process may be addressed.

In summary, there is little doubt that attention plays an important role in L2 learning. The early stages of SLA frameworks maintain that variables such as recognition, comprehension, prior knowledge, working memory, and depth of processing are all important factors in the learning process. While empirical studies have addressed all of these variables to some extent, future research still needs to better operationalize and define them in order to more clearly determine the role they play in the early stages of SLA.

*Awareness in input processing.*

The role of awareness in relation to intake is debatable. According to Tomlin and Villa (1994), awareness may enhance input processing, but unlike detection, it is not required. Likewise, McLaughlin (1987) theorizes that awareness is not always necessary at these early stages and the amount of controled or automatic processing may determine how much cognitive effort and depth or level of processing are required. On the contrary, Schmidt’s (1993) noticing
hypothesis holds that attention controls access to awareness and is responsible for noticing (attention plus a low level of awareness) that is crucial for intake to take place. Robinson’s model of the relationship between attention and memory reconciled these two positions by placing detection sequentially prior to noticing along the acquisitional process, indicating that awareness may play a more important role in input processing at the stage of intake. Given these positions, it seems plausible that depending on the type of intake, a low level of awareness may be involved.

_Empirical research on awareness at the level of noticing._

Several empirical studies employing online verbal reports have investigated awareness at the level of noticing.

Leow (1997b) (see section “Empirical studies” under the intake section of this current study for more information) found that posttest recognition scores were attributed to level of awareness, which was measured by the ability of participants to comment on the morphological rules of the target form. Leow concluded that increased awareness contributes to more recognition and accurate written production; he attributed this to the ability of awareness to enhance further processing of the target form.

As previously reviewed in the intake section “Empirical studies,” Rosa and O’Neill (1999) investigated level of awareness raised while processing input. Results showed that the higher the level of awareness demonstrated by participants, the stronger the effect on intake. The group showing awareness at the level of noticing and the non-think-aloud group showed no

---

4 Adhering to Leow’s (1997b, 1998a, 1998b) position that one can make claims regarding learners’ internal processes when there is evidence of such processes, the author has chosen to only include empirical studies on awareness at the level of noticing that use an online measure of awareness because this method arguably measures awareness without the having to account for veridicality (for more information, see section Concurrent vs. non-concurrent verbal reports).
instances of creating and testing hypotheses about the nature of the target structure, while this behavior was very common in participants showing awareness at the level of understanding. Furthermore, the think-alouds from the group with awareness at the level of understanding supported the association between higher levels of awareness (meta-awareness in the case of Leow, 1998b) and conceptually driven processing.

Rosa and Leow (2004a) also investigated whether different levels of awareness influenced L2 learners’ ability to recognize contrary-to-fact statements (see the section “Empirical studies” of the Intake section for more detail). Results indicated that higher levels of awareness were associated with more explicit conditions and were also more effective than lower levels (awareness at the level of noticing). The level of awareness reported through both think-aloud protocols and postexposure questionnaires had a differential effect on participants’ ability to recognize new ideas: participants who showed awareness at the level of understanding scored significantly higher on the generalization test than those demonstrated awareness at the level of noticing or who did not evidence awareness in their reports. Importantly, even those reporting awareness at the level of noticing scored significantly higher than those who did not think aloud, thus corroborating findings from Leow (1997b, 2001a) and Rosa and O’Neill (1999). However, these findings differ from those of Robinson (1995a, 1996), who reported that only awareness at the level of understanding had a significant impact on learning.

In summary, a majority of the SLA studies that have investigated awareness by the use of concurrent verbal reports appear to show that awareness facilitates L2 learning. Three levels of

---

5 While Rosa & Leow 2004a found that awareness at the level of noticing had a positive effect on learners’ performance, Robinson (1995a, 1996, 1997) found that only awareness at the level of understanding significantly impacted learning, however this discrepancy may be due to the different nature of assessment test used. Leow, Rosa and O’Neill, and Rosa and Leow tested learners’ performance via recognition tests measuring intake, Robinson assessed learning through grammaticality judgment tests.
awareness for grammatical items have been reported in these studies: awareness at the level of noticing, awareness at the level of reporting, and awareness at the level of understanding (Leow, 2001a; Rosa & Leow, 2004a; Rosa & O’Neill, 1999). These studies highlight two main ideas regarding awareness at the level of noticing: while awareness at the level of understanding facilitates intake (Leow, 1997b, 2001; Rosa & O’Neill, 1999) and production (Leow, 1997b, 2001a; Rosa & Leow, 2004a) to a greater degree than awareness at the level of noticing, participants who report awareness at the level of noticing still often score higher on the same measures than participants in a non-reported awareness group (Leow 2001a; Rosa & Leow, 2004a; Rosa & O’Neill, 1999). Second, it appears that awareness at the level of noticing may be indicative of certain types of processing: Leow (2001a) found that the absence of meta-awareness, which is characteristic of awareness at a higher level, was correlated with the absence of conceptually driven processing.

Considering the above links between intake and depth of processing, cognitive effort, and a low level of awareness, it can be concluded that these three variables all can be posited to play a role in intake.

Based on the findings from these aforementioned studies, future research should investigate this possible distinction between three levels of intake as supported by empirical studies and backed by theoretical approaches. As evidenced previously, these levels differ in their presence of awareness, depth of processing, level of cognitive effort, and type of attention.

One recent attempt to address the above is Leow’s (forthcoming) Model of the L2 learning process, which is presented below.

Figure 2: A Model of the L2 Learning Process (Leow, forthcoming)
Leow’s Model of the L2 Learning Process.

Leow’s (forthcoming) Model of the L2 learning process expands the course-grained stages of the learning processing that is supported by most theoretical perspectives (e.g., Gass, 1997; McLaughlin, 1987, Robinson, 2005; Schmidt, 1990; Tomlin & Villa, 1994; VanPatten, 2004) by adding the notion of learning as consisting of both processes and products and elaborating on the first two stages (elaborated below) of input processing and intake processing in light of cognitive processes postulated to play important roles in the learning process. It is premised on the role of attention in the process of L2 learning; in other words, the process of learning anything, especially in the L2 classroom context, is unlikely to occur unless the learner minimally pays attention to new stimuli in the L2 data.
Several variables that may accompany the allocation of attention resources serve to modulate learner attention to new linguistic information in the input; these include the depth of processing or the amount of cognitive effort employed while paying attention, cognitive registration, and level of awareness of the input. Leow’s model (forthcoming) contains both stages during which L2 linguistic data is processed (e.g., Chaudron, 1985; Faerch & Kasper, 1980; Gass, 1988; Slobin, 1985) and produced. There are three processing stages (input processing stage, intake processing stage, output processing stage), each of which is described below.

**Input processing stage.**

Input processing is the first processing stage in Leow’s model and is situated between the input and the intake of linguistic data. Dependent upon whether the learner’s attention is selective/focal or peripheral, this stage may be accompanied by depth of processing, cognitive registration, and level of awareness. Therefore, the input processing stage is divided into three phases: attended intake, detected intake, and noticed intake.

Attended intake is accompanied by peripheral attention paid to some linguistic data in the input but does not include any high level of processing, cognitive registration, or awareness of the linguistic data (Leow, forthcoming). The intake phase described here is similar to Chaudron’s (1985) description of the initial stages of perception of input and Gass’ (1988) notion of apperception, however attended intake in this model is characterized by an extremely low level of processing. Because input that is peripherally attended to is unlikely to be processed further, it is most likely to be discarded without being stored in worked memory or processed further (Corder, 1967).
Detected intake is the second phase of the input processing stage and presupposes that some amount of either peripheral or selective attention was minimally paid to the linguistic data, accompanied by a very low level of processing. This level of intake is characterized by a state of unawareness, even though the learner cognitively takes note of the new linguistic data. In line with Tomlin and Villa’s notion of detection, or cognitive registration without awareness, detected intake also has a potential for storage in working memory and further processing in comparison with attended intake. However, this may depend on the learner’s current working memory and whether or not a higher level of processing or cognitive effort is allocated subsequently to the detected intake (e.g., Leow, 2001; Bowles, 2005).

 Noticed intake is the third intake level of input processing and indicates that the learner attended to the linguistic data and cognitively registered them with the presence of a low level of awareness; however this level is still accompanied by only a low level of processing (Leow, forthcoming). This phase is in line with Schmidt’s (1990) notion of noticing (attention accompanied by a low level of awareness). Due to the relatively higher level of processing and cognitive effort in comparison with those of attended and detected intake, noticed intake has the highest potential to be stored in working memory and made available for further processing, possibly leading to incorporation into the learners’ grammar system; this has been reported in studies addressing the effects of noticing in L2 development (e.g., Leow, 1997b, 2000; 2001a; Leow et al., 2003; Martínez-Fernández, 2008; Rosa & Leow, 2004a; Rosa & O’Neill, 1999). Although noticed and detected intake, and to a lesser extent, attended intake, hold the possibility of being lodged in working memory and made available for subsequent recognition, they can also all be discarded if they are not minimally processed further (e.g., Bowles, 2005; Leow, 2001a).
Intake processing stage.

Attended, detected, or noticed intake can be considered preliminary intake and may be processed through either data-driven processing or conceptually-driven processing, depending on depth of processing and/or cognitive effort (Leow, forthcoming). Data-driven processing (cf. Robinson, 1995) allows the data to enter the learners’ internal system encoded as a chunk of language that is non-systemized without much cognitive effort expended in processing. If subsequent exemplars are also present without higher levels of processing, they may also be processed the same way, thus forming a collection of encoded discrete data that are lodged in the internal system. Ways to measure these entities include interpretation tests or simple controlled production assessment tasks of old exemplars (cf. Leow, 1997b).

On the contrary, preliminary intake can also be processed via conceptually-driven processing (cf. Robinson, 1995), another stage of further processing, if it occurs in the presence of relatively higher levels of processing such as encoding and decoding the linguistic data (Leow, forthcoming). Higher levels of awareness may accompany this stage and function to keep the linguistic data present in working memory in order to facilitate its potential incorporation into the internal system. There are two ways that this stage may minimally occur: through either linking the linguistic data to prior knowledge or by means of data-driven processing. In the case of the first option, the exemplar is cognitively linked to prior knowledge of a related linguistic datum that is then used to aid in the encoding and decoding to the linguistic information contained in the preliminary intake. Leow (forthcoming) gives the example of an L2 learner’s activation of the stem-changing verb forms in the present tense of the third person singular of morir (“to die”) > muere (“he dies” and the subsequent linking of this vocalic stem change to the irregular preterit third person form of the verb murió (“he died”). A conceptually-
driven process like this requires a higher level of processing in addition to a higher level of awareness. Subsequent exemplars that reactivate the same prior knowledge will cause the reduction of the level of awareness and depth of processing that are required to process the L2 data.

The second way that conceptually-driven processing may occur closely follows the data-driven processing path in that the linguistic data is encoded and lodged in the internal system but unsystemized. When subsequent exemplars are taken in, intake processing may occur in two ways: through activation of old (as described above) or new data. Regarding new data, the linguistic unsystemized data stored recently in the internal system can be reactivated by continued exposure to the same or similar linguistic data and may occur in the presence of a higher depth of processing, cognitive effort, and/or levels of awareness. Implicit restructuring may occur with a low depth of processing, in which case the L2 data would be stored in the grammatical system within the system learning component. However, this type of processing is dependent upon many factors, such as the presence of a great amount of exemplars in meaningful context and a long time period to allow for processing.

Higher depths of processing typically include hypothesis testing and rule formation (cf. Leow, 1997b; Rosa & Leow, 2004a; Rosa & O’Neill, 1999); potential level of awareness also increases from awareness at the level of noticing to awareness at the level of reporting to awareness at the level of understanding (Leow, 1997b). It is important to note that while higher depths of processing may assist in achieving higher levels of awareness, the relationship is not always true: it is logical that a higher depth of processing involves greater amounts of cognitive efforts when trying to hypothesize about an underlying grammatical rule or a lexical root, awareness at the level of understanding is only attained when the underlying rule obtained is
correct or the form-meaning connection of a lexical item is achieved. In other words, awareness at the level of understanding subsumes a high depth of processing but the opposite is not true, that is, a higher depth of processing does not necessarily lead to awareness at the level of understanding given that one may process a target item incorrectly at a high level of processing but does not achieve an awareness of the full understanding of that item. The linguistic data is restructured if necessary and stored into the grammatical system within the learning component with the help of prior knowledge activation, levels of processing, and the potential for higher level of awareness.

In summary, this second stage of further processing occurs between preliminary intake (attended, detected, and noticed) and the internal system. Cognitive effort, depth of processing, levels of awareness, amount of exemplars, conceptually-driven processing (prior knowledge), data-driven processing, and restructuring are all important variables in this intake processing stage. Other variables such as, for example, individual differences (e.g., working memory capacity, motivation etc.), and type of linguistic item may also play a role at this stage.

**Internal system.**

In the internal system, there are two kinds of product or stored linguistic knowledge of what has been processed thus far in the learning process: unsystemized data (discrete linguistic data) and systemized (internalized or learned) data. Gass’ (1988) postulation is similar to this separation of internalized data. Whether stored linguistic data becomes unsystemized or systemized data depends heavily on the levels of cognitive effort, depth of processing, and awareness (Leow, forthcoming). Although accuracy of the product is not necessary at this point, higher levels of awareness and increased accuracy have been reported to be correlated in several empirical studies (e.g., Leow, 1997b; Rosa & O’Neill, 1999; Rosa & Leow, 2004a, 2004b).
Both awareness at the level of understanding and automatization of the linguistic data through subsequent multiple exposures and meaningful practice result in a drastic reduction in the amount of cognitive effort that is required for processing the relevant linguistic data in the L2 input. This may lead to a secondary role for awareness and depth of processing during intake processing of linguistic data that was previously learned (Calderón, 2013).

There are two crucial features of this model of L2 learning. First, it is important to note that Leow’s (forthcoming) model is not linear given that learners’ output may also function as additional input, which is represented by the arrow going from output back to input in Error! Reference source not found.. Further input, whether it be in the form of additional exposure to the L2 linguistic data or learners’ own output (e.g., Swain’s output hypothesis), may serve to create more robust processing of the linguistic data during the learning process. This may be viewed as activation of prior knowledge of the product in the internal system and its relationship with the linguistic information taken in. The second lies in the existence of both processes and products depending upon the stage of the overall learning process, that is, from input to output.

In conclusion, Leow’s (forthcoming) Model of the L2 learning process provides a finer-grained explanation of the L2 learning process by adding the concept of learning as containing both products and processes, in addition to elaborating on the first two stages of input processing and intake processing, all the while highlighting the important role that cognitive processes play in the learning process. Most notably, Leow’s Model contains three phases of intake (attended, detected, noticed) at the input processing stage, each of which may be accompanied by cognitive registration and varying levels or depth of processing and awareness.

The numerous empirical studies reviewed earlier (see “Empirical studies” of the intake section) that appear to suggest the existence of three separate levels of intake fit Leow’s
(forthcoming) three phases of intake and therefore, from here on, a high level of intake will be referred to as “noticed intake”, a medium level of intake will be referred to as “detected intake”, and a low level of intake will be referred to as “attended intake.”

Given that the present dissertation seeks to address also one of the variables, namely, type of linguistic item, mentioned by Leow to potentially play a role in input and intake processing, the following section reviews the literature on type of linguistic item in SLA.

**Type of Linguistic Item**

Discussions of type of linguistic item often include several concepts in order to describe the differences between the linguistic items, such as saliency, meaningfulness, difficulty, complexity, and attention to form and meaning. Three of the overarching concepts are examined below: difficulty, attention to form and meaning, and saliency.

Determining the difficulty of linguistic items is one of the first steps in analyzing any study comparing differences in linguistic items. DeKeyser (2005) states that there appear to be a minimum of three factors involved in determining grammatical difficulty: complexity of form, complexity of meaning, and complexity of the form-meaning relationship. Meaning can be especially difficult for the L2 learner because of novelty or abstractness. According to DeKeyser, articles, classifiers, grammatical gender, and verbal aspect are especially difficult for L2 learners who do not have similar linguistic concepts in their L1. These elements are also notoriously abstract and are thus very difficult to infer from the input. Another issue contributing to the issue of grammatical difficulty for the L2 learner is the problem of form, which DeKeyser explains to be “the number of choices involved in picking all the right morphemes and allomorphs to express these meanings and putting them in the right place” (pp. 5-6). This problem is most pertinent in languages with rich inflection, such as Spanish. He
continues by explaining that learners tend to gloss over the morphology that is essential to sentence meaning. The third problem contributing to grammatical difficulty is that of form-meaning mapping, a step that can be very hard if the link between the form and the meaning is not transparent. Transparency involves at least three factors: redundancy refers to the form not being necessary because its meaning is also expressed elsewhere in the sentence, optionality refers to those optional grammar elements whose alternating presence or absence coupled with the same meaning, and opacity is a low form-meaning correlation, such as that occurring when many meanings are expressed by the same morpheme (e.g., third-person singular in English).

The concept of saliency is often mentioned in studies investigating type of linguistic item, however it is very complex construct and difficult to succinctly define. It can be described in terms of the physical or perceptual features of linguistic structure; Slobin (1971) and Brown (1973) stated that perceptual salience can depend on phonetic substance, stress, pitch, word position, and utterance position. Goldschneider and DeKeyser (2001) defined perceptual salience as “how easy it is to hear or perceive a given structure” (p. 22), they also loosely defined salience as a combination of phonological salience, semantic complexity, morphological regularity, and frequency. In a study of morpheme acquisition order, Goldschneider and DeKeyser (2005) found that salience accounted for the majority of the variance in the order of L2 acquisition.

**Empirical studies.**

In order to determine how different linguistic items are learned in the L2 setting, studies in the field of SLA (see Table 1) have used as target items a wide variety of grammatical forms (e.g., Ellis, 2007; Leeman, 2003; Leow et al., 2003; Ortega & Long, 1997; Shook, 1994; Yilmaz, 2012; Yilmaz & Yuksel, 2011) and to a lesser extent, lexical items (Martínez-Fernández, 2008).
The inclusion of these target structures and lexical items are often motivated by differing views regarding saliency (e.g., Leeman, 2003; Ortega & Long, 1997; Yilmaz, 2012; Yilmaz & Yuksel, 2011) and complexity (Uggen, 2012). However, there are very few studies that directly compare some measurement of grammatical versus lexical linguistic items. There are two strands of SLA research that have investigated type of linguistic item: the corrective feedback strand and the attentional strand. Although this current study most definitely falls under the attentional/processing/comprehension strands, this review of previous research on type of linguistic includes studies from both strands in an effort to review all pertinent studies and achieve a better understanding of the role of type of linguistic item in L2 development. Therefore, empirical studies focusing on corrective feedback that also include consideration of type of linguistic item will be presented first, followed by a review of studies in the attentional/processing/comprehension stand that consider type of linguistic item. Basic information on all studies reviewed is available in Table 1.

**Type of linguistic item in corrective feedback studies.**

A sizable amount of the studies comparing type of linguistic item, whether it be grammatical-grammatical or grammatical-lexical, are studies on corrective feedback. The following first summarizes several previous studies comparing either two grammatical items or two lexical items, and then more detail is given on the five studies that classified the linguistic focus of corrective feedback into grammatical categories (Loewen & Philp, 2006; Mackey, Gass, & McDonough, 2000; Nabei & Swain, 2002; Smith, 2012; Smith & Renaud, 2013), thus comparing target grammatical forms with target lexical forms.

As seen in Table 1, the majority of empirical studies on corrective feedback that compare two or more linguistic items have as their targets two grammatical forms (Ellis, 2007; Leeman,
2003; Ortega & Long, 1997; Yilmaz, 2012; Yilmaz & Yuksel 2011). Of these studies, the target linguistic items of several differed in saliency (more or less salient) (Leeman, 2003; Yilmaz, 2012; Yilmaz & Yuksel, 2011). Overall, increased salience seems to have had a positive effect on several aspects of L2 development (Leeman, 2003; Yilmaz & Yuksel, 2011); only Yilmaz (2012) found no difference at all in posttest scores for two grammatical forms differing in saliency. Two additional studies compared two grammatical targets: Ellis (2007) found that feedback had a greater effect on the English comparative (-er) than on the past (-ed) and findings from Ortega & Long (1997) revealed that while there was no learning of topicalization, participants were able to recognize and produce exemplars of adverb placement.

In addition to the studies investigating differences in salience and meaningfulness, there are five empirical studies (Loewen & Philp, 2006; Mackey, Gass, & McDonough, 2000; Nabei & Swain, 2002; Smith, 2012; Smith & Renaud, 2013) comparing type of linguistic item that did not have specific target linguistic items but rather classified the linguistic focus of corrective feedback into grammatical categories. In a study on interactional feedback, Mackey, Gass and McDonough (2000) investigated learners’ perceptions regarding morphosyntactic, lexical, phonological, and semantic errors. Participants were 10 adults of mixed L1 who were learning English as a second language and seven L1 English adults learning Italian. The L2 English learners had an average of 9.3 years of previous English study while the L2 Italian learners had an average of 1.8 years of previous Italian study. Each participant worked individually with a native speaker or a near-native speaker to perform a video-taped two-way information exchange activity in the L2, during which the (near-) native speaker provided feedback on errors in the form of negotiation and recasts. Morphosyntactic errors included grammatical aspects such as agreement and tense and lexical errors were those that contained comments about a known or
unknown word; phonological and semantic errors are not reviewed in this current study.

Following the information exchange activity, the participant completed a recall procedure in English in which the researcher asked participants to recall their thoughts at the time of each feedback episode, thus obtaining information regarding participant perceptions at the time of the feedback. Results revealed that participants were most accurate about their perceptions of lexical and phonological feedback; in other words, they most often realized that lexical feedback and phonological feedback were indeed regarding the lexicon and phonology, respectively. They were overwhelmingly inaccurate regarding their perceptions of morphosyntactic feedback; the L2 English participants most often perceived it as semantic while the L2 Italian participants thought it was lexical. Mackey et al. also analyzed uptake, operationalized by participants modifying their output in response to feedback. Results showed that uptake and accurate perception occurred for 82% of the lexical errors and only 33% of the morphosyntactic errors. A possible explanation for this difference offered by the authors is that morphosyntax can be unimportant in the ultimate goal of understanding, however it is also important to consider that morphosyntactic feedback that was generally not perceived as being morphosyntactic was most often provided via recasts, while lexical feedback was most often provided via negotiation. The authors conclude by stating that future research that measures development is needed to answer the many questions remaining, including the differences in development between different types of linguistic items.

The second feedback study using a set of categories to classify feedback targets was Nabei and Swain (2002), who set out to determine what connections exist among teacher recasts, student’s awareness, and that student’s learning. Since this was a case study, there was one single participant, a native Japanese speaker, who took English during high school and
extensively used English in social settings. After six cycles of classroom observations, GJT, and stimulated recall, all targets of feedback were divided into grammatical, lexical, and phonological categories. The authors claimed that the effect of recasts was influenced by the linguistic elements of the feedback, however the only mention of type of linguistic item in the results section (besides a breakdown of the linguistic focus of the researcher’s recasts) was that the participant was observed ignoring grammatical correction provided by the recasts and instead focusing on lexical information. Overall, this study needs to be regarded with caution considering that it is a case study of one individual.

Although they did not specifically investigate type of linguistic item, Loewen and Philp (2006) also classified feedback targets according to linguistic type (lexical, morphological/syntactic, phonological, and combination groups). Their participant base numbered 73 L2 English learners of varied L1s. Participants completed meaning-focused activities during which corrective feedback was provided and then completed individual tests consisting of correcting sentences with errors that they previously produced and reading aloud their original transcribed utterance. They found that morphosyntactic, lexical and phonological errors each received roughly equal levels of feedback.

Smith (2012) aimed to determine the relationship between the linguistic category of the recast target and noticing and therefore the targets of recasts were divided into morphological errors, semantic errors, and syntactic errors (see section “Eye-tracking as a measure of processes related to the attentional strand.” for more details on this study). After the treatment task of watching a video and chatting online with a NS, participants wrote a summary of the story and

---

6 Loewen and Philp (2006) only reported the percent of feedback that was classified into each linguistic category; they did not look at type of linguistic item as a separate variable. Therefore, the study cannot contribute much to the current discussion beyond the fact that type of linguistic item was taken into account, but it is being included in this current study simply because it did indeed consider linguistic item.
then completed a stimulated recall session with the researcher. Results of the stimulated recall showed that the probability of noticing syntactic recasts and semantic recasts was significantly higher than noticing morphological recasts, however there was no significant relationship for linguistic category and noticing in the eye-tracking condition. Smith explains the advantage of semantic and syntactic recasts over morphological recasts by referring to VanPatten’s (2007) argument that learners prefer process lexical information over other input features.

Smith and Renaud (2013), as detailed extensively in the section “Eye-tracking as a measure of processes related to the attention strand”, looked at type of linguistic item in their study on L2 recasts, noticing, and learning during SCMC. The target of each recast was coded for type of linguistic items (lexical, agreement, tense, spelling, other). Results showed that the linguistic focus of the target did not affect whether or not participants fixated on it.

In summary, a number of feedback studies have investigated type of linguistic item, and in many cases, the comparison was between two grammatical items. The overwhelming majority of these studies comparing two types of grammatical forms found some difference between forms (Ellis, 2007; Leeman, 2003; Ortega & Long, 1997; Yilmaz & Yuksel, 2011); only Yilmaz (2012) found no difference at all in posttest scores for two grammatical forms. In the empirical studies comparing the linguistic category of the focus of corrective feedback, Mackey et al. (2000) found participant morphosyntactic scores to be significantly lower than lexical scores; Smith (2012) reported that participants noticed syntactic and semantic recasts significantly more than morphological recasts.

**Type of linguistic item in attentional/processing/comprehension studies.**

Several SLA studies in the attention/processing/comprehension stands have also investigated the effects of type of linguistic item on L2 learners’ attention, processing, and
comprehension of targets found in the input. The majority of these studies has compared two distinct grammatical items (Collentine, 1997; Leow, 1993, 1995; Leow et al., 2003; Shook, 1994; Uggen, 2012). Martínez-Fernández (2008) compared two types of lexical items, but only four studies (Greenslade et al., 1999; Leow et al., 2008; Morgan-Short et al., 2012; VanPatten, 1990) specifically compared two grammatical forms and a lexical item. Therefore, the following first provides an overview of the studies comparing the same type of linguistic item and then looks in more detail at the studies that compare a grammatical form and a lexical item.

Among the studies comparing the same type of linguistic item, Leow (1993, 1995) motivated the use of the present perfect and present subjunctive as target forms by the hypothetically higher communicative value that the present perfect form has in comparison to the present subjunctive form. He found no significant differences in recognition of these grammatical forms when the text was written (1993) but when the mode was aural, recognition of the form with higher communicative value was significantly higher; Leow concluded that the perfect tense may carry a higher communicative value when the input is aural. Leow et al. (2003) also looked at the same grammatical target forms in a reading passage with think-aloud protocols followed by recognition and comprehension tests; results showed that the more salient grammar form (present perfect) had significant benefits in terms of the amount of noticing of the target; type of linguistic item did not significantly affect comprehension or recognition. Regarding meaningfulness, Shook’s (1994) results showed that participants were better able to recognize the more meaningful target (present perfect tense) than the less-meaningful target (relative pronoun). Collentine (1997) compared regular versus irregular forms of the present perfect in a word arranging task and found that a significantly longer time was needed for the irregular form, however there was no difference between the two forms in regards to accuracy.
In a study investigating whether completing production tasks impacts learners’ attention to target structures in subsequent input, Uggen (2012) found that grammatical complexity positively affected noticing; she explains this by pointing out that the simpler structure may have been less salient. Only one study investigated two types of lexical items: Martínez-Fernández (2008) compared concrete nouns with abstract nouns and found that all experimental groups performed higher on concrete nouns than on abstract nouns and that retention was significantly lower for abstract nouns than for concrete nouns.

Table 1: Summary of Past Studies Comparing Two Grammatical Items or Two Lexical Items

<table>
<thead>
<tr>
<th>Name &amp; Year</th>
<th>Research Questions (RQs) and Hypotheses (H)</th>
<th>Participants (N, L1)</th>
<th>Linguistic items</th>
<th>Treatment</th>
<th>Measure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>VanPatten (1990)</td>
<td>H: Comprehension will be adversely affected if learners are required to focus on form in the input.</td>
<td>N=202 L1=English Level: 1st &amp; 4th semester, 3rd year</td>
<td>Lex-Gram-Gram (inflación, la, -n in Spanish)</td>
<td>2 reading passages</td>
<td>attention: write a check mark upon hearing targets</td>
<td>for comprehension, performance on inflación was significantly better than on la and –n; no difference between performance on la and –n</td>
</tr>
<tr>
<td>Leow (1993)</td>
<td>H2: Learners will take significantly more present perfect forms than present subjunctive forms in the input.</td>
<td>N=137 L1=English Level: 1st &amp; 4th semester</td>
<td>Gram-Gram (present perfect, present subjunctive in Spanish)</td>
<td>2 reading passages</td>
<td>multiple choice recognition test</td>
<td>no difference between present perfect and present subjunctive</td>
</tr>
<tr>
<td>Shook (1994)</td>
<td>RQ2: What effect does type of grammatical item in the input have on L2 learner-readers’ intake?</td>
<td>N=125 L1 not stated but assumed to be English Level: 1st &amp; 4th semester</td>
<td>Gram-Gram (more-meaningful present perfect tense, less-meaningful relative pronoun in Spanish)</td>
<td>exposure to 2 reading texts</td>
<td>production and recognition tasks for each of the targets</td>
<td>participants were better able to recognize the more meaningful target than the less-meaningful target</td>
</tr>
</tbody>
</table>

7 In an attempt to present only information pertinent to the current discussion of linguistic item, information regarding experimental conditions is not included in this table. The experimental conditions in these studies were largely based on different types of feedback.
8 The information in this table only includes those details relating the linguistic item variable of the study in question. Information regarding other variables is excluded.
<table>
<thead>
<tr>
<th>Study</th>
<th>Methods</th>
<th>Findings</th>
<th>Additional Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leow (1995)</td>
<td>N=213; L1=English; Level: 1st &amp; 4th semester</td>
<td>2 aural passages; multiple-choice recognition task</td>
<td>Intake of present perfect was greater than that of present subjunctive</td>
</tr>
<tr>
<td>Collentine (1997)</td>
<td>N=30; L1=English; Level: 2nd year</td>
<td>8 sentences in a word arranging task; processing time, word arranging task</td>
<td>For processing time, significantly longer time was needed for irregular form; for accuracy, no difference between regular and irregular</td>
</tr>
<tr>
<td>Ortega &amp; Long (1997)</td>
<td>N=30; L1 not stated but assumed to be English</td>
<td>Gram-Gram (direct object topicalization, adverb placement in Spanish)</td>
<td>N/A</td>
</tr>
<tr>
<td>Greenslade et al. (1999)</td>
<td>N=53; L1=English; Level: 3rd semester</td>
<td>2 reading passage; circle, underline, or put a check mark on target in reading passage; recall comprehension task</td>
<td>N/A</td>
</tr>
<tr>
<td>Mackey et al. (2000)</td>
<td>N=17; L1=mixed and English; average years English study and Italian study</td>
<td>Information exchange activity with (near) native speaker with feedback on errors</td>
<td>Best accuracy on perceptions of lexical and phonological feedback, worst on morphosyntactic</td>
</tr>
<tr>
<td>Study</td>
<td>Research Question 3 (RQ3)</td>
<td>Participants, Level, and Settings</td>
<td>Methodology and Procedure</td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Nabei &amp; Swain (2002)</td>
<td>RQ3: What connections are there among the teacher recasts, Shoko’s awareness and her learning?</td>
<td>N=1, L1=Japanese, Level=English classes during high school; extensive social use</td>
<td>targets varied and were divided into categories: grammatical, lexical, phonological</td>
</tr>
<tr>
<td>Leeman (2003)</td>
<td>RQ3: Does exposure to input with enhanced salience of positive evidence lead to greater L2 development than exposure to input with unenhanced positive evidence? RQ4: If exposure to recasts leads to greater L2 development than exposure to input with unenhanced positive evidence, can these benefits be attributed to either negative evidence or enhanced salience of positive evidence?</td>
<td>N=74, L1=English, Level: Gram-Gram (noun-adjective agreement in Spanish, varying in enhanced or unenhanced saliency)</td>
<td>information-gap activities completed individually with researcher, who provided recasts</td>
</tr>
<tr>
<td>Leow et al., (2003)⁹</td>
<td>RQ3: Do readers exposed to present perfect forms report significantly more noticing than readers exposed to present subjunctive forms? RQ4: Is there a significant relationship between the reported noticing of present perfect or present subjunctive forms and the immediate recognition of these forms? If so, are the two more salient linguistic item had significant benefits in terms of the amount of noticing of the target; type of linguistic item did not significantly affect comprehension or recognition</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

⁹ Leow (1998b) used two different grammatical forms (third person singular and third person plural of stem-changing preterit, but there was no analysis of any measure to investigate the difference between the two forms.
<table>
<thead>
<tr>
<th>Study</th>
<th>RQ</th>
<th>Participants</th>
<th>Intervention</th>
<th>Measures</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loewen &amp; Philp (2006)</td>
<td>RQ6: Does exposure to type of linguistic item have a significant effect on L2 readers’ comprehension?</td>
<td>N=73, L1=varied non-English Level: intermediate to upper intermediate</td>
<td>Targets varied and were divided into categories: lexical, morphological/syntactic, phonological and combination</td>
<td>Individual tests in which participants corrected incorrect sentence or read aloud original transcribed utterance</td>
<td>Morphosyntactic, lexical, and phonological errors each received roughly equal levels of feedback</td>
</tr>
<tr>
<td>Ellis (2007)</td>
<td>Does metalinguistic feedback have a differential effect on the acquisition of the target?</td>
<td>N=34, L1=English</td>
<td>Gram-Gram (English past tense ‘-ed’ and comparative ‘-er’)</td>
<td>4 instructional tasks (storytelling and comparative statements)</td>
<td>Oral imitation test, GJT, metalinguistic knowledge test</td>
</tr>
<tr>
<td>Leow et al. (2008)</td>
<td>Does type of attentional condition have a differential effect on adult L2 readers’ subsequent comprehension of text content?</td>
<td>N=72, L1=English Level: 2nd semester</td>
<td>Lex-Lex-Morph-Gram (sol, la, lo, -n)</td>
<td>Reading text for meaning only or meaning plus circle instances of target while thinking aloud</td>
<td>Multiple-choice comprehension assessment</td>
</tr>
<tr>
<td>Martínez-Fernández, (2008)</td>
<td>RQ2: What effect do tasks with different degrees of involvement load have on vocabulary development, and is that effect different for concrete and abstract nouns?</td>
<td>N=45, L1=English Level: 4th semester</td>
<td>Lex-Lex (4 concrete nouns &amp; 4 abstract nouns in Spanish)</td>
<td>Retell in writing contents of reading text; 4 translation and recognition tests</td>
<td>All groups performed higher on concrete than abstract nouns; retention was significantly lower for abstract than for concrete nouns on all tests</td>
</tr>
<tr>
<td>Yilmaz &amp; Yuksel (2011)</td>
<td>H2: Learners will score higher on the more salient target than on the less salient target after</td>
<td>N=24, L1=English Level: beginner</td>
<td>Gram-Gram (plural morpheme /-1Ar/)</td>
<td>2 oral picture description tasks</td>
<td>H2: no difference on posttest scores for two targets when...</td>
</tr>
<tr>
<td>Study</td>
<td>Research Question</td>
<td>Participants</td>
<td>Procedure</td>
<td>Dependent Variables</td>
<td>Findings</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------</td>
<td>--------------</td>
<td>-----------</td>
<td>--------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Morgan-Short et al. (2012)</td>
<td>RQ2: Does simultaneous attention to form and meaning in L2 written input affect comprehension?</td>
<td>N=361, L1=no stated, Level: 3rd semester</td>
<td>Reading text for meaning only or meaning plus circle instances of target while thinking aloud</td>
<td>Multiple-choice comprehension assessment</td>
<td>No significant comprehension difference among groups</td>
</tr>
<tr>
<td>Smith (2012)</td>
<td>RQ: What is the relationship between the linguistic category of the recast target and noticing?</td>
<td>N=18, L1=varied, Level: not stated</td>
<td>Writing summary of a story, correction of incorrect sentences</td>
<td>Noticing events during stimulated recall</td>
<td>Probability of noticing syntactic recasts and semantic recasts were significantly higher than for morphological recasts</td>
</tr>
<tr>
<td>Uggen (2012)</td>
<td>RQ3: Does the past hypothetical conditional, as a morphologically more complex, and perhaps more difficult, structure promote greater noticing compared to the present hypothetical conditional, a morphologically less complex, and perhaps less difficult, structure?</td>
<td>N=30, L1=varied, Level: 5.7 average years of English</td>
<td>Written essay to elicit target form</td>
<td>Reading aloud and underlying activity, 2nd written essay, stimulated recall</td>
<td>Grammatical complexity positively affected noticing</td>
</tr>
<tr>
<td>Yilmaz (2012)</td>
<td>RQ4: Does degree of salience of the target on</td>
<td>N=48, L1=English</td>
<td>Picture description</td>
<td>OP test, comprehension</td>
<td>RQ4: effect of salience for all</td>
</tr>
<tr>
<td>Study</td>
<td>RQ2: What is the relationship between recast features and eye gaze?</td>
<td>N=16 L1=English Level: low-intermediate</td>
<td>Spanish and German targets varied and were divided into categories: lexical, agreement, tense, spelling, other</td>
<td>SCMC writing conference with NS providing recasts</td>
<td>participant had to decide if their non-target-like SCMC utterances were correct or not and rewrite</td>
</tr>
</tbody>
</table>
One of the four attentional/processing/comprehension studies that compared two different types of linguistic items was VanPatten (1990), who aimed to determine whether comprehension would be negatively affected if L2 learners were required to focus on form in input. A total of 202 L1 English students of Spanish listened to two informative reports in Spanish under one of four input processing conditions: processing for meaning, for meaning and a key lexical item (inflación, “inflation”), for meaning and a bound morpheme (-n, third person plural marker of present indicative), and for meaning and an unbound morpheme (la, female singular definite article). VanPatten stated that his motivation in selecting the two grammatical morphemes was based on their different structural properties. Participants were instructed to mark all instances of the target assigned to their condition; this served as the operationalization of noticing. Results revealed that conscious attention to grammatical forms had negative effects on text comprehension, however conscious attention to lexical items did not significantly affect comprehension. Overall, VanPatten’s results showed that it is difficult for L2 learners to process input for meaning while consciously attending to morphological forms.

In a replication of VanPatten (1990), Greenslade et al. (1999) intended to determine if the results from VanPatten (1990) could be generalized to the written mode. Using the same materials and procedures with the exception of the mode of presentation of the experimental passage, Greenslade et al.’s participant sample included 53 adult learners of L2 Spanish. Results showed that conscious attention to grammatical forms had a negative effect on text comprehension, but conscious attention to lexical items did not have this same effect. Interestingly, results also revealed that the two grammatical targets were not equal: while the comprehension scores of the la grammatical group were significantly lower than those of the lexical inflación group, there were no significant differences in comprehension between the
lexical group and the \textit{–n} grammatical group. The authors explain this difference by referring to the fact that \textit{–n} does not stand alone and may have been processed together with the semantic information of the verb, while \textit{la} is isolated and has negligible semantic value.

Leow et al. (2008) investigated simultaneous attention to form and meaning of lexical items (\textit{sol, la}), a lexical-morphological item (\textit{lo}) and a verbal morpheme (\textit{–n}) to determine whether attentional resources during input processing are in competition to be allocated to either form or meaning (see “Empirical accounts” of the section “Depth of processing” for more information). Results showed now significant difference in comprehension between processing for meaning while paying attention to the different linguistic items in the input. The authors attribute this finding to possibly modality (written instead of aural) or a lack of depth of processing.

Morgan-Short et al. (2012) replicated Leow et al. (2008) and also corroborated their results (see “Empirical account” from the section of depth of processing for more information). They found that attending to the grammatical or lexical form while reading for meaning did not affect comprehension.

In summary of the type of linguistic item in the attentional/processing/comprehension strands, previous research has shown mixed results for the effect of type of linguistic item. While Leow (1995) and Shook (1994) found that recognition of the grammatical form with the higher communicative value and meaningfulness, respectively, was greater than that of the form with lesser communicative value/meaningfulness, results from Leow et al. (2003) revealed that the more salient linguistic item had significant benefits only terms of the amount of noticing; type of linguistic item did not significantly affect recognition or comprehension. Uggen (2012) also found that grammatical complexity and saliency positively affected noticing. Furthermore,
In Collentine’s (1997) word-arranging task, processing time was significantly longer for the irregular grammatical form in comparison to that of the regular grammatical form, however there were no differences between grammatical forms in terms of accuracy. Regarding types of lexical items, Martínez-Fernández (2008) reported that retention was significantly lower for abstract nouns than for concrete nouns.

In conclusion of the four studies looking at two different types of linguistic item in the attentional/processing/comprehension strand, VanPatten found that conscious attention to grammatical forms negatively affected comprehension, however attention to the lexical item did not have this same effect. Greenslade et al.’s (1999) results mirrored this finding regarding comprehension, but they also found that the comprehension of the two grammatical forms was not equal and hypothesized that the group showing conscious attention to the grammatical form had significantly lower comprehension scores than the group showing conscious attention to the lexical item due to the arguably negligible semantic value of the grammatical form. On the other hand, employing a more robust design that first established that participants were indeed processing for meaning before statistically addressing attentional conditions, Leow et al. (2008) and Morgan-Short et al. (2012) reported no significant differences between experimental groups.

**Summary.**

In summary of some of the theoretical issues related to type of linguistic item, DeKeyser (2005) states that complexity of form, complexity of meaning, and complexity of the form-meaning relationship all play a role in determining grammatical difficulty. Furthermore, salience generally refers to the ease of hearing or perceiving a given structure (Goldschneider & DeKeyser, 2001).
Previous studies in both the corrective feedback strand and the attentional/processing/comprehension strand have investigated the effects of type of linguistic item on several aspects of L2 development, however the great majority of these studies compare two of the same type of linguistic item (e.g., grammatical vs. grammatical, lexical vs. lexical) instead of comparing a grammatical form to lexical items. In the feedback strand, most studies (Ellis, 2007; Leeman, 2003; Ortega & Long, 1997; Yilmaz & Yuksel, 2011) comparing two grammatical forms did find differences between the forms; Yilmaz (2012) was the only study to find no difference at all between two grammatical forms. There were also a handful of feedback studies comparing the linguistic category of the focus of corrective feedback: Mackey et al. (2000) reported that participant morphosyntactic scores were significantly lower than lexical scores and in the same grain, results from Smith (2012) revealed that participants noticed syntactic and semantic recasts significantly more than morphological recasts, as inferred from stimulated recalls. However, Smith and Renaud (2013) found that the linguistic target of the recast did not affect learner eye fixation. Finally, Nabei & Swain (2002) claimed that the linguistic elements of the feedback influenced the effect of recasts, however caution must be exercised when extrapolating these results because it was a case study of one L2 learner.

As discussed earlier, regarding the attentional/processing/comprehension strand, there seem to be mixed results for the effect of type of linguistic item on L2 comprehension and development. VanPatten (1990) found that conscious attention to grammatical forms, not lexical item, negatively affected comprehension and results from Greenslade et al. (1999) supported this. On the contrary, both Leow et al. (2008) and Morgan-Short et al. (2012) reported no difference between attentional conditions. Although Leow (1995) and Shook (1994) found that recognition of the grammatical form with the higher communicative value was greater than that
of the form with lower communicative value, Leow (1993) found no such results in the written mode. Also, Leow et al. (2003) reported that type of linguistic item did not significantly affect recognition or comprehension, but Uggen (2012) found that that grammatical complexity positively affected noticing. In the only study looking at two types of lexical items, Martínez-Fernández reported that retention was higher for concrete than for abstract nouns.

Given the number of studies in both the feedback strand and the attentional/processing/comprehension strands that have shown a difference between different combinations of linguistic items, it seems worthwhile to further investigate type of linguistic item as a variable in SLA. Furthermore, both the lack of studies comparing grammatical vs. lexical items and the interesting yet conflicting results of those few studies that do make this comparison (Greenslade et al., 1999; Leow et al., 2008; Mackey et al., 2000; Morgan-Short et al., 2012; Smith, 2012; VanPatten, 1990) motivate future studies to incorporate grammatical forms as well as linguistic items as targets in order to shed light on differences between the two. In addition, since only a handful of the previous studies of type of linguistic item looked at recognition, it would be of interest to further investigate differences between the recognition of a grammatical form and lexical items.

**Methodological Approaches to Operationalizing and Measuring Attention.**

Attention and/or noticing can either be measured online (concurrently) or offline (retrospectively). Logically, it makes more sense to measure this construct where it occurs (at the stage of construction) instead of after it occurs (at the stage of reconstruction). Therefore, the trend within the field of SLA is starting to focus on collecting concurrent data, or triangulating data by collecting both concurrent and retrospective data. The following sections will discuss two popular techniques for measuring attention and/or noticing: online verbal reports and eye-
tracking.

**Concurrent vs. retrospective verbal reports.**

The use of concurrent verbal reports in SLA has led to a wealth of new information regarding the roles of attention, noticing, and awareness in L2 learning and therefore it is important to have a clear idea of what they entail. At the same time, it is also necessary to understand why the use of retrospective verbal reports may not always be the best methodological choice.

Verbal reports are one way to collect introspective data on the thought processes involved in performing a task. Corder (1973) states that only introspective measures allow a glimpse of what learners think about their own production. This, in his opinion, is crucial in order to understand how language works. Not all researchers, however, concur with him regarding the use of verbal reports. Selinker (1972) stresses that only observable data like a learner’s output should be used to formulate theories and conduct research about SLA. Given the difficulty in determining the reasoning behind the L2 learner’s production, such reasoning often has to be inferred in the absence of verbal reports. Gass and Mackey (2000) warn against making such inferences because production phenomena can often be explained in several different ways. Verbal reports allow researchers to not only avoid making inferences but also access conscious cognitive processes that are inaccessible by other means. Furthermore, Schmidt (2001) supports the use of verbal reports in the study of attention by stating that “the clearest evidence that something has exceeded the subjective threshold and been consciously perceived or noticed is concurrent verbal report” (p. 20).

Ericsson and Simon (1984, 1993) in the field of cognitive science classified verbal reports as either concurrent or retrospective. Whereas concurrent verbal reports (“think-alouds”)...
are produced by the participant in real time during the execution of the task, retrospective verbal reports are those in which participants verbalize at some time after the completion of the task in question. Ericsson and Simon recommend that concurrent (online) verbal reports be used whenever possible due to veridicality. Bowles (2010) explains that non-veridicality has a greater potential to occur with the use of retrospective verbal reports: participants’ thought processes may not be accurately reflected because they cannot recall what they were thinking while completing the task. Essentially, this involves attempting to access information at two different times and the possibility of memory decay. It is important to point out that although Ericsson and Simon’s (1984, 1993) work with verbal reports and fabrication did not focus exclusively on the use of verbal reports to investigate language acquisition, their efforts have resulted in the use of verbal reports as standard procedure in fields ranging from accounting (e.g. M. Anderson, 1985) to SLA (e.g., Leow 2001a, 2001b, Leow, Hsieh, & Moreno, 2008; Rosa & Leow, 2004a, 2004b).

Although their strand of work did not focus exclusively on the use of verbal reports to investigate language acquisition, Ericsson and Simon (1984, 1993) in the field of cognitive psychology also proposed a further classification of think-aloud verbalizations into Type 1, Type 2, and Type 3. Type 1 verbalizations are “non-metalinguistic” while they classified Types 2 and 3 as “metalinguistic verbalizations.” While non-metalinguistic verbalizations are simply thoughts, metalinguistic verbalizations are characterized by the requirement of explanations and justifications. Previous SLA research (Bowles, 2008; Bowles & Leow, 2005) has also explicitly used this classification. Ericsson and Simon (1993) predicted that non-metacognitive verbalizations of thoughts will be, for the most part, non-reactive. In other words, they will accurately reflect the nature of cognitive processes but may slightly slow processing. These
hypothetically non-reactive verbalizations are referred to as Type 1 verbalizations. On the other hand, metacognitive verbalizations are termed “Type 2” and “Type 3” and provide justifications or additional specific information regarding the thought processes. Ericsson and Simon admit that these may show more reactivity: they have the potential to slow cognitive processing and also cause changes in the processing.

**Empirical studies using concurrent verbal reports.**

The first published mention of verbal protocol use in SLA to address the process of noticing was Alanen (1995), who combined verbal reports with two retrospective measures. The attention and awareness strand of SLA (e.g., Leow, 1997a, 1998a, 1998b, 2000, 2001a, 2001b; Rosa & Leow, 2004a, 2004b; Rosa & O’Neill, 1999; Sachs & Suh, 2007) gained momentum soon after and has depended heavily on concurrent verbal reports as the main methodological tool to measure learners’ awareness of and attention to certain features of the L2 input. Considering that verbalizations from these concurrent verbal reports are able to show the learner’s cognitive processes while interacting with the L2 (cf. Bowles, 2010 for a review of the issue of reactivity; Leow, Grey, Marijuan, & Moorman, 2014), the verbalizations have then been used to evidence different levels of several cognitive concepts such as levels of awareness (Leow, 2001a, 2001b; Rosa, 1999; Rosa & Leow, 2004a, 2004b; Rosa & O’Neill, 1999; Schmidt, 2001), levels or depth of processing (e.g., Leow et al., 2008; Morgan-Short et al., 2012), attention (e.g., Leow, 1998b, 2001a), activation of prior knowledge (e.g., Leow, 1998), and so on.

Although the majority of studies using concurrent verbal reports have used non-metacognitive think-alouds, a few have employed metacognitive think-alouds. Bowles and Leow (2005) compared the comprehension and item/system learning of a metacognitive group
and non-metacognitive group, finding that only metacognitive verbalization caused a significant decrease in text comprehension but not for production. In Bowles’ (2008) study of the effects of completing non-metacognitive and metacognitive verbal reports while performing an L2 problem-solving task, results showed that the metacognitive group took significantly more time and their ability to produce exemplars of the target structure was hindered. Yoshida (2008) had participants in a think-aloud produce only metacognitive think alouds while reading and performing a recall task; the think-aloud group performed similarly to the silent group.

*The issue of reactivity.*

Although several strands of SLA frequently employ concurrent verbal reports, they have been at the center of debate regarding the possibility of reactivity accompanying them. Considering that a concurrent verbal report involves participants thinking their thoughts aloud while completing a task, the issue of reactivity describes whether the act of thinking aloud alters the cognitive processes involved in performing the task. If it does alter cognitive processes, then this act may not be a true reflection of normal cognitive processing (Bowles, 2010). Some researchers appear to support this argument of the reactivity of concurrent verbal reports, stating that they are not a pure measure of thought processes (e.g., Godfroid et al, 2010; Ellis, 2001; Jourdenais, 2001).

This discussion of reactivity is comprised of a review of SLA studies investigating reactivity, a general summary of reactivity using Bowles’ (2010) meta-analysis, and then summaries of reactivity specifically in studies measuring intake.

In the first SLA study to empirically examine the issue of reactivity in concurrent verbal reports, Leow and Morgan Short (2004) investigated the effects of think-alouds on text comprehension, intake, and written production of the formal imperative in Spanish. Participants
included 77 beginning learners of Spanish in either a control condition or a think-aloud condition. The think-aloud group produced non-metacognitive think-alouds while reading and completing a comprehension task, a fill-in-the-blank task, and a multiple-choice recognition task. Given that the control and think-aloud groups did not differ significantly on post-test assessments, the authors concluded that reactivity did not play a significant role in performance. Time on task was not measured, so no latency information is available. However, the authors did recommend that research designs employing think aloud procedures include a non-TA group to address the issue of potential reactivity.

Bowles and Leow (2005) expanded on Leow and Morgan-Short (2004) by examining two types of verbalizations, metalinguistic and nonmetalinguistic, in terms of their impact on learner performance. The participant group consisted of 45 fifth-semester Spanish learners who were randomly assigned to either a control group or one of the two experimental conditions. All participants read a passage and then completed the same tasks as in Leow and Morgan-Short (2004). Reactivity was not found in comparisons between groups; however, metalinguistic verbalization seemed to cause a significant decrease in comprehension in comparison with the non-metalinguistic group.

Rossomondo (2007) investigated the role of lexical temporal indicators in the incidental acquisition of the Spanish future tense. First-semester students of Spanish comprised a silent group and a non-metacognitive think-aloud group. After a sample TA and a warm-up, participants read the passage and either said their thoughts aloud or were silent, depending on their respective group. Participants then performed a multiple-choice comprehension test in English, after which they did either a form-recognition task or a form production task. No reaction time data was collected. Results showed reactivity on the recognition and production
tests; more specifically, reactivity facilitated performance. Participants thinking aloud did not have significantly different comprehension scores from the silent group.

Sachs and Suh (2007) found no reactivity for think-alouds. A total of 30 ESL learners were divided into a think-aloud group and a non-think-aloud group. All tasks took place between individual participants and the researcher via computer-mediated communication (CMC). A set of pretests was followed by a story-retelling task either with or without textually-enhanced recasts, and with or without the requirement to think aloud. Then participants completed an interactive story-retelling post-test and a text completion post-test. Results showed that the think-aloud group and the non-think-aloud group had no significant differences in the amount of time spent on take, thus indicating non-reactivity for latency. The authors concluded that there was no reactivity in the data.

Sachs and Polio (2007), however, reported mixed results in their investigation of the reactivity of think-alouds performed in the L2 on a three-stage L2 writing task. In an initial experiment, results showed that when participants were silent during the comparison stage, they revised more errors than when they thought aloud and thus evidenced negative reactivity. A second experiment also focused on the issue of reactivity and compared two reformulations groups (with or without think-alouds) in accuracy in an L2 writing revision task. Results showed no significant difference between the silent and think-aloud group. The conflicting results of the first and second experiments may be due to the different experimental designs; however it is still possible that the think-alouds may have a negative effect on learner performance. In addition, this is the only study that required participants to think aloud in the L2, which could have played a role in the results.

Bowles (2008) took the reactivity strand one step further by examining the effects of
performing non-metacognitive and metacognitive think-alouds during completion of an L2 problem-solving task on subsequent written production of the target form. Participants included 194 first-semester Spanish students who were randomly assigned to one of six experimental groups that differed in type of verbalization (metacognitive, non-metacognitive, silent) and type of feedback (implicit or explicit). Results showed that the metacognitive think-aloud group spent significantly more time on task than the non-metacognitive verbalization group and the silent group. The metacognitive verbalization group also had a significantly decreased ability to produce the target structure; the same did not hold for production of novel exemplars. Overall, results indicate that non-metacognitive verbalization was non-reactive, thus supporting Ericsson and Simon’s (1984, 1993) model.

In two separate experiments, Sanz, Lin, Lado, Bowden, and Stafford (2009) examined the reactivity of concurrent verbal protocols in an L2 instructional lesson. In Experiment 1, a total of 24 L1 English adult learners of Spanish were randomly assigned to either a metacognitive think-aloud group (n=11) or a silent group (n=13). The treatment task was a computerized lesson on the Latin case system. Post-test assessments consisted of an aural interpretation test, a written grammaticality judgment test (GJT), and a sentence production test. Sanz et al. used a very precise measurement on latency that involved measuring mean reaction times on correct responses from all pre- and posttest and calculating a grand mean. Results showed main effects only for time, thus indicating that both the think-aloud and the silent group learned as a result of the lesson, and that the verbalizations on the think-aloud were neither facilitative nor detrimental to performance. The think-aloud group did have longer reaction times on just one of the three pretests, the GJT, than did the silent group. Therefore, the authors reported mixed results for latency. The second experiment used the same number of participants per group and the same
pre- and post-tests, however a less explicit version of the treatment was employed. Instead of
being exposed to an explicit grammar lesson on the target form like in Experiment 1, Experiment
2 had participants rely on knowledge they gained from task-essential practice and explicit
feedback. Similar to Experiment 1, results showed that the treatment caused participants to
improve their interpretation and production abilities with regard to the target form. Unlike
Experiment 1, reactivity facilitated performance in Experiment 2: participants in the think-aloud
group performed significantly better than those in the silent group on the GJT and the production
test. No latency effects were found. Overall, the authors concluded that reactivity depends on
not only the task and assessment tool but also whether the dependent variable is accuracy or
latency. Although another study (Stafford et al., 2012) that employed a similar research design
reports contradictory findings with respect to Sanz et al.’s (1999) reactive findings, it is
important to note that the two participant samples were very different.10

The issue of reactivity is best addressed using Bowles’ (2010) meta-analysis. She
compiled data from several reactivity studies and one unpublished empirical study (Polio &
Wang, unpublished11) in the field of SLA to determine if concurrent verbal reports are reactive
for accuracy and/or latency when used in conjunction with verbal tasks. All the studies
described above except Stafford et al. (2012) together provide the data for this meta-analysis.
Her investigation summarized that overall, thinking aloud while performing a verbal task has
only a small effect on post-task performance; those thinking aloud perform only slightly better or
slightly worse. The results for time on task, however, were more decisive: thinking aloud

---

10 Note 3 of Stafford et al. (2012) explained that 26 participants thought aloud during the treatment,
however these data were not included. A series of ANOVAs showed no significant main effect for
verbalization and no significant interaction with Time and/or Treatment.

11 Bowles (2010) included the unpublished paper of Polio and Wang (unpublished) in an effort to be as
inclusive as possible while still abiding by her selection criteria (see Bowles, 2010, p. 80); Bowles also
followed the methods described in Rosenthal (1994) for determining inclusion protocol.
increases time on task. With effect sizes ranging from small \((d=0.16)\) to very large \((d=1.16)\), the largest latency effects occurred when participants thought aloud while performing reading tasks. In addition, because it was not possible to get the effect sizes of all studies due to a violation of the assumption of homogeneity, there is no definitive answer regarding the question of reactivity and concurrent verbal reports. Bowles further explains that while subsequent analyses performed aiming to identify the sources of between-study variance showed that some could be attributed to the type of report (metacognitive vs. non-metacognitive), other sources of variance may be due to the unique research context. Most notably, L2 proficiency level and explicitness of instruction accounted for a significant amount of variance between studies.

Considering that Bowles (2010) concluded that reactivity depends on a host of variables, it is also helpful to examine reactivity in light of intake, one of the variables discussed in the current work. Four studies (Leow & Morgan-Short, 2004; Sachs & Suh, 2007; Rossomondo, 2007; Sanz et al., 2009) of those summarized above were used to determine the effects of type of task on receptive form learning in Bowles’ more specific meta-analysis. But because the meta-analysis showed that the assumption of homogeneity of distribution was violated \((Q=11.44, df=3, p<.05)\), Sanz et al.’s (2009) study with learners of beginning proficiency was removed from the data analysis. This resulted in a homogeneous distribution among the three remaining studies, all of whose participants were of beginner or intermediate proficiency. Overall, Bowles (2010) found that thinking aloud during tasks centered on a reading passage seems to facilitate receptive form learning such as measures of intake.

Only two studies (Bowles & Leow, 2005; Leow & Morgan-Short, 2004) from Bowles’ (2010) meta-analysis measured written production scores of think-aloud groups and compared them to those of silent groups. Results from the meta-analysis showed that the effect size
distributions from the two studies met the assumption of homogeneity; effect size was $d = -.14$. Since all confidence intervals were in the negative range, results show that thinking aloud while completing a written production task has a small, negative effect on productive form learning when compared to performing the task silently.

Regarding comprehension, Bowles’ (2010) meta-analysis analyzed data from the six studies that included a reading with a measure of text comprehension (Bowles, 2008; Bowles & Leow, 2005; Leow & Morgan-Short, 2004; Polio & Wang, in review; Rossomondo, 2007). The assumption of homogeneity of distribution was met and the effect size was $d = .04$. Since the 95 per cent confidence interval ranged from -.23 to .31, results revealed that thinking aloud during a reading task slightly improves comprehension of the text when compared to performing the task silently. Bowles states that additional research is required regarding the effects of type of task on test comprehension due to the positive and negative values of the confidence interval.

Furthermore, several studies (Bowles, 2008; Bowles & Leow, 2005; Leow & Morgan-Short, 2004) have shown that concurrent verbal protocols did not significantly affect L2 learners’ processing in reading tasks (except in latency). More recently, however, Bowles (2010) reviewed four studies in which the task involved a reading passage. She reported that two of these studies showed nonreactivity (Bowles & Leow, 2005; Leow & Morgan-Short, 2004) while one showed positive reactivity (Rossomondo, 2007). An exploratory meta-analysis of L2 tasks and post-task assessments resulted in Bowles (2010) reporting that nonmetalinguistic TAs have a small effect ($d = 0.21$) on post-task measures of comprehension. This effect size, however, is not considered reliable due to the inclusion of zero in the range of the 95% confidence intervals. Bowles concluded overall that that “thinking aloud while completing a verbal task has a small effect on post-task performance” (2010, p. 110).
Following the publication of Bowles’ (2010) meta-analysis, a study by Goo (2010) investigated how individual differences in working memory capacity (WMC) mediate the potential occurrence of reactive effects on reading comprehension and rule learning for L2 learners of Spanish. Participants were 42 adult learners of first-semester Spanish who completed a reading task containing 20 instances of the target linguistic form. There was a think-aloud group and a non-think-aloud group. Comprehension was tested by short-answer and multiple-choice questions; written production was measured by a fill-in-the-blank test. WMC was measured via listening and operation span tasks in English. Results indicated that WMC appeared to have mediated reactive effects of think-alouds on learner performance. The TA condition showed a trend toward negative reactivity; Goo stated that there were “more or less reactive effects on learner comprehension” (Goo, 2010, 739). More specifically, Goo reported that the concurrent verbalization requirement may have affected only the high-WMC learners.

Although the topic of a study by Stafford et al. (2012) was not reactivity, but rather explicit instruction and explicit feedback, they were still able to report results regarding reactivity. In their study, 65 Spanish-English bilinguals completed input-based, task-essential practice with interpreting agent/patient roles and also received either more or less explicit feedback. There were two interpretation tests (one written and one aural), a grammaticality judgment test, and a production test. Of the 65 participants, 26 performed concurrent verbal reports during the treatment to rule out reactivity. Results for these 26 participants showed no significant main effect for verbalization (thinking aloud) and no significant interaction with Time or Treatment, however given the small number of participants, these results need to be interpreted with caution.
A study by Morgan-Short, Heil, Botero-Moriarty, and Ebert (2012) addressing simultaneous attention to form and meaning in L2 written input also provided information on the reactivity of concurrent verbal reports. A pool of 361 third-semester L2 Spanish learners attended to either lexical or grammatical forms while reading for meaning alone or reading for meaning while attending to either lexical or grammatical forms. Approximately half of the participants performed concurrent verbal reports, which were then coded for depth of processing by counting the instances of circling and/or making certain comments on the target form. The reading was followed by a multiple-choice comprehension test. Results revealed that thinking aloud had a statistically significant reactive effect for the experimental group that read for comprehension while paying attention to a particular form in the passage. In other words, negative reactivity was found for comprehension. But because the effect size was minimal and having thought aloud did not alter the statistical pattern of results between groups in terms of the effect of Condition, the authors conclude that use of think-alouds did not appear to compromise the internal validity of results related to simultaneous attention to form and meaning (cf. Leow et al., 2008).

In summary, Bowles’ (2010) meta-analysis showed that overall, thinking aloud while performing a verbal task has only a small effect on post-task performance, meaning that those thinking aloud performed only slightly better or slightly worse. More specifically, thinking aloud during a reading passage appears to facilitate both comprehension and receptive form learning like intake; however thinking aloud while performing a written production tasks seems to have a small, detrimental effect on productive form learning in comparison to completing it silently. Although Morgan-Short et al. (2012) did report negative reactivity for comprehension,
a low effect size and failure to find other statistical patterns among groups annulled this finding. Additionally, Goo (2010) found a trend toward negative reactivity regarding comprehension.

Regarding latency, Sachs and Suh (2007) and Sanz et al.’s (2009) Experiment 2 found non-reactivity for latency, in contrast to Bowles and Leow (2005). Sanz et al.’s (2009) Experiment 1 found mixed results for reactivity for latency. However, it is important to remember that latency was defined differently in Sanz et al., thus perhaps explaining some of the contradictory findings across studies. On the other hand, Leow and Morgan-Short (2004) and Rossomondo (2007) did not measure reaction time and therefore reactivity of latency could not be determined.

In conclusion to the discussion on reactivity, concurrent verbal reports afford the researcher a unique glimpse of online processing and, more particularly, processes and therefore should be a staple component of SLA studies that involve constructs such as depth of processing and awareness in the attentional strand of SLA, even if the data is only used as a support that participants were in fact complying with instructions. Several studies (e.g., Gass et al., 2003; Greenslade, Bowden, & Sanz, 1999; Leow, 1993; Shook, 1994; VanPatten, 1990; Wong, 2001) failed to measure concurrent processing by means of think-aloud protocols or eye-tracking. Without these online data, it can only be assumed that participants were performing as indicated in the instructions. Furthermore, Chaudron (1985) recommended that future research regarding intake adopt more standardized measures that contrast the degree and type of processing; given that concurrent verbal reports can arguably contrast degree and type of processing through different coding scheme, concurrent verbal reports along with eye-tracking may be helpful measures in the future (Leow et al., 2014). Regarding reactivity, Leow et al. (2014), following Leow and Morgan-Short (2004), suggest that all studies using concurrent verbal reports also
include a silent group to rule of the possibility of reactivity.

**Coding noticing, awareness, and similar constructs.**

Studies using concurrent verbal reports code these reports in order to obtain data on constructs such as noticing and awareness (in addition to depth of processing, which will be discussed in a later section). Past studies have used several similar coding criteria; however each differs in important ways. Details on coding are divided into coding of grammatical items and coding of lexical items.

In coding for awareness of a target grammatical item (third person preterit of irregular –ir verbs in Spanish), Leow (1997) established three levels of awareness: awareness at the level of understanding (+ cognitive behavior, + meta-awareness, + morphological rule), awareness at the level of reporting (+ cognitive behavior, + meta-awareness, - morphological rule), and awareness at the level of noticing (+ cognitive behavior, + meta-awareness, + morphological rule). While awareness at the level of understanding involved a high level of processing evidenced by hypothesis testing and rule formation, awareness at the level of reporting was characterized by a medium level of processing and some cognitive effort. Lastly, awareness at the level of noticing involved a low level of processing and minimal cognitive effort (cf. Leow, 2012).

Leow (2000) used concurrent verbal reports to establish the presence or absence of awareness of the irregular third-person singular and plural preterit forms of selected stem-changing –ir verbs in Spanish. Any participant who provided a report of being aware of the targeted forms or who provided some form of metalinguistic description of the underlying rule was assigned to the aware group. Participants who did not meet the criterion were determined to be unaware.

reports for levels of awareness of contrary-to-fact conditional sentences in Spanish. Awareness at the level of noticing was based on Schmidt’s (1990) noticing hypothesis and was operationalized as a verbal reference to the target structure without any mention of rules. The learner had to give clear indication that they were paying focal attention to the target form by either reading the target and pausing or making a comment on it, or explicitly referring to the target without reading it. Awareness at the level of understanding was operationalized as the verbal expression of an explicit rule of the target. Also, the participant had to explicitly make a connection between present or past time frames and the respective verb form.

Both Leow (2001b) and Leow et al. (2003) operationalized noticing of the formal imperative command in Spanish as any translation (correct or incorrect) of targeted verbs, circled targeted verbs, verbal references to and/or written remarks related to targeted verbs in the input as revealed in their think-aloud protocols and on participants’ versions of the text. Likewise, Hama and Leow (2010) operationalized noticing as a verbal comment regarding some aspect of animacy of the target form (artificial determiners). Participants were coded as understanding if they mentioned correct animacy-related rules.

Bowles (2004) operationalized noticing as any translation of the target form (Spanish formal imperative), comment about the target form, or marks on participants’ versions of the text relating to the target verbs. Simply reading a target form aloud in Spanish did not count as noticing because the act of decoding a word does not necessarily imply awareness.

The only published study that coded lexical items in concurrent verbal reports is Martínez-Fernández (2008). Following Bowles (2004), Martínez-Fernández operationalized noticing as reading glosses aloud or making a comment about the target item.

Critically, one similarity among the reviewed coding criteria is the required absence of
any rule verbalization in order for a verbalization to be coded as noticing (Leow, 1997; Rosa & O’Neill, 1999; Rosa & Leow, 2004). It is important to note that Leow’s (2000) coding of awareness conflates the three levels differentiated in other studies (Leow, 1997; Rosa & O’Neill, 1999; Rosa & Leow, 2004). Additionally, Leow (2001a), Leow et al. (2003), and Martínez-Fernández (2008) all operationalized noticing as making a comment about the target item; Martínez-Fernández added to this operationalization the act of reading the glosses of the target lexical items aloud.

Overall, it appears that coding for awareness and noticing in past studies greatly overlap, however at this same time this highlights the need for fine-grained coding criteria in future studies.

Eye-tracking as a measure of processes related to the attentional strand.

Prior research (Carpenter & Just, 1976; Rayner, 1998) has confirmed the link between the eye and the mind: eye movements are directly related to underlying cognitive processes. In a theoretical account of the eye fixation sequence and duration during cognitive tasks, Carpenter and Just reported that the eyes commonly fixate on the external referent as its corresponding internal representation is processed\(^\text{12}\). Eye-tracking is the process of measuring the point of gaze of the eye and the motion of the eye in relation to the head and is a useful measure of the allocation of attention that fits in well with the selective attention feature of current computational models (Rehder & Hoffman, 2005a). It is an uncontroversial measure of overt attention (Blair, Watson, Walshe, & Maj, 2009) and furthermore, several studies have reported a close link between covert attentional processes and eye movement (see Godfroid et al., 2013, Rayner, 1998, and Wright & Ward, 2008 for review). Hyrskykari (2006) affirms that “the point

\(^{12}\) They used gaze duration to operationalize processing time and gaze locus to indicate which item is being processed at any moment.
of gaze reveals the focus of the user’s visual attention” (p. 660). The majority of researchers agree that eye-tracking data are a reliable and useful tool for measuring “moment-to-moment processing of individual words and larger segments of text” (Starr & Rayner, 2001, p. 156). Although under some circumstances attention can dissociate from eye gaze (Posner, 1980), changes in attention are usually followed by the corresponding eye movements (e.g., Kowler, Anderson, Dosher, & Blaser, 1995). Evidence exists showing that even the simplest stimuli cause a dissociation between attention and eye movements (Deubel & Schneider, 1996). Furthermore, Kahneman (1973) posited that eye-tracking can be used to measure cognitive effort by means of intensity and time, both of which can be captured by eye-fixation location and eye-movement time (refer to the section entitled Definitions of Terms for definitions of all eye-tracking measures).

There are several ways to measure processing time via eye-tracking, however none is without controversy. Rayner (1998) explains that if the research interest is analyzing a unit larger than a word, then the total first-pass fixation time (referred to as ‘gaze duration’ in the present study) of the unit of interest is typically used as the main measurement. Second pass time can also be of interest in instances in which a reader enters a text region, exits that region, and then regresses back to it (Rayner & Sereno, 1994). The problem of multiple fixations on one unit of analysis motivates the use of gaze duration as an eye-tracking measure; it represents the sum of all fixations on a word during the first pass before the reader performs a saccade to the next word (does not include second pass fixations). This measure, along with first fixation duration (duration of the first eye fixation in an interest area when that area is first encountered during forward reading), is one of the most frequently-used measures when the unit of analysis is the word (Rayner, 1998). Regarding these two measures, Rayner and Pollatsek (1987)
suggested that slower cognitive processes are more likely to affect gaze duration, while faster cognitive processes affect first fixation duration. Yang and McConkie (2005) argue that gaze duration is the measure best suited for capturing processing difficulties. Rayner (1998) highlights this controversy that surrounds the most appropriate measure of processing time, but also suggests that researchers examine many measures in one study in order to reveal a closer look at the reading process (Rayner, Sereno, et al., 1989; Schmauder, 1992).

The eye-tracking data from readers can conceivably show the online ease or difficulty of processing in addition to highlighting “the processes that underlie incremental integration of words and phrases in developing sentence representation” (Witzel et al., 2012, p. 3). Yet another advantage of eye-tracking is its naturalness in comparison with other tasks that have readers read in ways that are never actually done in everyday life (like self-paced reading). This technique does little to interfere with natural reading; it even allows participants to employ the reading strategies that they typically use, such as stopping their reading to regress to a prior part of the sentence to re-inspect content.

The invisible boundary paradigm is a technique used in eye-tracking that enables manipulation of the characteristics of an upcoming word. In other words, an invisible interest area is created around a target word; once the reader crosses the boundary into the interest area, then the word either becomes visible or is masked for the duration of the fixation on the interest area. Rayner (1975) was among the first to employ this technique in eye movement studies; he found that when a reader fixated 7-12 character spaces prior to a word, the reader was only able to pick up gross visual characteristics like word shape.

The invisible boundary paradigm is useful in creating a condition of peripheral attention: since when participants fixate on something during reading, they pay focal attention to it, if the
invisible boundary paradigm does in fact impede fixation on a certain word, then that word can only be attended to peripherally.

The relationship between language processing and eye behavior takes two main theoretical directions: the direct measurement position and the oculomotor control principle, each of which posit varying degrees of cognitive influence.

Following the cognitive tradition, the direct measurement position maintains that the fixation time on a word is a good measure of the latency for internal cognitive processes associated with that word (Yang, 2006). This position includes three assumptions. First, the majority of saccades in reading are triggered by a cognitive event; saccade latency thus reflects the processing time necessary for the event (Just & Carpenter, 1980). Second, gaze usually shifts from one processed word to the next and aims at an optimal part of the word, as posited by the word-based targeting assumption (Reichle, Rayner, & Pollatsek, 1999). Third, a reader attends to and processes text in a sequential manner; attention is shifted to change the focus of text processing (Morrison, 1984). Whereas some new models in oculomotor research complement the direct measurement position, others deviate from its main assumptions. The following is a summary of some of these models.

The E-Z Reader (Reichle, Liversedge, Drieghe, Blythe, Joseph, White, & Rayner, 2013; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Pollatsek, & Rayner, 2006; Reichle, Rayner, & Pollatsek, 1999; Reichle, Warren, & McConnell, 2009) is a processing model that accounts for how word identification, visual processing, attention, and oculomotor control jointly influence when and where the eyes move during processing. It posits a separation between attention shifts and saccade programming in an attempt to account for the frequency distributions of fixation durations and saccade lengths. This model consists of four processes: a familiarity
check, lexical access completion, saccade programming, and the actual saccade. It posits that attention during reading is like a spotlight that shifts serially from word to word; attention shifts to the next word when lexical access of the currently fixated word is complete. Overall, the E-Z Reader model serves as a unified theory that explains why various aspects of eye-movement control operate certain ways, and therefore assist in understanding how both oculomotor and linguistic variables affect eye movement control during reading (Reichle, Rayner, & Pollatsek, 1999).

Reichle, Warren, and McConnell (2009) address a major limitation of the E-Z Reader model, namely, that the E-Z Reader model (Pollatsek et al., 2006) does not consider the issue of how eye movements are affected by higher level, postlexical language processing. By means of a series of simulations based on simple assumptions, they show that the E-Z reader can, in fact, explain how the systems responsible for higher level, postlexical language processing may work in conjunction with those in charge of various functions such as shifting attention and identifying words in order to produce the eye movement patterns observed during the act of reading. More specifically, the authors posit that postlexical processing can lag behind ongoing lexical processing and thus leave no visible effects on eye progression through the text. However, the difficulty that is inherent in postlexical processing will occasionally show itself very quickly in the form of pauses and regressions back to the text area where processing difficulty first occurred. They also postulate that postlexical processing difficulty may occur in the event that the processing of a word is not completed before the next word has been identified.

Compatible with the direct measurement position, the SWIFT model (Engbert, Longin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005; Kliegl & Engbert, 2003) maintains that eye movement triggering depends on a random waiting time function but is regulated by
foveal processing difficulty. Therefore, the fixation duration directly reflects the processing of the fixated word. The SWIFT model, however, assumes simultaneous processing instead of serial processing. The Glenmore model (Reilly & Radach, 2003, 2006) posits parallel processing of word units and disregards the proposed relationship between processing difficulty and the time between two saccades. Instead, the time required to execute a saccade is determined by a random waiting time signal and regulated by a general level of activation that is associated with word processing. The word-based targeting assumption is the only part of the direct measurement position that is preserved.

Feng (2006) proposed the stochastic (random) and hierarchical architecture for reading eye-movement (SHARE) model, which purports that eye movements in reading are observable outcomes of latent stochastic processes. Because of common parameters that affect both, fixation duration and saccade length are allowed to vary. This model can be compatible with the direct measurement position in terms of the serial-processing and word-based targeting assumption.

The competition/interaction (C/I) model posits that saccade initiation in reading is based on internal reading strategies and therefore eye movements are initiated independently of ongoing linguistic processing (McConkie & Yang, 2003; Yang & McConkie, 2001, 2004, 2005). Autonomous triggering signals influence saccades; saccade latency is controlled by internally regulated parameters. Cognitive factors influence eye behavior through processing-related inhibition, parametric adjustment, and direct cognitive control (McConkie & Yang, 2003). When linguistic processes cause processing difficulty, saccade initiation is inhibited and there is a reduction in saccade probability and a shortening of saccade length. The latency of difficulty detection determines the onset time of saccade inhibition. Crucially, direct cognitive control
happens at a longer latency and affects the eye more strongly when the cognitive processing problems are detected and deliberately handled. This model holds that cognitive influence on eye behavior is largely due to processing-related inhibition. Engbert, Nuthmann, Richter and Kliegl (2005) classify the C/I model as a “primary oculomotor model” (p. 778), not a cognitive-control model.

In summary, the referenced models differ foremost in their views of whether a reader attends to and processes text sequentially or simultaneously. While the E-Z Reader model (Reichle et al., 1998) and the SHARE model (Feng, 2006) argue for sequential processing, SWIFT and the C/I Model (McConkie & Yang, 2003; Yang & McConkie, 2001, 2004, 2005) assume simultaneous processing. Second, these models do not concur on the word-based targeting assumption (Reichle et al., 1999), as upheld by the direct measurement position, which states that gaze usually shifts from one processed word to the next and aims at an optimal part of the word. The word-based assumption is instantiated in the E-Z Reader model, the Glenmore model (Reilly & Radach, 2003, 2006), and the SHARE model, but is discounted in the C/I model. Finally, the models differ in their views of what triggers saccades in reading and, consequently, what fixation duration reflects. The direct measurement position maintains that saccades are triggered by a cognitive event and latency thus reflecting processing time. Similarly, the SWIFT model proposes that fixation duration may reflect processing and therefore processing difficulty causes delays in saccades; the C/I model also maintains that linguistic processing difficulty causes an inhibition of saccade initiation and a shortening of the saccade length. On the contrary, the Glenmore model posits no relationship between processing difficulty and saccade execution. In conclusion, due to the close relationship posited between linguistic operations and eye behavior during reading by the E-Z Reader (Reichle, Polletsek, &
Rayner, 1999), this is the model that the current investigation follows.

Using data on eye movements during reading, Yang and McConkie (2001) investigated how cognitive processing affects saccadic eye movement. A text read by 36 adults was occasionally replaced by one of six alternate stimulus patterns for the duration of single eye fixations; frequency distributions of the fixation durations were recorded. Results showed two saccade disruption times, thus suggesting three types of saccades: early, normal, and late. Yang and McConkie made several conclusions. First, late saccades are the only ones to be affected by cognitive processes related to the contents of the words seen during the critical fixations. Also, cognitive influences caused saccade delays instead of saccade triggers. The authors conclude that the C/I model seems to best account for eye movement during reading.

In a study of category learning, Rehder and Hoffman (2005a) used eye-tracking to investigate whether participants would limit their attention only to the stimulus dimensions required to classify each structure. Attention was operationalized as the presence of eye fixations on spatially separated stimulus dimensions. They predicted that participants would cease to fixate on dimensions irrelevant to the correct classifications. A total of 72 undergraduate students were randomly assigned to four category structures. Participants completed learning trials in which they classified exemplars according to one of two color categories; trials continued until the participant completed four errorless blocks or after a maximum of 28 blocks. Feedback was given after each block that told participants how close they were to achieving an errorless block. The eye-tracking measures (dependent variables) were number of dimensions fixated, fixation time proportion (time spent fixating each dimension

---

13 There were two saccade disruption periods: an Early Disruption Period and a Late Disruption Period. During the Early Disruption Period, all saccades are cancelled beginning in the 125-150 ms interval due to a blank screen or a homogeneous string of X’s. The Late Disruption Period consists of cancellation of saccades in the X’s and Non-word+ conditions, but beginning about 50 ms later.
divided by the total time spent fixating all dimensions), and the relative priority that captured the ordering of fixations. Results showed that attention to the dimensions was eventually allocated in a way that optimized the participants’ ability to discriminate categories. These findings cross-validate the use of eye-tracking as a measure of attention in category learning.

An innovative experiment by Kuo, Hsu, and Day (2009) used eye-tracking to investigate cognitive effort and the framing effect, which refers to the phenomenon that varying the presentation of the same problem can systematically affect the choice one makes. The aim of the study was to determine if participants would exert more cognitive effort, as measured by a) the processing time per word and/or b) number of fixations per word to process a negatively framed problem than a positively framed problem. The 56 participants were randomly divided into two groups, each of which was exposed to two positively framed problems and two negatively framed problems. Eye movements were tracked and participants had to select an option to the problem. Results showed that the framing effect is more likely to exist as the cognitive effort asymmetry (as measured by differences in processing time per word between the positive and the negative frame) increases. The researchers conclude that eye movement data can be used to measure cognitive effort.

Theeuwes and Van der Stigchel (2009) carried out a series of three experiments in which between six and twelve participants completed exogenous cueing tasks. The experimental design consisted of a training session and an experimental session; the task involved an abrupt onset cue indicating the target location at chance level and participants had to discriminate between the letters E, S, and H. Attention was measured in terms of saccades, saccade latency, and saccade trajectory. Results revealed that participants had slower and less accurate responses to a target presented at a cued rather than an uncued location when there was a delay between the
cue and target; this indicates the occurrence of inhibition-of-return. For the trials which had participants move their eyes to a location in space, results showed that no saccade deviation away from the location was inhibited due to inhibition-of-return, except in the case that participants had to process the target letter at the inhibited location. These findings are consistent with the notion that inhibition-of-return may reduce the input of signals going into the saccade map. The authors conclude that the strength of saccade deviation is an important measure that can reveal how much attentional processing is occurring at any location in time.

In a pilot study, Godfroid et al. (2010) explored what they defined as learners’ noticing of pseudo words in written input via online eye-tracking and stimulated recalls. Nine L1 Dutch speakers learning English read a subset of pseudo words in each of four conditions; each pseudo word was read in only one condition due to the Latin square design. Participants read several texts at their own pace while an eye tracker recorded eye movements. Assessment tasks included a vocabulary-retention posttest and a content posttest. The first administration of a multiple-choice vocabulary retention post-test resulted in ceiling effects. The second was cued-production, resulting in floor effects. The third was a multiple-choice test with 18 options that discriminated between participants in a suitable manner. Sessions ended with a stimulated recall protocol. In regards to the results, Godfroid et al. (2010) showed eye-tracking data from four experimental paragraphs with target items to illustrate that the target word received notably more attention than other words on the screen, as operationalized by first fixation duration, first run dwell time, and overall dwell time. The participants displayed an increased level of cognitive engagement with the target word, which the authors interpret as a noticing event. The combined results of the vocabulary retention posttest versions resulting in the floor and ceiling effects suggest that even a single exposure to a lexical item can lead to some intake. Furthermore,
participants sometimes recognized words on the post-test even though the eye-tracking data show no marked increase in viewing for that target word during the reading. Godfroid et al. explained this by suggesting the possibility of a critical boundary in terms of viewing activity: input below this boundary would be ‘perceived’ and input above it would be ‘noticed.’ In other words, it is possible that the participants who recognized words on the post-test without showing increases in viewing for those targets during the reading did in fact pay attention to them during the reading but at a baseline level (as evidenced by the eye-fixation data). Godfroid et al. (2010) concluded by cautioning that eye-tracking data are quantitative data and therefore do not tap into the quality of cognitive operations that result in extended processing times. Eye fixation durations may be used as an index of depth of processing in the sense that Craik and Lockhart (1972) viewed it as the amount of attention paid to a stimulus; however eye-tracking cannot contribute any information regarding what learners are doing in terms of qualitative operations upon extensively processing a stimulus. Therefore, the triangulation of eye-tracking data with verbal report data may be very revealing.

Godfroid, Boers, and Housen (2013) expanded their 2010 study by aiming to measure instances of focal attention by means of eye-tracking and to assess whether learners' fixation times on novel words are positively associated with subsequent recognition of those words. Attention was operationalized as a quantitative variable reflected in eye fixation times during reading; based on this premise, they evaluated the weaker version of Schmidt's (1994, 1995, 2001) original noticing hypothesis (1990) in which noticing is said to be minimally facilitative of intake. A participant group of 28 L1 Dutch female learners of second and third year college-

\footnote{In describing the ‘weaker’ version of Schmidt’s (1994, 1995, 2001) original noticing hypothesis, Godfroid et al. (2013) opted for a continuous approach to noticing. In other words, instead of a categorical model of intake such as Leow’s (forthcoming) three-level model, Godfroid et al. uphold a positive, linear}
level English read 20 short English paragraphs containing only known words except for the 12 unknown critical words. The eye movement measures collected during the reading included first fixation duration, gaze duration, second pass time, and total fixation time. As in their 2010 study, Godfroid et al. created different conditions (including a control condition) between which eye fixations could be compared. Therefore, Godfroid et al. (2013) operationalized noticing "in relative, rather than absolute terms; that is, as an increase in fixation time relative to a baseline condition." A vocabulary post-test (as piloted in Godfroid, et al. (2010)) to measure recognition followed. Results showed that the pseudo words were fixated significantly longer than the control words. In response to the second research question, there was a significant positive relationship between fixation duration on a novel word and post-test recognition of that word. More specifically, for every one additional second that participants looked at a novel word during the reading, they were 8% more likely to correctly recognize that word on the vocabulary test. Godfroid et al. concluded by stating that eye movement data are quantitative and therefore cannot be used to gain additional information on the cognitive operations that are behind awareness. They did, however, state that they could “safely assume that our participants were consciously perceiving (i.e. visually aware of) the unfamiliar words at which they were looking” (p. 510).

Witzel et al. (2012) aimed to address methodological concerns of eye-tracking by examining the processing of three sentence types containing temporarily ambiguous structural configurations (relative clause attachment ambiguity, adverb attachment ambiguity, and noun phrase versus sentence coordination ambiguity). A total of 32 university students read sentences in four reading tasks. Findings of eye-tracking were compared with findings from three other relationships between attention, operationalized as fixation duration, and word learning (with no discrete stages).
experimental tasks (serial probe recognition task, G-maze task, and L-maze task) that also use eye-tracking in order to determine whether the tasks were able to reveal processing differences among the sentence types. All eye movements were recorded and five measures were used to analyze data: first fixation duration, first-pass reading time (‘gaze duration’), go-past reading time, right-bounded reading time, and total reading time (‘total fixation time’). Results showed that the eye-tracking data from the reading task was able to indicate processing differences at the exact location of temporary ambiguity in two of the three sentence types and delayed effects for the third sentence type. The researchers concluded that the richness of data in the eye-tracking task was able to compensate for the occasional insensitivity to processing differences of some of the five measures.

In a study of noticing of corrective feedback in the form of explicit recasts, Smith (2012) aimed to determine whether eye-tracking technology could measure noticing of this feedback during native speaker (NS) – non-native speaker synchronous computer-mediated communication (SCMC). The participant group was made up of 18 college students learning English as an L2 in the United States; participant L1 was varied. After completing an online English proficiency test during the first session, participants were trained in recasts a few days later. On that same treatment day, they also completed the treatment task, immediate posttest, and performed a stimulated recall. The treatment task consisted of watching a short animation clip and then chatting online with the NS to retell the story. The NS provided full recasts when appropriate; eye-tracking was also performed during this online interaction. The immediate posttest asked participants to write a narrative retelling of the video, while the stimulated recall involved answering questions about their possible noticing of anything particular in the NS
recast utterances. The delayed posttest\textsuperscript{15} was unique to each participant in that it included a transcript of each participant’s half of the chat interaction; it then asked the participant whether each sentence was correct or contained an error. All items targeted for recast were coded for linguistic category (morphological, semantic, or syntactic). Smith operationalized noticing by instances of what he called “noticing events” in both the eye-tracking data and the stimulated recall data. In the eye-tracking data, a noticing event consisted of the salient part of the recast text being positioned under some amount of blue (or hotter) shaded area of a heat map generated by the eye-tracker. Regarding noticing in the stimulated recall, an earlier non-target-like item was coded as ‘noticed’ if participants said that they noticed a difference between their own output and the interlocutor’s recast. Results showed that the heat map and the stimulated recall measures were strong predictors of immediate posttest success, and, to a certain extent, delayed posttest success. Furthermore, both measures of noticing, when considered independently, are positive predictors of both immediate and delayed posttest success. Regarding linguistic category and noticing, the probability of participants noticing semantic and syntactic recasts on the stimulated recall measure were significantly higher than the probability of noticing morphological recasts. There was no significant relationship between linguistic category and noticing for the heat map measure. Critically, Smith found evidence of what he called “cognitive registration” of the stimuli: there were items coded as being noticed according to the eye-tracking measures but were not noticed according to the stimulated recall measures. Smith concludes that “eye tracking methodology may help researchers untangle the lower levels of the noticing continuum” (p. 72).

Smith and Renaud (2013) used eye-tracking to examine the relationship between L2

\textsuperscript{15} Smith (2012) did not include information on the time elapsed between the treatment session and the delayed posttest.
recasts, noticing, and learning during synchronous computer-mediated communication (SCMC). A total of 16 low-intermediate level L2 students of either German or Spanish at an American university completed a pre-treatment questionnaire, treatment task, and immediate posttest over two sessions. The pre-treatment questionnaire established comparability among students regarding their target language experience. The treatment task consisted of 15-minute individual SCMC writing conferences with a native speaker (NS) of the target language (either German or Spanish) in which the NS provided full recasts in response to participant errors. Eye movements were recorded during this writing conference. The posttest, occurring one week later, was unique to each participant in that it included the non-target-like participant utterances from the SCMC conference. Participants were to determine whether each item was target-like or not; in the case that it was not, they were required to rewrite the sentence in a target-like way. Notably only the recast sentences that were rewritten in a way that followed the form specified by the NS during the SCMC were given a positive point value. This allowed for data to be gathered regarding whether participants were able to productively use the information contained in the NS recast. Recasts were coded for the number of targets found in the recast, the target focus (linguistic type: lexical agreement, tense, spelling, other), and projected difficulty. The two measures of eye-tracking were fixation number and total fixation time. Results revealed that participants fixated on the NS-provided recasts about 72% of the time; they then scored correctly on the corresponding item on the pretest between one-fifth and one-third of the time. In addition, the strongest predictor of posttest success of all variables considered was a participant fixation on the targeted item for at least 200 milliseconds. Fixation time and number was not affected by the target focus, the number of targets found in the recast, or the projected difficulty. The

---

16 A fixation was operationalized as having a duration of at least 200 milliseconds.
researchers concluded that eye-tracking may be an effective methodology to examine the noticing of recasts and should be used in conjunction with concurrent verbal reports in the future.

Winke (2013) conceptually replicated Lee’s (2007) study of the effects of input enhancement on grammar learning and comprehension by using eye-tracking; noticing of the English passive construction enhancement was operationalized as a significant increase in visual foci over a baseline (control) condition. Therefore, in addition to examining the effects of textual enhancement on learners’ form learning and comprehension, the study aimed to determine whether input enhancement significantly drew learners’ visual attention to the targeted linguistic forms, as measured by eye-tracking. Participants included 55 college students of intermediate L2 English proficiency. A pretest and immediate posttest based on Lee (2007) included 10 sentences that assessed passive-related form errors\(^\text{17}\). A modified authentic news text comprised the experimental task; one version was enhanced and the other was unenhanced. Eye movements were recording during reading of the news text. Eye-tracking measures included total fixation time, number of visits to the target area, first-pass reading time, and rereading time. A free-recall task was used to assess comprehension. Results showed that enhancement significantly affected participants’ noticing of the passive forms through longer gaze durations and rereading times. However, participants exposed to the enhanced text did not show significantly increased first-pass reading time. Winke concluded by suggesting that the study demonstrated that future investigations looking at enhancement and noticing require the sensitive and objective measure of noticing that eye-tracking provides.

Godfroid and Uggen (2013) set out to determine whether L2 German learners would attend to and notice vowel-changing verbs (stem-changing verbs versus regular verbs) during

\(^{17}\) No information was included regarding the amount of time elapsed between the pretest and the experimental treatment.
sentence processing and, if so, to what extent they would attend and notice. The participant
group was made up of 40 beginning learners who wrote sentences containing the target verbs
that were shown in a picture for both the pretest and the posttest. Learning was operationalized
as the difference between the pretest and posttest scores. The experimental task via computer
consisted of 24 simple sentences, each of which contained either a regular or irregular verb;
participants were told to read for meaning. After each sentence, participants were presented with
three on-screen pictures and had to look at the one that corresponded to the action described in
the preceding sentence. Eye movements were recorded during this process: increased attention
was operationalized as either significantly longer subtracted fixation times for irregular than for
regular verbs or a direct visual comparison between a possibly vowel-changing form and a non-
vowel changing form. Results of the fixation time data reveal that the participants did in fact pay
more attention to the irregular verbs; the authors thus suggest that it may be possible for
beginning L2 learners to pay attention to small morphological features with no semantic meaning
during a written production task. Furthermore, verb irregularity affected subtracted total time
but not subtracted first fixation or gaze duration; this suggests that the increased attention on the
irregular verbs occurred late in the reading process. Results also showed that total time spent
fixating on the irregular verbs positively impacted learning of that form; the authors pointed out
that this provides evidence for the eye-mind link hypothesized to exist for reading.

In a study triangulating eye-movement recordings and verbal reports, Godfroid and
Schmidtke (2013) aimed to determine the differential contributions of attention and awareness to
receptive vocabulary learning. Participants were 29 advanced L2 English learners who read for
meaning 20 English paragraphs containing 12 novel pseudowords while their eye movements
were tracked, thus operationalizing attention. Target word retention was measured using a
surprise posttest in which participants saw the original sentence in which the target appeared, along with 18 possibilities for the missing word. Post-task interviews were used to measure conscious recollection of the reading of each pseudoword; these interviews operationalized awareness. Results revealed that total fixation time on the pseudoword and recollection predicted word recognition. Also, when fixation times and awareness levels were entered into a single regression model, awareness alone sufficed to predict recognition. The authors conclude that attention, in the form of looking at a word, induced awareness (encoding of a processing episode), a strong predictor of vocabulary learning.

In summary, the above studies highlight the array of constructs and cognitive processes that eye-tracking can help investigate. While Rehder and Hoffman (2005a, b), Theeuwes and Van der Stigchel (2009), and Godfroid and Schmidtke (2013) further investigated attention, Kuo et al. (2009) studied cognitive effort, Witzel et al. (2012) looked at lexical and sentence processing and Yang and McConkie (2001) focused on cognitive processes. Godfroid et al. (2010), Godfroid and Uggen (2013), Smith (2012), Smith and Renaud (2013) and Winke (2013) investigated noticing while Godfroid et al. (2013) studied focal attention, one of the component processes of noticing. The overarching conclusion for these studies as a group is the unique possibility of eye-tracking to measure noticing (Godfroid et al., 2010; Godfroid & Uggen, 2013; Smith, 2012; Smith & Renaud, 2013; Winke, 2013), attention (Godfroid et al., 2013; Godfroid & Schmidtke, 2013; Rehder & Hoffman, 2005a, 2005b; Theeuwes & Van der Stigchel, 2009), cognitive effort (Kuo et al., 2009), and lexical and sentence processing (Winkel et al., 2012). The methodologies and results also highlight the need to take into consideration a variety of eye movement variables when studying the effect of the independent variable.

The studies reviewed underline the necessity of further investigations into cognitive
processes like attention and cognitive effort by means of eye-tracking methodology.

Furthermore, although numerous studies have looked at attention during reading, only Godfroid et al. (2010), Godfroid et al. (2013), and Godfroid and Schmidtke (2013) have reported on the relationship between fixations on novel words and subsequent recognition (see also Brusnighan & Folk, 2012; Williams & Morris, 2004). Godfroid and Uggen (2013) used eye-tracking during an experimental reading task and comprehension questions to establish that participants paid more attention to the irregular verbs and suggest the possibility that beginning L2 learners may be able to attend to small morphological features with no semantic meaning during a reading task. Future studies in this direction can help to develop a finer-grained concept of attention. Godfroid et al. (2010) highlighted the need for studies that “tease apart the lower end of the attentional gradient and see how it relates to learning” (Godfroid 2010, p. 186). However, since intake can occur with no further processing and without entering the developing system, it is uncertain to what extent Godfroid et al.’s statement relates to intake unless, like in previous studies (e.g., Leow, 1997, 2000; Rosa & Leow, 2004), the MC recognition assessment task is viewed as a measure of intake instead of learning. The data obtained in that study assisted in identifying clear cases of heightened attentional processing, resulting in a straightforward measurement of the higher end of the attentional scale. Future studies, therefore, should investigate the lower end of this attentional scale, which would address the proposed fine-grained version of intake.

Furthermore, while Witzel et al. (2012, p. 3), asserted that eye-tracking data from readers can conceivably show the online ease or difficulty of processing and results from Theeuwes and Van der Stigchel (2009) show that the strength of saccade deviation can reveal how much attentional processing is occurring, Godfroid et al. (2010) concluded that neither eye-tracking nor
post-test data can provide any definitive information about the type and depth of cognitive processing that results from noticing; Godfroid et al. (2013) continued this argument by stating that eye movement data cannot be used to provide information about the quality of processing.

Employing Godfroid et al.’s (2013) operationalization of noticing in addition to concurrent think-aloud protocols may serve as a strong model for operationalizations in future studies. By operationalizing noticing as an increase in fixation time relative to a baseline condition, Godfroid et al. avoided limitations of previous studies such as Smith’s (2010) study of noticing of recasts: Smith operationalized noticing as an eye fixation of greater than 500 ms. This operationalization, however, proves problematic given Rayner’s (2009) affirmation that effects smaller than 250 ms are indeed abundant. Furthermore, Godfroid et al. indicated in the case of Smith’s operationalization, that "the noticing threshold will be reached more easily for some (long, infrequent, unpredictable, etc.) words than for other (short, frequent, predictable, etc.) items" (2013, 490). Although Godfroid et al. (2010) also investigated noticing, their investigation was a pilot test and they did not clearly operationalize noticing other than saying that "higher reading times for pseudo words would have to be interpreted as behavioral evidence of noticing” (179).

In summary, eye-tracking is one concurrent data elicitation procedure that clearly provides not only a more fine-grained assessment of attention to form during exposure to written input but also some indication of cognitive processing in relation to target items. While there are many measures of eye-tracking and numerous arguments supporting different measures, it appears prudent to follow the suggestion of examining multiple measures in one study in order to reveal a closer look at the reading process (Rayner, Sereno, et al., 1989; Schmauder, 1992).
Conclusions for the Role of Noticing in SLA

There is little doubt that attention plays a crucial role in SLA; Schmidt (2001) went so far as to say that SLA is driven by what learners pay attention to and notice. While the seven current theoretical frameworks describing the early stages of the learning process in SLA differ in their views of what role awareness plays, they all concur that learning begins with exposure and attention to the L2. Noticing can be measured retrospectively by means of offline verbal reports, however, given that concurrent verbal reports are able to show the learner’s cognitive processes while interacting with the L2 (Bowles, 2010), the attentional strand of SLA (e.g., Leow, 1997a, 1998a, 1998b, 2000, 2001a, 2001b; Rosa & Leow, 2004a, 2004b; Rosa & O’Neill, 1999), has relied on concurrent verbal reports as the main methodological tool to measure learners’ awareness of and attention to certain features of the L2 input. Initial concerns regarding the possible reactivity of concurrent verbal reports have been addressed by Bowles’ (2010) meta-analysis that found that thinking aloud while performing a verbal task has only a small effect on post-task performance, and that those thinking aloud performed only slightly better or slightly worse. But regardless, as indicated in Bowles’ meta-analysis and recommended in the original reactivity study in SLA (Leow & Morgan-Short, 2004), any study employing concurrent verbal reports should also use a silent control group, especially considering that the full range of variables that could lead to reactivity are still unknown. In addition, future studies need to consider coding criteria. Overall, it appears that coding for noticing in past studies greatly overlaps, which highlights the need for fine-grained coding criteria in future studies. In comparison with verbal reports, eye-tracking has been purported to be an uncontroversial measure of overt attention (Blair et al., 2009) and furthermore, several studies have reported a close link between covert attentional processes and eye movement (Rayner, 1998). Empirical
studies such as Godfroid et al. (2010) highlight the need to investigate the lower end of the attentional scale and Leow’s (forthcoming) Model of the L2 learning process proposes three phases of intake, all of which appear to have substantial empirical support. Furthermore, as stated by Robinson, Mackey, Gass, and Schmidt (2012, p. 261), “verbal reports are unlikely to faithfully reflect everything that learners are attending to…” Also, empirical evidence appears to show that depth of processing facilitates several aspects of L2 learning in the early stages (e.g., Craik & Tulving, 1975; Leow et al., 2008; Morgan-Short et al., 2012; Qi & Lapkin, 2001; Shook 1994). In addition, the few studies that have compared grammatical versus lexical items (Greenslade et al., 1999; Leow et al., 2008; Mackey et al., 2000; Morgan-Short et al., 2012; Smith, 2012; VanPatten, 1990) have obtained conflicting results regarding the role of type of linguistic item in intake and its relation to comprehension, therefore motivating future studies on the topic. Taking into consideration this discussion on attention, the next logical methodological step seems to be to follow Leow et al.’s (2014) advice to triangulate the concurrent data collection measures of online verbal reports and eye-tracking to investigate level of intake in L2 learners, taking into account variables such as depth of processing and type of linguistic item and also to address empirically the issue of reactivity.

**Rationale and Research Questions**

**Rationale for the present study.**

Leow’s (forthcoming) Model of the L2 learning process makes several postulations regarding the early stages of the L2 learning process that clearly need to be tested empirically. More specifically, his view of the construct of intake as a multi-leveled one and the role of depth of processing at these early stages warrant empirical support. In addition, due to the inconsistent findings on whether type of linguistic item plays a role in input processing, further research into
the difference between grammatical forms and lexical items in recognition and production tasks is necessary in order to achieve a better understanding of this variable. Finally, the study seeks to empirically address the role of reactivity in L2 performance.

To this end, the following lines of research constitute the impetus for the present investigation: (a) research on the possibility of different levels of intake, as measured by both think-aloud protocols and eye-tracking, (b) research on depth of processing as evidenced in concurrent verbal reports, (c) the role different types of linguistic items play on intake, (d) the relationships between these variables, and (e) the role of reactivity in L2 performance.

**Research questions.**

The main research questions guiding this dissertation are the following:

RQ1) Does thinking aloud while performing a reading task have any detrimental effect on beginning Spanish readers’ a) recognition, b), production, and c) comprehension of a grammatical form and lexical items when compared to readers not thinking aloud?

RQ2) Does level of intake\(^\text{18}\) have an effect on adult L2 learner’s recognition of a) a grammatical form and/or b) lexical items from a written text as measured by recognition tests?

RQ3) Does level of intake have an effect on adult L2 learner’s controlled written production of a) a grammatical form and/or b) lexical items from a written text as measured by controlled written production tests?

\[\text{\^{18} In an effort to avoid lengthy and convoluted research questions, the operationalization of the independent variable level of intake has been omitted from the research questions and instead is available here: level of intake is operationalized by time spent on experimental reading passages and is confirmed in the Grammatical and Lexical Noticed Intake Groups by clicking data.}\]
RQ4) Are there any relationships between L2 learners’ eye fixation time measures (first fixation duration, gaze duration, second pass time, total fixation time) and their abilities to recognize a) a grammatical form and/or b) lexical items in written L2 input as measured on recognition posttests?

RQ5) Is there a relationship between depth of processing, as measured by form-meaning connections in think-aloud protocols during a reading task, and subsequent recognition and written production of a) a grammatical form and/or b) lexical items embedded in the reading task?

RQ6) Is there a relationship between depth of processing, as measured by form-meaning connections in think-aloud protocols during a reading task, and comprehension of a reading passage, as measured by a multiple-choice comprehension test?

RQ7) Are there any differences between recognition of a) a grammatical form and b) lexical items in beginning Spanish learners from a written text as measured by recognition tests?
Chapter Three: Research Design and Methodology

This third chapter provides information on the pilot study in addition to a detailed description of the proposed research method to address the research questions.

Pilot Study.

The researcher conducted a pilot study of the current study approximately six months before the main data collection began. The pilot study was on a much smaller scale and included only some of the methodological components of the current study.

Four research questions were included in the pilot study:

RQ1) Does level of intake have an effect on adult L2 learners’ recognition of a) a grammatical form and/or b) lexical items from a written text as measured by recognition tests?

RQ2) Are there any significant relationships between level of intake and fixation time measures (first fixation duration, gaze duration, second pass time, and total fixation time) of a) a grammatical form and/or b) lexical items?

RQ3) Are there any relationships between L2 learners’ eye fixation time measures (first fixation duration, gaze duration, second pass time, total fixation time) and their abilities to recognize a) a grammatical form and/or b) lexical items in written L2 input as measured on recognition posttests?

RQ4) Are there any differences between recognition of a) a grammatical form and b) lexical items in beginning Spanish learners?

Pilot methodology.

Data were collected from a total of ten adult speakers of English. English was the first language of eight of the participants and Chinese was the first language of two participants. No
participant had more than one year of previous formal exposure to Spanish. There were three experimental conditions: a noticed intake condition \( (n = 5) \), a detected intake condition \( (n = 3) \), and an attended intake condition \( (n = 2) \). Each participant contributed grammatically and lexically. In other words, a participant in the noticed intake condition was part of the grammatical noticed intake group and the lexical noticed intake group.

The noticed intake group for the target grammatical form and the target lexical items had five participants, the detected intake groups had three participants, and the attended intake group had two participants.

There was only one experimental reading passage ("Juanito") and it contained 10 instances of the same target grammatical form (se quedó, lavó, charló, tomó, tocó, bailó, regresó, se acostó, manejó, miró) and 10 target lexical items (bolsa, novia, gente, entrada, carnet, canción, despertador, cita, dolor, fecha). There were no control items. The reading passage was presented via computer, one sentence per screen. The noticed group had no time limit per sentence, the detected group had nine seconds per sentence, and the attended group had six seconds per sentence. These time limits were established based on informal piloting of these times among beginning learners of Spanish.

Regarding the materials, the pretests included a verb stem meaning pretest and an item pretest with both grammatical and lexical items. Posttests included a recognition posttest and a

---

19 There were no separate intake conditions or groups for grammatical level of intake and lexical level of intake, nor were participants reassigned into intake level groups based on the time they spent on the reading.

21 All pretests and posttests used in the pilot study had the same format as those used in the current study. The only difference was that some instances of the target grammatical form and the target lexical items were different.
comprehension posttest. All tasks except the experimental reading passage were performed as pen and paper tasks.

**Pilot procedure.**

The pilot study consisted of two experimental sessions: one during which the pretests and a language background questionnaire were administered and the second, approximately five days later, during which the experimental exposure and immediate posttests took place. During the second experimental session, participants first completed a practice think-aloud task and a practice eye-tracking task. They then read the experimental reading passage while having their eye movements tracked. Participants also thought their thoughts aloud while reading. The experimental reading passage was followed by the posttests and debriefing questionnaire.

**Pilot coding.**

All tests were scored using the same criteria as the current study. The measures used for eye-tracking were total duration, first fixation duration, gaze duration, and second pass time. The think-aloud protocols were coded using criteria similar to Leow et al. (2008) and Morgan-Short et al. (2012) and had three levels (regardless of whether the target was grammatical or lexical) based on the completeness of the form-meaning connection.

**Pilot results and discussion.**

Descriptive statistics were used to further investigate the research questions. Regarding the possible effect of level of intake on recognition, the grammatical detected group took in the highest amount of target grammatical forms on the recognition posttest and the grammatical attended group took in the lowest. For the target lexical items, the lexical noticed intake group had the highest recognition posttest scores while the lexical attended intake group had the lowest.

---

22 No written production tests were used.
One plausible explanation might have been the relative amount of time allocated per level, which, based on the think alouds, appeared to have permitted processing in many instances previously allocated to the level of noticing. To this end, it was quite challenging to address the issue of depth of processing between these three levels and this issue will be discussed in the section outlining proposed methodological changes to the research design.

Results from the second research question showed that there seems to be a trend for the total fixation time measure: both the grammatical and lexical noticed intake groups exhibited considerably greater means than did the detected intake groups, and in turn the detected intake groups had greater means than the attended groups.

Descriptive statistics from the third research question were computed by summed fixation times for all grammatical/lexical targets for each participant. They suggest a positive relationship between correct recognition and both gaze duration and second pass time of the target grammatical form. Likewise, it also appears that there may be a positive relationship between incorrect recognition and both first fixation duration and total fixation time for the target grammatical form. These relationships do not seem to cross over to target lexical items: there appears to be a positive relationship between all four fixation time measures and correct recognition of the target lexical items.

Last of all, the fourth research question found that participants had a mean score almost two points higher on recognition of target lexical items (7.70) in comparison with recognition of the target grammatical form (5.10). This hints that lexical linguistic items may be easier for beginning learners to recognize than grammatical linguistic items.
Pilot modifications.

This pilot study served as a first attempt to use two forms of concurrent data to investigate the proposed levels of intake demonstrated in several previous intake studies. While the conditions designed to elicit specific levels of intake were innovative and did create a strong starting point for future studies, these conditions needed to be refined. As the descriptive results show, while overall there did appear to be a hierarchy of performance according to level, there was seldom a clear distinction between noticed, detected, and attended intake. It is probable that one of the main issues causing this were the time limits per sentence assigned to the experimental groups. While the noticed group had no time limit, the detected group had nine seconds per sentence and the attended group had six seconds per sentence. The think aloud protocols, especially used to code for depth of processing, appear to indicate that such time limits did provide participants in both the attended and detected intake groups enough time to report processing that approximates that reported in studies that have addressed the construct of noticing via concurrent verbal reports. To address this issue, a perceived time limit with a “prize” was used in the noticed, detected, and attended conditions of the main study. This change was made in hopes of creating more defined differences between the levels of intake. Furthermore, a greater number of participants in the main study would help to offset any individual differences that may have affected the data of the pilot study due to the small number of participants per cell.

The experimental conditions of the main study were also modified. There were still noticed, detected, and attended intake groups, however none of these groups had corresponding think-aloud groups (see below for further explanation). Given the greater intake associated with levels above noticing documented in previous studies, a noticed+ TA group was added, in which
participants thought aloud; reactivity was controlled for in a silent noticed + group. These two noticed+ groups differed in that there was no eye-tracking and participants had no perceived time limit.

Since the think aloud protocols and descriptive statistics did not appear to have adequately distinguished the three levels of intake, the experimental conditions were adjusted via the use of additional instructions and textual enhancement, in addition to the addition of the perceived time limits as discussed previously. More specifically, the instructions of the reading passage will vary across experimental conditions in terms of: 1) how the participant is instructed to read, 2) textual enhancement, 3) whether or not the participant is told to pay attention to the textual enhancement, 4) the verbalizations the participant is asked to make, and 5) whether or not the participant is told to use the mouse to click on the textually enhanced words.

The researcher noted during the pilot study that some of the instances of the target grammatical form and the target lexical items needed to be changed for the main study. For example, tocó was removed as a target because tocar can mean both “to touch” and “to play” and therefore might confuse participants. Regarding the target lexical items, despertador was replaced because the researcher found that too many participants confused it with despertarse. Novia and gente were replaced because many participants already had accurate prior knowledge of these words; cita was also replaced because it is too short. Lastly, dolor was also replaced because participants tended to translate tenía dolor de cabeza as “his head hurt”, thus completely avoiding any translation of dolor.

Also, a control grammatical form and control lexical items was added to the experimental reading passage in order to provide control data with which to compare the fixation measurements of the target grammatical form and target lexical items. This data was to be
analyzed similar to Godfroid et al. 2013), who measured noticing as a significant increase in fixation from the control to the target. A pretest in which participants had to match words to their meanings ensured that all participants had a similar knowledge of non-target words.

**Current Study**

The design and execution of the current study was carried out by taking into account both the methodologies of past studies and the researcher’s experience with the pilot study.

**Method**

*Participants.*

Participants were 233 adult speakers of English enrolled in their first semester of Spanish at one of two American universities (one Midwestern and one East Coast). Recruitment occurred through participants’ Spanish classes; either the researcher or a research assistant visited all sections of first-semester Spanish to explain the opportunity to participate in the study. Compensation for completing all three experimental sessions was extra credit points applied to either a quiz or the midterm, depending on the school.

L1 status was confirmed by a modified version of Freed, Dewey, Segalowitz, and Halter’s (2004) Language Contact Profile (Appendix A: Language Background Questionnaire); this background questionnaire also solicited information regarding their prior experience with Spanish and other languages. In order for their data to be included in the study, participants had to meet the following criteria: a) speak English as a native language; b) speak no other language with a self-rated proficiency of greater than 1 on scale of 0-3; c) have taken no more than two high/middle/elementary school years of Spanish; d) never have studied abroad in a Spanish-speaking country or have traveled in a Spanish-speaking country with the purpose of learning.
Spanish; and e) complete both experimental sessions. The data of any participant not meeting this qualification were excluded from the study.

After applying the participation criteria, a total of 96 participants remained and the data from these participants were used in the statistical analyses. Of the participants whose data was excluded from the statistical analysis, in 60 cases it was due to participants demonstrating prior knowledge of the target grammatical form on the pretests and/or having had more than two years of formal Spanish instruction. In 33 cases it was due to participants who were not native English speakers or who grew up speaking English and another language. In 30 cases it was because participants’ self-rated proficiency of another L2 was higher than 1. In 2 cases it was because the participants had traveled extensively abroad with the intent of learning Spanish, and in 11 cases it was due to participants not completing all three experimental treatments. The data of one participant was excluded because s/he had a vision disorder, and the data from one additional participant was excluded because s/he was more than two standard deviations away from the mean age of participants. The remaining 95 participants’ biographical information is summarized in Error! Reference source not found. below; their proficiency scores appear in Table 3.

Table 2: Participant Characteristics, Overall and by Intake Group

<table>
<thead>
<tr>
<th>Grammatical intake groups</th>
<th>Age</th>
<th>Gender (Male, Female)</th>
<th>Years of Formal Study in Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noticed+ TA Intake Group</td>
<td>19</td>
<td>17 / 20</td>
<td>18.26 (.65)</td>
</tr>
<tr>
<td>Noticed+ Silent Intake Group</td>
<td>19</td>
<td>18 / 27</td>
<td>20.05 (2.17)</td>
</tr>
<tr>
<td>Noticed Intake Group</td>
<td>19</td>
<td>18 / 27</td>
<td>19.74 (2.05)</td>
</tr>
<tr>
<td>Detected Intake Group</td>
<td>20</td>
<td>18 / 25</td>
<td>20.00 (2.13)</td>
</tr>
<tr>
<td>Attended Intake Group</td>
<td>19</td>
<td>18 / 22</td>
<td>19.16 (1.26)</td>
</tr>
</tbody>
</table>

Lexical intake groups
<table>
<thead>
<tr>
<th>Intake Group</th>
<th>Listening</th>
<th>Oral</th>
<th>Reading</th>
<th>Writing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noticed + TA Intake Group</td>
<td>.32 (.48)</td>
<td>.32 (.48)</td>
<td>.58 (.51)</td>
<td>.37 (.50)</td>
</tr>
<tr>
<td>Noticed + Silent Intake Group</td>
<td>.58 (.51)</td>
<td>.42 (.51)</td>
<td>.53 (.51)</td>
<td>.47 (.51)</td>
</tr>
<tr>
<td>Noticed Intake Group</td>
<td>.47 (.51)</td>
<td>.37 (.50)</td>
<td>.47 (.51)</td>
<td>.37 (.50)</td>
</tr>
<tr>
<td>Detected Intake Group</td>
<td>.65 (.49)</td>
<td>.60 (.50)</td>
<td>.65 (.49)</td>
<td>.65 (.49)</td>
</tr>
<tr>
<td>Attended Intake Group</td>
<td>.63 (.50)</td>
<td>.42 (.51)</td>
<td>.68 (.48)</td>
<td>1.05 (2.46)</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td>.53 (.50)</td>
<td>.43 (.50)</td>
<td>.58 (.50)</td>
<td>.58 (1.19)</td>
</tr>
</tbody>
</table>

The experimental treatment occurred between the sixth and seventh week of the semester. Both Spanish language programs at both universities used communicative and cognitive approaches to L2 learning. All participants had several exposures to authentic Spanish reading passages like the ones used as the experimental treatment in this study prior to participation.

*The target grammatical form and target lexical items.*

This study used both a target grammatical form and target lexical items. The following is a description of both and the rationale for their selection.
The target grammatical form.

The target grammatical form in this experiment was the third-person singular form of regular –ar verbs in the preterit. The mood was indicative. The preterit is previous to the moment of speech time and refers to the past (e.g., Alonso & Henríquez Urea, 1974). Considered to be the absolute past or the remote past (RAE, 1973), the preterit has no relation to the present and is used to express an event or situation performed in the part, an action that interrupts another, and an action considered to be far from the moment of the speech act.

There are three conjugations of verbs in Spanish: those that end in –ar, those that end in –er, and those that end in -ir. Verbs that are conjugated regularly in the past both drop their ending and add a suffix, depending on the person and number to be reflected, as seen in Table 4.

<table>
<thead>
<tr>
<th></th>
<th>Suffix</th>
<th>Example</th>
<th>Suffix</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st person</td>
<td>-é</td>
<td>estudié</td>
<td>-amos</td>
<td>estudiamos</td>
</tr>
<tr>
<td>2nd person</td>
<td>-aste</td>
<td>estudiate</td>
<td>-asteis</td>
<td>estudiateis</td>
</tr>
<tr>
<td>3rd person</td>
<td>-ó</td>
<td>estudió</td>
<td>-aron</td>
<td>estudiaron</td>
</tr>
</tbody>
</table>

The third person singular form of preterit was the target grammatical form of the current study; only regular –ar verbs were used. The twenty forms of the target grammatical form are: se quedó (s/he stayed), lavó (s/he washed), charló (s/he chatted), tomó (s/he took), escuchó (s/he listened), bailó (s/he danced), regresó (s/he returned), se acostó (s/he went to sleep), manejó (s/he drove), miró (s/he looked), se levantó (s/he got up), navegó (s/he surfed), llegó (s/he arrived), pensó (s/he thought), desayunó (s/he ate breakfast), nadó (s/he swam), preguntó (s/he asked), habló (s/he spoke), descansó (s/he rested), caminó (s/he walked).

The preterit in Spanish has characteristics that make it both easy and difficult to learn. First, it can be considered an easy form to learn because it is a simple tense, in comparison with
compounds, and requires no complicated temporal explanations for beginning students who have not yet learned the imperfect. Furthermore, the mood is indicative and therefore avoids any discussion of the more challenging subjunctive or imperative. Also, it is relatively easy for beginning learners of Spanish to comprehend because it does not differ in mood or aspect from the present indicative, the first tense they learn; it only varies in tense. The accented final vowel of regular verbs in the third-person singular form of the preterit may serve to draw learner’s attention to the ending, thus facilitating learning. Hulstijn and De Graaff (1994) posit that simple formal phenomena may be salient enough in the input to be discovered by L2 learners spontaneously, without the help of explicit instruction.

Regardless, there are also characteristics of the preterit that make it difficult to learn. The preterit is often accompanied by context clues and temporal adverbs that may complicate the acquisition of this verb tense because they render the verb inflection redundant. Also, there is little difference between regular third-person singular verbs in the present and in the preterit: *escucha* (s/he listens, third-person singular present) versus *escuchó* (s/he listened, third-person singular preterit) differs in only one phone. Beginning Spanish learners struggling to make meaning out of input may very well focus solely on the lexical meaning of *escuchó* and disregard the tense completely.

Given the methodology of the current study, the ideal situation would be one in which some participants would be able to decipher the meaning of the target grammatical form on their own, and therefore a simple grammatical form is most conducive to this. Conceptually, the preterit is a very simple form of the past to learn because it does not require understanding complex temporal relations, mood, or aspect. Second, it is a form typically found in the narrative genre and it would lend itself well to use in a reading passage for beginning learners. Last of all,
this target grammatical form aligned well with the existing syllabus in that it is presented early in the semester and thus allowed for the participants to have had only very limited and controlled exposure to the L2 before the time of data collection.

*The target lexical items.*

The 20 target lexical items included eleven masculine and nine feminine nouns, all appearing in the singular form. The accompanying article was not considered to be a part of the target lexical item; in the reading passages, four target lexical items were preceded by a definite article, seven by an indefinite article, four by a possessive determiner, one by an adverb of comparison, and four were not preceded by any article or determiner. The twenty target lexical items are: *abrazo* (hug), *apellido* (last name), *cabezada* (nap), *canción* (song), *carnet* (identification card), *cartera* (purse), *cuna* (crib), *duda* (doubt), *entrada* (entrance), *feriado* (holiday), *fichero* (file cabinet), *mascota* (pet), *mesón* (diner), *negocio* (business), *parentela* (relatives), *paseo* (avenue), *relato* (story), *reloj* (clock), *taza* (mug), *traje* (suit).

The reading passages in which the twenty target lexical items appeared facilitated the understanding of these items. In addition, all other words included in the experimental reading passages were comprised of vocabulary from the first two chapters of the textbook used in the participants’ Spanish class.

*Rationale for the selection of the target lexical items.*

There were several motivations behind the choice of the specific target lexical items. First, none appeared in the textbook or vocabulary lists of participants’ Spanish classes until after data collection for the current study. Second, all target lexical items are concrete concepts; this
seemed appropriate for beginning learners of Spanish as the labels of concrete concepts are easier to acquire than those of abstract concepts (De Groot & Keijzer, 2000).

**The control grammatical and lexical items.**

Twenty instances of the control grammatical form and twenty control lexical items were also included in the experimental reading passages in order to provide data against which to measure noticing of the target grammatical form and target lexical items (cf. Godfroid et al, 2013). This data allowed a comparison (see the Results chapter) between the target grammatical form and the control grammatical form, in addition to a comparison between the target lexical items and the control lexical items, regarding recognition, production, and eye fixation measures. In this way, it was possible to determine whether participants did in fact have higher scores and values on targets than on control.

Each of the twenty instances of the control grammatical form and twenty control lexical items appeared in a different sentence; control items were not in sentence-initial nor sentence-final position. All exactly substituted the corresponding target.

**The control grammatical form.**

The control grammatical form was the third person present form of regular –ar verbs. Each of the twenty instances of the control grammatical form appearing in the two reading passages was the third person present singular form of one of the twenty instances of the target grammatical form (third person preterit form of regular –ar verbs). In other words, the only difference between the target and control grammatical forms was that the target was preterit and the control was present. Participants had already extensively studied this grammatical form in their Spanish class and they had already been tested on it in class by the time they participated in the experiment. The twenty instances of the control grammatical form included the following: *se*
Qued (s/he stays), lava (s/he washes), charla (s/he chats), toma (s/he takes), escucha (s/he listens), baila (s/he dances), regresa (s/he returns), se acuesta (s/he goes to sleep), maneja (s/he drives), mira (s/he looks), se levanta (s/he gets up), navega (s/he surfs), llega (s/he arrives), piensa (s/he thinks), desayuna (s/he eats breakfast), nada (s/he swim), pregunta (s/he asks), habla (s/he speaks), descansa (s/he rests), camina (s/he walks).

The control lexical items.

The control lexical items were all basic Spanish words that all participants should already have known. Most were part of the first chapter vocabulary from the participants’ Spanish class; others, such as amigo, are words that most Americans who have not even studied Spanish could recognize. There were twenty control lexical items, however due to a researcher error, two of these items (mensaje, tarea) appeared in both reading passages and therefore each participant was exposed to them twice, not once. These control grammatical items were: amigo (friend), cama (bed), escuela (school), fiesta (party), dinero (money), leche (milk), libro (book), lunes (Monday), madre (mother), mensaje (message), noche (night), periódico (newspaper), puerta (door), sábado (Saturday), siesta (nap), sombrero (hat), tarea (homework), trabajo (job).

Experimental conditions

Level of intake conditions.

As addressed in the previous chapter, there are both theoretical accounts (Chaudron, 1985; Faerch & Kasper, 1980; Slobin, 1985) and empirical studies (Faretta-Stutenberg and Morgan-Short, 2011; Godfroid et al., 2010; Lee, 1998; Leow, 1998b, 2000, 2001a; Martínez-Fernández, 2008, Rosa & Leow, 2004; Rosa & O’Neill, 1999; Williams, 2004, 2005) that suggest the existence of different levels of intake in the early stages of L2 learning. In the current study, all participants were initially placed randomly into one of five level of intake
conditions (attended intake, detected intake, noticed intake, noticed+ think-aloud (TA) intake, noticed+ silent intake). These conditions differed in how the participants were instructed to read the reading passages and were based on Leow’s (forthcoming) intake phase characteristics (see Table 5 and Table 6). More specifically, the instructions of the reading passage varied across experimental conditions in terms of: 1) how the participant was instructed to read, 2) the perceived time allotted for exposure to the reading passage, 3) textual enhancement\textsuperscript{23}, 4) whether or not the participant was told to pay attention to the textual enhancement, 5) the verbalizations the participant was asked to make, 6) whether or not the participant was instructed to click on any unknown words, and 7) the use of an invisible boundary paradigm. See Table 6

\textsuperscript{23} Textual enhancement in this study was at the word level: if there was textual enhancement, then the entire target word was textually enhanced, not just the ending.
for an overview.

<table>
<thead>
<tr>
<th>Experimental condition characteristics</th>
<th>Verbalization</th>
<th>Perceived time limit</th>
<th>Reading textual enhancement</th>
<th>Explicit instruction to attend to enhancement</th>
<th>Click on unknown word</th>
<th>Eye-tracking</th>
<th>Invisible boundary paradigm</th>
</tr>
</thead>
<tbody>
<tr>
<td>noticed intake condition</td>
<td>none</td>
<td>yes</td>
<td>bolded and enlarged</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>detected intake condition</td>
<td>none</td>
<td>yes</td>
<td>bolded</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>attended intake condition</td>
<td>none</td>
<td>yes</td>
<td>bolded</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>noticed+ TA condition (higher depth of processing condition with TA)</td>
<td>TA</td>
<td>no</td>
<td>bolded and enlarged</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>noticed+ silent condition (silent higher depth of processing)</td>
<td>none</td>
<td>no</td>
<td>bolded and enlarged</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 5: Intake Characteristics

<table>
<thead>
<tr>
<th>Level of intake</th>
<th>Depth of processing</th>
<th>Attention</th>
<th>Awareness</th>
<th>Cognitive effort</th>
</tr>
</thead>
<tbody>
<tr>
<td>noticed</td>
<td>very low</td>
<td>selective/focal</td>
<td>+</td>
<td>low</td>
</tr>
<tr>
<td>detected</td>
<td>nonexistent or very low</td>
<td>selective</td>
<td>-</td>
<td>very low</td>
</tr>
<tr>
<td>attended</td>
<td>nonexistent</td>
<td>peripheral$^{24}$</td>
<td>-</td>
<td>nonexistent</td>
</tr>
<tr>
<td>noticed+ TA</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>noticed+ silent</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

$^{24}$ As operationalized by a lack of fixation through the invisible boundary paradigm.
Table 6: Experimental Condition Characteristics

<table>
<thead>
<tr>
<th>Experimental condition characteristics</th>
<th>Verbalization</th>
<th>Perceived time limit</th>
<th>Reading textual enhancement</th>
<th>Explicit instruction to attend to enhancement</th>
<th>Click on unknown word</th>
<th>Eye-tracking</th>
<th>Invisible boundary paradigm</th>
</tr>
</thead>
<tbody>
<tr>
<td>noticed intake condition</td>
<td>none</td>
<td>yes</td>
<td>bolded and enlarged</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>detected intake condition</td>
<td>none</td>
<td>yes</td>
<td>bolded</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>attended intake condition</td>
<td>none</td>
<td>yes</td>
<td>bolded</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>noticed+ TA condition</td>
<td>TA</td>
<td>no</td>
<td>bolded and enlarged</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>noticed+ silent condition</td>
<td>none</td>
<td>no</td>
<td>bolded and enlarged</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Given that noticed intake is characterized theoretically and empirically by a very low level of awareness that presupposes a low depth of processing and the potential for cognitive effort, in comparison to detected intake and attended intake, noticed intake was operationalized in this study as instructions to read for overall meaning and to pay attention to and click on enlarged words in bold print. Having participants click on textually enhanced words ensured that all items were attended to in order to minimally promote some processing; this clicking is similar to the instruction to circle in Leow et al. (2008). It is important to note that no feedback on the meaning of the word or a gloss was provided in any experimental condition when participants clicked on a word. All participants in this group had their eye movements tracked. Also, there was a perceived time limit: participants were told to try to finish before the time limit, without specifying exactly what that time limit was. In other words, the goal of the perceived time limit was to have participants perform under time pressure. Participants were told they would receive a small prize if they finished before the time limit, however in reality, all participants received the small prize of a chocolate bar. The perceived time limit was intended to incite participants to process the target in question but avoid reaching higher depths of processing. Furthermore, textual enhancement was employed in an attempt to promote some level of cognitive registration of these items in the input and also to ensure noticing. The target grammatical form and target lexical items were both bolded and enlarged in the reading passage of the noticed intake condition. The instructions of the reading passage advised participants to pay attention to these enlarged and bolded words. The instructions of the noticed intake conditions were the following:

You are about to read two passages. Be sure to read for meaning. You will be tested afterward on your OVERALL comprehension, so it is NOT IMPORTANT that you understand EVERY SINGLE word. There is a time limit. You will receive a small
prize if you finish before the time limit. Pay attention to the enlarged words in bold print and use the mouse to click these words. Each paragraph is presented on a separate screen. There are 12 screens in total, six for each reading passage. When you are finished with one screen, press OK to advance to the next screen. We will not be able to help you in any way. Don’t forget to finish before the time limit and get your prize so read quickly!

Detected intake is characterized by some selective or focal attention being paid to the linguistic data along with a lack of cognitive effort, depth of processing, and awareness. Therefore, the instructions for this condition told participants to read quickly for overall comprehension. As in the noticed intake condition, there was also a perceived time limit and eye-tracking was used. The detected intake condition only saw target words bolded instead of enlarged and bolded. Furthermore, the detected intake condition received instructions to pay special attention to bolded words. They did not have the instruction to click on any words in order to discourage them from processing more deeply and because of the time limit. The instructions for the detected intake condition we as follows:

You are about to read two passages. Be sure to read for meaning. You will be tested afterward on your overall comprehension, so it is NOT IMPORTANT that you understand every single word. There is a time limit so read quickly. You will receive a small prize if you finish before the time limit. Pay attention to the words in bold print. Each paragraph is presented on a separate screen. There are 12 screens in total, six for each reading passage. When you are finished with one screen, press OK to advance to the next screen. We will not be able to help you in any way. Don’t forget to finish before the time limit and get your prize so read quickly!
The experimental condition of attended intake was characterized by peripheral attention and a lack of depth of processing, awareness, and cognitive effort. In order to operationalize this low level of intake, the instructions of the reading passage told participants to only skim for meaning. There was a perceived time limit and eye-tracking was used. In this condition, there was no instruction telling participants to pay attention to or click on words. It was hypothesized that these instructions would be most likely to create a condition in which the majority of participants employed only peripheral attention, as incited by the very short amount of time participants had to spend on the reading, and no depth of processing or cognitive effort to process the target grammatical form and target lexical items in the input.

Critically, the attended intake condition also utilized an invisible boundary paradigm for all instances of the target and control grammatical forms and the target and control lexical items; this served as the operationalization of peripheral attention. In other words, each target and control appeared normally in the reading passages until the participant fixated on any part of a particular target or control, at which moment that target or control disappeared and was replaced by upper-case Xs until the eye fixated on another part of the reading passage. This happened as many times as the participant fixated on a certain target or control, not just the first time. Since conscious perception can occur if a reader looks at something for 50-60 milliseconds during reading, the invisible boundary paradigm enabled the targets and controls to be processed only parafoveally (i.e., never fixated on).

The instructions for the reading passage for the attended intake condition were the following:

You are about to read two passages. Be sure to skim for meaning. You will be tested afterward on your OVERALL comprehension, so it is NOT IMPORTANT that you
understand EVERY SINGLE word. **There is a time limit so read quickly.** You will receive a small prize if you finish before the time limit. Each paragraph is presented on a separate screen. There are 12 screens in total, six for each reading passage. When you are finished with one screen, press the space bar to advance to the next screen. We will not be able to help you in any way. **Don’t forget to finish before the time limit and get your prize so read quickly!** Please note that some words may disappear and reappear as you read.

The noticed+ TA condition included all the same experimental condition characteristics as the noticed intake condition with two exceptions: there was no perceived time limit and participants were instructed to think aloud as they read. Since participants thought aloud, there was no eye-tracking used. Furthermore, participants were not instructed to click on any words in order to facilitate a more natural style of reading. Participants in this condition were told they had unlimited time to complete the experimental reading passage and they were also instructed to think aloud. The purpose of this condition was to address the issue of depth of processing in relation to subsequent processing and thus collect data from participants who reached considerably higher levels of awareness and depth of processing in comparison with the noticed intake condition. The instructions for this condition were:

You are about to read two passages. Be sure to read for meaning. You will be tested afterward on your **OVERALL** comprehension. **THERE IS NO TIME LIMIT.** You may take as much time as you need. Pay attention to the **enlarged words in bold print.** In this experiment we are also interested in what you think about as you complete this task. In order to find out, we are going to ask you to **THINK ALOUD** in English from the time you start the task to when you finish the task. We would like you to talk
CONSTANTLY. We don't want you to try to plan out what you say or try to explain to us what you're saying. Just act as if you are alone in the room speaking to yourself.

What's most important is that you keep talking, and talk clearly and loudly enough into your microphone. We will not be able to help you in any way.

The noticed+ silent condition as exactly the same as the noticed+ TA condition with the only exception that the silent deeper processing condition did not think aloud. There was also no eye-tracking. This condition served as a control condition to ensure that there was no reactivity from thinking aloud. The instructions for this condition were:

You are about to read two passages. Be sure to read for meaning. You will be tested afterward on your OVERALL comprehension. THERE IS NO TIME LIMIT. You may take as much time as you need. Pay attention to the enlarged words in bold print. We will not be able to help you in any way.

Determination of final group membership.

It is important to note that the experimental conditions were designed to have participants perform at a certain level of intake. Each participant was randomly assigned to an experimental condition, however participants were then assigned to a grammatical intake group (Grammatical Noticed Intake Group, Grammatical Detected Intake Group, Grammatical Attended Intake Group, Grammatical Noticed+ TA Group, Grammatical Noticed+ Silent Group) and a lexical intake group (Lexical Noticed Intake Group, Lexical Detected Intake Group, Lexical Attended Intake Group, Lexical Noticed+ TA Group, Lexical Noticed+ Silent Group) based on the time they spent reading the grammatical reading passage and the lexical reading passage, respectively. More specifically, after all data was collected, the researcher compared the time that each
participant of the noticed, detected, and attended intake conditions took to read each
reading passage. These times were the basis for group assignment since the main
distinguishing characteristic of each condition was the time spent reading.
The bottom third times for each reading passage comprised the attended intake group, the
middle third for each reading passage comprised the detected intake group, and the top
third for each group comprised the noticed intake group. Therefore, the term “condition”
is used to refer to the condition to which a participant is initially randomly assigned,
while “group” refers to the final intake group that a participant is part of due to the time
s/he spent reading the reading passages. Reading time best quantifies level of intake
because the characteristics of level of intake (depth of processing, awareness, cognitive
effort) seem to depend highly on time spent in interaction with the reading passage. It is
important to note that participants in the noticed+ TA condition and the noticed+ silent
condition were all transferred directed to the Noticed+ TA Intake Group and the
Noticed+ Silent Intake Group, respectively, unless they did not perform according to the
specifications of each condition25.

In summary, each participant was initially assigned randomly to one of five conditions
(noticed intake condition, detected intake condition, attended intake condition, noticed+ TA
intake condition, noticed+ silent intake condition). Due to the Latin square design discussed in
detail in the next section, each participant saw one reading passage with the target grammatical
form and the other reading passage with the target lexical items. The time for the reading
passage with the target grammatical form for each participant determined their placement into

25 As will be discussed in an upcoming section, all participants in these conditions did in fact perform
according to condition specifications. They also all had much longer reading times than did the other
intake conditions and groups, so there was no need to change their groups.
their grammatical intake group (Grammatical Noticed Intake Group, Grammatical Detected Intake Group, Grammatical Attended Intake Group, Grammatical Noticed+ TA Group, Grammatical Noticed+ Silent Group), while the time for the reading passage with the lexical target items for each participant placed them into their lexical intake group (Lexical Noticed Intake Group, Lexical Detected Intake Group, Lexical Attended Intake Group, Lexical Noticed+ TA Group, Lexical Noticed+ Silent Group).

*Latin square design.*

As will be described in the next section, there were two versions of each of two reading passages: a target grammatical version (target grammatical, control lexical) and a target lexical version (control grammatical, target lexical). While each participant read both reading passages (“Juanito” and “Laura”), that participant only read one version of each reading passage. For example, if a participant read the target grammatical version (target grammatical, control lexical) of “Juanito,” then the same participant read the target lexical items version of “Laura.” In order to control for reading passage order, half of participants read “Juanito” first, while the other half read “Laura” first. This design resulted in a Latin Square design in which each participant was randomly assigned to one of four reading conditions per intake condition (five possible), thus resulting in 20 conditions (see Table 7).

Table 7: Latin Square Design

<table>
<thead>
<tr>
<th>Intake Condition</th>
<th>Reading passage presented First</th>
<th>Reading passage presented Second</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noticed+ TA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Juanito</td>
<td>Laura</td>
</tr>
<tr>
<td></td>
<td>Target grammatical form</td>
<td>Control lexical items</td>
</tr>
<tr>
<td></td>
<td>Control grammatical form</td>
<td>Target lexical items</td>
</tr>
<tr>
<td>2</td>
<td>Laura</td>
<td>Juanito</td>
</tr>
<tr>
<td></td>
<td>Control grammatical form</td>
<td>Target lexical items</td>
</tr>
<tr>
<td></td>
<td>Target grammatical form</td>
<td>Control lexical items</td>
</tr>
</tbody>
</table>

155
<table>
<thead>
<tr>
<th></th>
<th>Control grammatical</th>
<th>Target lexical</th>
<th>Target grammatical</th>
<th>Control lexical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
<tr>
<td>4</td>
<td>Laura</td>
<td>Juanito</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Target grammatical</td>
<td>Control lexical</td>
<td>Control grammatical</td>
<td>Target lexical</td>
</tr>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
<tr>
<td>Noticed+ Silent</td>
<td>1</td>
<td>Juanito</td>
<td>Laura</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Target grammatical</td>
<td>Control lexical</td>
<td>Control grammatical</td>
<td>Target lexical</td>
</tr>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
<tr>
<td>2</td>
<td>Laura</td>
<td>Juanito</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control grammatical</td>
<td>Target lexical</td>
<td>Target grammatical</td>
<td>Control lexical</td>
</tr>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
<tr>
<td>3</td>
<td>Juanito</td>
<td>Laura</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control grammatical</td>
<td>Target lexical</td>
<td>Target grammatical</td>
<td>Control lexical</td>
</tr>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
<tr>
<td>4</td>
<td>Laura</td>
<td>Juanito</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Target grammatical</td>
<td>Control lexical</td>
<td>Control grammatical</td>
<td>Target lexical</td>
</tr>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
<tr>
<td>Noticed Intake</td>
<td>1</td>
<td>Juanito</td>
<td>Laura</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Target grammatical</td>
<td>Control lexical</td>
<td>Control grammatical</td>
<td>Target lexical</td>
</tr>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
<tr>
<td>2</td>
<td>Laura</td>
<td>Juanito</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control grammatical</td>
<td>Target lexical</td>
<td>Target grammatical</td>
<td>Control lexical</td>
</tr>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
<tr>
<td>3</td>
<td>Juanito</td>
<td>Laura</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control grammatical</td>
<td>Target lexical</td>
<td>Target grammatical</td>
<td>Control lexical</td>
</tr>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
<tr>
<td>4</td>
<td>Laura</td>
<td>Juanito</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Target grammatical</td>
<td>Control lexical</td>
<td>Control grammatical</td>
<td>Target lexical</td>
</tr>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
<tr>
<td>Detected Intake</td>
<td>1</td>
<td>Juanito</td>
<td>Laura</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Target grammatical</td>
<td>Control lexical</td>
<td>Control grammatical</td>
<td>Target lexical</td>
</tr>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
<tr>
<td>2</td>
<td>Laura</td>
<td>Juanito</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Control grammatical</td>
<td>Target lexical</td>
<td>Target grammatical</td>
<td>Control lexical</td>
</tr>
<tr>
<td></td>
<td>form</td>
<td>items</td>
<td>form</td>
<td>items</td>
</tr>
</tbody>
</table>
Materials

*Experimental reading passage.*

Two Spanish reading passages, “Juanito” and “Laura”, served as the experimental reading passages (see Appendix B: Experimental Reading Passages for the reading passage and Appendix C: Reading Instructions by Condition for the instructions for each experimental condition). Each reading passage had two versions: one containing the target grammatical form and control lexical items while the second contained the control grammatical form and target lexical items. The “Juanito” reading passage versions included 317 (target grammatical, control lexical) and 318 words (control grammatical, target lexical) and contained 11.32 and 11.44 mean number of words per sentence, respectively. Both versions of the “Juanito” reading passage had
28 sentences. The “Laura” reading passage versions contained 315 (target grammatical, control lexical) and 313 words (control grammatical, target lexical); the mean number of words per sentence was 11.59 and 11.67, respectively.

The topics of both reading passages fit within the scope of what the participants were covering in the Spanish class from which they were recruited and also built upon the vocabulary they had already been exposed to in previous lessons that semester. “Juanito” told the story of a shy and boring college student who goes to a club one night and has an uncharacteristically good time. In the reading passage “Laura,” a private detective solves the case of an interesting client who has lost her “baby;” the baby actually turns out to be the client’s beloved dog.

The placement of the target and control grammatical form and lexical items was handled carefully. For the target grammatical form version and the target lexical items version of each reading passage, there were ten instances of the respective target spaced evenly throughout the reading passage. Likewise, the target grammatical form version contained ten control lexical items and the target lexical items version contained ten instances of the control grammatical form. Each instance of the control grammatical form and control lexical item substituted a respective target. This substitution was done in a way that, in most cases, did not in any way alter any of the surrounding words. In the few cases in which an alteration in the surroundings was necessary in the substitution from the target to the control, the alternation was minimal, such as the change from a masculine indefinite article to the feminine indefinite article. The first and last sentences of each reading passage contained no targets or controls. No sentence contained more than one target or control, regardless of whether it was grammatical or lexical, however some sentences contained a target and a control. All instances of the targets and controls appeared in the middle third of the sentence. No glosses were provided.
The reading of the reading passage occurred via computer. Each reading passage was divided into six screens so that each screen contained approximately four to five sentences. No sentences were split between two screens. Participants advanced through each screen by pressing the space bar; once participants advanced a screen, they were unable to return to previous screens. This division of the reading passage served to discourage backtracking as much as possible. Each screen also included a counter showing which number screen the participant was currently on in relation to the total number of screens of that reading passage (e.g., 2/6). A short drift correction screen appeared between each screen.

The instructions for the reading passage changed according to the level of intake condition and thus facilitated the characteristics of each condition, as described in Table 6.
As discussed later in the chapter, the data provided by the participants was then examined and assigned into intake groups according to time spent on reading.

Crucially, only the Noticed+ TA condition thought aloud while reading the reading passages.

**Pretest and immediate posttest assessments.**

In order to determine if the target grammatical form and target lexical items were taken in, participants completed a series of pretests and immediate posttests. The pretests included a verb stem meaning pretest and an item pretest of both the target grammatical form and target lexical items. The immediate posttests included a recognition posttest and a comprehension posttest. All assessments were pen-and-paper. There were no time limits on any tests, however participants recorded the time they spent on each test.

*The controlled written production pretest and immediate posttest.*

The controlled written production pretest and immediate posttest were exactly the same with the exception of item order (items on the immediate posttest were randomized). The purpose of both assessments was to determine whether participants had any production ability of the target grammatical form and target lexical items. The format of the controlled written production pretest and controlled written production immediate posttest was a column of the English translations of the instances of the target grammatical form, target lexical items, and distracters; instructions told participants to write the Spanish translation of the English word in the corresponding row. All English translations of verbs included the Spanish infinitive in parenthesis; this was included because the reason for testing the target grammatical form was to determine if participants knew the 3rd person singular preterit of regular -ar verbs, not if they knew the meaning of the verb. Instructions told participants not to guess or backtrack to pages
that they had already completed. The immediate posttest version included in the instructions that participants should provide the Spanish translation of the English words, “as they may have appeared in the reading,” thus orienting participants to the reading passages.

The assessments consisted of 100 items, 20 of which were instances of the target grammatical form and 20 of which were instances of the target lexical items. The remaining 60 items were distracters. Of these distracters, 30 were singular nouns and 30 were verbs. Of the verbs, 20 were preterit (four 1st person singular –ar verbs, three 2nd person singular –ar verbs, one 2nd person singular –er verbs, four 3rd person singular –er verbs, two 3rd person singular –ir verbs), five were present (one 1st person singular –ar verb, one 2nd person –ar verb, one 3rd person singular –ir verb, 1 1st person plural –ar verb, one 3rd person plural –er verb), and five were future (one 1st person singular verb, two 2nd person singular –ar verbs, one 2nd person singular –er verb, one 3rd person plural –ar verb).

*The items pretest.*

In order to minimize drawing participants’ attention to the target grammatical and lexical items, the grammatical and lexical items pretest consisted of a list of all grammatical and lexical targets that appeared in all versions of the experimental reading passage (40) together with 80 distracter words, totaling 120 words (see Appendix D: Items Pretest). Instructions told participants that for each word, they should mark the most appropriate column: whether they have never seen the word before, whether they have seen the word but do not know the meaning, or whether they know the meaning. They were also prompted to write the meaning of the word in the case that they did know it; this was to determine if they actually knew the meaning of the word or were mistaking it with another word. The purpose of this pretest was to determine
whether participants had any receptive prior knowledge of the target grammatical form and the target lexical items before participation in the study.

The distracters included a combination of 40 nouns and 40 verbs, many of which include written accents in an attempt to best conceal the target grammatical form. More specifically, of the 40 distracter nouns, 14 had accents on the final letter (5 –ú, 2 –é, 7 –í) and seven had word-internal accents. Regarding the distracter verbs, ten were the 1st person singular present form of the same instances of the target grammatical form appearing in the “Laura” reading passage. Thirty additional distracter verbs were a combination of 1st, 2nd, and 3rd person singular and plural forms of the present indicative, present subjunctive, pluperfect, and future. There was an even mixture of –ar, -er, and –ir verbs. Of the 40 distracter verbs, three had word-final accents and seven had word-internal accents.

The verb stem meaning pretest.

The purpose of this pretest was to ensure that participants knew the meanings of the verb stems for all target grammatical items. In other words, it was essential that participants already knew the meaning of the infinitive forms of the target grammatical items but not target grammatical form (third person singular form of regular –ar verbs in the preterit) prior to exposure to the experimental reading passage. The verb stem meaning pretest included a review of the meanings of the infinitive forms of the verbs used to express the target grammatical form and also a test of the meanings. First, the review consisted of a presentation of the infinitive of each of the verbs used to express the target grammatical form (20) alongside the English translation; these infinitives appeared alongside 20 distracter –ar infinitive verbs. The distracter words were infinitive verbs taken from the first two chapters of the textbook used in the Spanish class in which participants were enrolled. The assessment portion consisted of four sets of
matching tests: each test had a list of 10 of the distracters and target, and participants had to match each with the correct English translation in a separate column. The assessment was broken down into four sets so as to avoid having an overwhelming matching task of 40 items. Any participant not scoring 100% on the 20 targets during the first administration of the verb stem meaning pretest took the test again until obtaining a perfect score.

*Immediate recognition posttest.*

This posttest aimed to determine if participants were able to recognize the instances of the target grammatical form and the target lexical items that they were exposed to in the experimental reading passage, or in other words, to determine which forms and items they took in. Therefore, this is the measure that directly tested recognition of the target grammatical form and target lexical items. The immediate recognition posttest contained the same items as the items pretest, however all items were randomized. An additional difference is that instead of indicating whether have never seen, have seen but don’t know the meaning, or have seen and known the meaning, the immediate recognition posttest had participants mark whether each word “DID appear in text,” “DID NOT appear in text,” or “not sure” (see Appendix E: Recognition Posttest).

In addition to the target grammatical form and the target lexical items, the immediate recognition posttest also included the control grammatical form and the control lexical items.

As on the items pretest, there were ten distracter verbs ending in –o, all of which were simply first person singular present forms of target –ó (third person singular preterit). These were included to provide a contrast to the final –ó of the target grammatical form; an analysis of how many –o distracters were recalled by participants compared with how many –ó targets were
recalled could show that participants were actually recalling the --ó, the target grammatical form, instead of just the final --o.

**Comprehension posttest.**

The comprehension posttest was a multiple-choice task with four options per item (see Appendix F: Comprehension Posttest and Appendix G: Comprehension Posttest Answer Sheet). Participants completed sentences by choosing the option that best reflected the content of the reading passage. This task aimed to determine how much of the reading passage the participants comprehended and thus ascertain whether they actually read/skimmed for meaning, as instructed by the instructions of the reading passage. The comprehension posttest was administered in English to ensure that the language of task did not interfere with comprehension. There was one comprehension question specifically addressing each sentence in which an instance of the target grammatical form and target lexical items appeared. In addition to the 20 questions addressing the 20 instances of the target grammatical form and target lexical items in each reading passage, there were also two global comprehension questions per reading passage for a grand total of 44 comprehension questions.

The comprehension questions were phrased in a way that tied the comprehension directly to the instance of the target grammatical form or the target lexical item: the correct answer to the question was always the target. However, in an attempt to avoid the sentential comprehension being dependent on the comprehension of the target, the multiple-choice options of each comprehension question listed three incorrect (yet plausible, given the reading passages) answers and one “none of the above.” Therefore, all the participant had to know to answer a comprehension question correctly is that the answer was not one of the three options listed, but rather was another option (“none of the above”). This way, the participant could still get the
comprehension question correct without having comprehended the target. However, the global comprehension questions had the correct answer listed as an option.

Questionnaires.

Debriefing questionnaire.

A debriefing questionnaire administered after the posttest assured that prior knowledge and/or recognition of the target grammatical and lexical forms did not contaminate the data (see Appendix H: Debriefing Questionnaire). The form was the same as that used in Leow (1993, 1995). It provided the target grammatical form and target lexical items and asked participants if they had any external exposure to these items outside of the experimental environment and during the experimental period, that is, between the pretest and the posttest.

More specifically, the debriefing questionnaire stated that the study investigated the third person singular form of preterit –*ar* regular verbs. Participants then indicated whether they did or did not “know/recognize this form before doing the exercises.” If a participant indicated that s/he already knew/recognized the target grammatical form before participating in the study, then that participant’s data was investigated further to see if their data should be excluded (details following). In addition, a list of all target lexical items and their English translations was provided; participants were asked to indicate for every item whether they had never seen the word before reading the passage, had seen the word before reading the passage but didn’t know what it meant, or had seen the word before reading the passage and also knew what it meant. If they indicated that they knew what it meant, then they were prompted to write the meaning next to it.
Language contact profile.

A shortened version of Freed, Dewey, Segalowitz, and Halter’s (2004) Language Contact Profile was also used to ensure that English was indeed a first language of all participants and that participants had not had extensive experience with the Spanish language even though they were enrolled in a beginning level class; participants also should not have had more than two semesters of study of any other Romance language.

The version of the Language Contact Profile used in this study included questions regarding country of birth, first language(s), home language(s), study abroad and living abroad experiences, self-rating in English and Spanish (listening, speaking, reading, writing on a 4-point scale), years spent studying Spanish, and communication in Spanish to native Spanish speakers.

Practice tasks.

Practice TA protocol task.

A practice TA protocol task provided participants in the deeper processing TA condition with practice in thinking their thoughts aloud (see Appendix I: Practice Think-aloud Protocol Task). A math word problem was used, given its ability to naturally facilitate thinking one’s thoughts aloud. The practice TA protocol task asked participants, in English, to determine the total cost of several groceries on a grocery list while talking aloud and refraining from writing anything down.

Practice eye-tracking task.

A practice reading task helped familiarize participants with the act of reading while having their eye movements registered (see Appendix J: Practice Eye-tracking Task). Participants were instructed to read five sentences of a Spanish reading passage while having their eye movements tracked. Participants were also instructed to follow the same instructions as
those that corresponded to their experimental condition. In other words, this task was the same as the experimental reading passages for each condition, except that the reading passage was shorter and on a different topic.

**Procedure**

The current study took place over two sessions (see Table 8). The first session consisted of the pretests; participants returned one week later to complete the experimental session followed immediately by the posttests. The pretest sessions took place in a large computer lab and the experimental session occurred in a small eye-tracking lab.

One week prior to exposure to the experimental treatment, all participants reported to the language lab to complete the pre-tests (generalized production pretest, items pretest, verb stem meaning pretest) in addition to the language contact profile.

On the day of the experimental treatment, participants were randomly assigned to an experimental condition upon reporting to the language lab. The experimental treatment was administered individually to one participant at a time. Participants first signed the Institutional Review Board Waiver. The researcher then explained to participants how to position themselves while using the video-based SR-Research EyeLink 1000 eye-tracker and also performed the calibration of the eye-tracking equipment (only for the noticed, detected, and attended intake conditions). Next was the administration of the practice think-aloud protocol (only for the deeper processing TA condition) and practice eye-tracking task (only for the noticed, detected, and attended intake conditions, see Appendices I and J). Participants then read the experimental reading passage. Eye movements were recorded for the noticed, detected, and attended intake conditions, while the two noticed+ conditions read the reading passages on the computer without having their eye movements tracked. The concurrent verbal reports for the Noticed+ TA
condition were audio-recorded. Participants then completed the controlled written posttest, the recognition post-test, and the comprehension posttest.

Table 8: Methods of Experimental and Control Conditions

<table>
<thead>
<tr>
<th>Condition 1: Noticed intake</th>
<th>Session 1 (pretests)</th>
<th>Session 2 (experimental exposure and immediate posttests)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. Controlled written production pretest</td>
<td>1. IRB form</td>
</tr>
<tr>
<td>Condition 2: Detected intake</td>
<td>2. Items pretest</td>
<td>2. Eye-tracking calibration</td>
</tr>
<tr>
<td></td>
<td>4. Language contact profile</td>
<td>4. Experimental reading passage with eye-tracking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5. Controlled written production posttest</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6. Recognition posttest</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7. Comprehension posttest</td>
</tr>
</tbody>
</table>

| Condition 4: Noticed+ TA    | 1. Controlled written production pretest | 1. IRB form |
|                             | 2. Items pretest                       | 2. Think-aloud practice                                |
|                             | 3. Verb stem meaning pretest           | 3. Experimental reading passage with think-aloud (no eye-tracking) |
|                             | 4. Language contact profile            | 4. Controlled written production posttest               |
|                             |                                           | 5. Recognition posttest                                 |
|                             |                                           | 6. Comprehension posttest                               |

| Condition 5: Noticed+ Silent| 1. Controlled written production pretest | 1. IRB form |
|                             | 2. Items pretest                       | 2. Experimental reading passage (no eye-tracking and no think-aloud) |
|                             | 3. Verb stem meaning pretest           | 3. Controlled written posttest                           |
|                             | 4. Language contact profile            | 4. Recognition posttest                                 |
|                             |                                           | 5. Comprehension posttest                               |

**Coding and Scoring**

All pre- and posttests were scored and eye-tracking data was extracted from the eye-tracker. The concurrent verbal reports of the Noticed+ TA condition were transcribed and the resulting data were coded for depth of processing. The debriefing questionnaire, although not coded, confirmed either the presence or absence of prior knowledge of the targets. The following is a detailed explanation of the scoring, coding, and assignment into intake groups.

**Pre- and posttest measures.**

All pre- and posttests (controlled written pretest/posttest, items pretest, recognition posttest, verb stem meaning pretest, comprehension posttest) were scored; one correct answer equaled one point (distracters excluded) and incorrect answers were awarded zero points. Due to
the Latin square design of the study that exposed each participant to only half of the entire set of instances of the target grammatical form and target lexical items, for the controlled written pretest/posttest, items pretest, and the recognition posttest, only the target items to which each participant was exposed in their condition were scored. The other target items that they did not see due to their condition were not scored. The same was true for the control grammatical form and the control lexical items on the controlled production written pre/posttests, the items pretest, and the recognition posttest: participants were awarded one point for each correct answer of a control, but only for the controls that belonged to their condition. Similarly, for the comprehension posttest, only target items to which each participant was exposed in their condition were scored, in addition to both overall comprehension question for each reading passage. For the verb stem meaning pretest, all 20 instances of the target grammatical form were scored, since it was essential that participants knew the meaning of the infinitive verb form of the target for comprehension of the reading passages.

The controlled production pretest and the items pretest only served to confirm that participants had no prior knowledge of the target grammatical form or the target lexical items. If a participant indicated prior knowledge of a target lexical item on the items pretest (“I have seen this word before and know its meaning”) that corresponded to the targets belonging to their condition, and if s/he correctly produced that item as prompted with no more than two letters varying from the correct orthography, then it was considered that the participant possessed correct prior knowledge of that item and it was subsequently removed from the statistical analysis for that participant. Regarding the target grammatical items, if a participant correctly identified -ó as being the 3rd person singular form of the preterit on the controlled production pretest or the items pretest, then all that participant’s data was excluded from statistical analysis.
Because the grammatical target form and the target lexical items were considered two separate types of targets in this study, each assessment provided at least two scores: a target grammatical form score and a target lexical score. In addition, the controlled written production pre/posttest, the items pretest, and the recognition posttest all had two additional scores: a control grammatical form score and a control lexical score.

Criteria for correct answers.

In determining the correctness of answers on all production tests, a two-letter window of error was allowed. More specifically, participants could incorrectly write up to two letters and still be awarded the one point for a correct answer; incorrect spelling totaling three or more incorrect letter was regarded as incorrect and no points were assigned. For example, the target lexical item *parentela* was awarded one point on the production tests and considered to be correctly known if it were spelled *perentela* (1 incorrect letter) or *prentelo* (1 incorrect letter, 1 missing letter) but it was not considered correct if it was spelled *perenta* (1 incorrect letter, 2 missing letters) or *perentalo* (3 incorrect letters).

Additionally, the researcher decided to accept certain synonyms for select targets on the controlled written production pre- and posttests: ‘restaurant’ was accepted for *mesón*, ‘vacation’ and ‘free day’ for *feriado*, ‘ID’ for *carnet*, ‘talked’ for *charló*, ‘street’ for *paseo*, and ‘cabinet’ for *fichero*.

Criteria for inclusion of data in statistical analysis.

The language background questionnaire initially screened participants for characteristics that might cause their data to be excluded from the data analysis (speak English as a native language; speak no other language with a self-rated proficiency of greater than 1 on scale of 0-3; have taken no more than two high/middle/elementary school years of Spanish; never have
studied abroad in a Spanish-speaking country or have traveled in a Spanish-speaking country with the purpose of learning Spanish; and complete both experimental sessions), and there were also additional criteria that participants had to meet via their performance on the pre- and posttests in order to have their data included in the data analysis. If a participant demonstrated correct prior knowledge of the target grammatical form (operationalized as at least one correct production of the target grammatical form) on the controlled production pretest, all of their data (grammatical and lexical) was excluded from the data analysis. If a participant showed correct prior knowledge of a target lexical item on that same pretest, all data regarding that specific target lexical item from all assessments was excluded. Regarding the items pretest, a participant indicating that s/he knew the meaning of a grammatical target had all their data excluded from data analysis only if s/he was correctly able to produce that word. If the participant indicated that they knew the meaning of the word but s/he did not correctly produce the meaning, then the data was included in the analysis. Regarding target lexical items, the same criteria applied, however in the case that the participant demonstrated correct prior knowledge of a word, all data on only that individual word was excluded from data analysis.

The debriefing questionnaire was used for triangulation purposes and assisting in determining if a participant had any correct prior knowledge of the target grammatical form or the target lexical items before exposure to the reading passage. However, more importance was placed on the controlled production pretest and the items pretest: if a participant indicated prior knowledge of the target grammatical form or the target lexical items on the debriefing questionnaire, that knowledge must also have been demonstrated on either the controlled production pretest or the items pretest in order for their data to be excluded. In other words, if a participant indicated prior knowledge of the target grammatical form or a target lexical item on
the debriefing questionnaire but did not indicate prior knowledge of the same item on the controlled written pretest or the items pretest, that that item remained in the participant’s analysis.

**Assignment into grammatical and lexical intake groups.**

As described earlier, each participant was randomly assigned to an experimental condition (noticed, detected, attended, noticed+ TA, noticed+ silent), however the time each participant spent on each of the two readings determined to which grammatical intake group and lexical intake group s/he was assigned. Due to the Latin Square design of this study, each reading passage for each participant was either grammatical target/lexical control or grammatical control/lexical target. Given this, it was possible to use the time spent on the reading passage with the grammatical target form to represent the time spent reading for grammatical intake group placement, and time spent on the reading passage with the lexical target form to represent the time spent reading for the lexical intake group placement\(^26\). This was done by roughly dividing all the grammatical reading passage times and all the lexical reading passage times in three; the top third for the grammatical reading passage times was assigned to the Grammatical Noticed Intake Group while the top third for the lexical reading passage times was assigned to the Lexical Noticed Intake Group. Likewise, the middle third for each reading passage times was assigned to the Grammatical Detected Intake Group/Lexical Detected Intake Group, and the bottom third was assigned to the Grammatical Attended Intake Group/Lexical Attended Intake Group. An exact one-third division was not always possible due to some participants spending the same

---

\(^26\) Time spent reading the entire experimental reading text was used to place participants into intake groups because time spent on the entire text arguably reflects whether participants read the text according to the instructions of their condition. If they read the text according to their experimental condition, then it is being assumed that they also read every individual word of the text, including instances of the target grammatical form and target lexical items, according to condition’s instructions.
amount of time on readings, and it was also desirable to have a gap of at least a few seconds between groups, so the one-third division was made flexibly. Descriptive statistics regarding time spent on the reading passages and group assignment is available in Table 9.

Table 9: Time Spent (in Seconds) on Reading Passages

<table>
<thead>
<tr>
<th></th>
<th>Grammatical</th>
<th>Lexical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Noticed+ TA Intake Group</td>
<td>19</td>
<td>461.54 (149.77)</td>
</tr>
<tr>
<td>Noticed+ Silent Intake Group</td>
<td>19</td>
<td>257.14 (138.09)</td>
</tr>
<tr>
<td>Noticed Intake Group</td>
<td>19</td>
<td>99.06 (20.40)</td>
</tr>
<tr>
<td>Detected Intake Group</td>
<td>20</td>
<td>76.25 (4.61)</td>
</tr>
<tr>
<td>Attended Intake Group</td>
<td>19</td>
<td>57.95 (7.34)</td>
</tr>
</tbody>
</table>

All participants in the noticed+ TA condition and the noticed+ silent condition were assigned to the Grammatical Noticed+ TA Intake Group/Lexical Noticed+ TA Intake Group and the Grammatical Noticed+ Silent Intake Group/Lexical Noticed+ Silent Intake Group because the experimental design of eye-tracking, think-alouds, and non-think-aloud characteristics did not allow them to change, but this was not an issue because their mean times on the reading passages were, in fact, considerably longer than those of the lower intake groups. Any participant who did not record their reading times was simply left in the same level of intake group as their intake level for the experimental condition; this happened in 13 cases (of a total of 96; 14%) for both the grammatical and lexical intake groups.

Since the noticed intake condition was the only condition in which participants were instructed to click on textually-enhanced targets and controls, there is only limited data on
clicking. This data was used to verify correct assignment into intake groups: all participants who were originally in the noticed intake condition, had clicking data, and who were also assigned to the Grammatical and/or Lexical Noticed Intake Group, had to have clicked on at least 50% of all textually enhanced words in order to be assigned to this group. Since clicking was not an instruction for the other intake conditions, there was no clicking data for the detected and attended experimental conditions so it was not possible to use this same criteria for participants from these two conditions who were assigned to the Grammatical and/or Lexical Noticed Intake Group.

As shown in Table 9, nine participants in the noticed condition were assigned to the Grammatical Noticed Intake Group based on their reading times; this means that clicking data was available for these nine participants. An additional 12 participants in the noticed condition were assigned to the Lexical Noticed Intake Group, but clicking data was only available for 11 of these 12 participants due to computer malfunction. These nine and eleven participants, respectively, met the criteria of clicking on least 50% of all textually enhanced words. In the Grammatical Noticed Intake Group, the nine participants clicked an average of 37.67 textually-enhanced words out of a possible 40 (SD = 2.65). In the Lexical Noticed Intake Group, the eleven participants clicked an average of 38 words out of a possible 40 (SD = 2.49). Therefore, with the exception of the one participant whose clicking data was unavailable due to computer malfunction, all participants who were originally in the noticed condition and who were also assigned to the Grammatical and/or Lexical Noticed Intake Groups met the criteria of clicking on last 50% of the textually-enhanced words to be in the Grammatical and/or Lexical Noticed Intake Groups.
Depth of processing think-aloud protocols.

The current study expands on the operationalizations of depth of processing from previous studies (i.e., Leow et al., 2008; Morgan-Short, et al., 2012; Qi & Lapkin, 2001) by employing a more current operationalization from Leow (forthcoming) that uses two separate sets of criteria to describe depth of processing: one for describing depth of processing of a grammatical form and one for describing depth of processing of a lexical item. Generally, depth of processing is understood to be the amount of attention, cognitive effort, or time spent processing or elaborating on the target item in the input, and/or noticing induced by different types of experimental tasks or conditions (cf. e.g., Robinson, 2003; Hulstijn, 2001; Laufer & Hulstijn, 2001; Leow, Hsieh, & Moreno, 2008) or, more simply, the process of making a form-meaning connection. However, given the apparent differences between grammatical forms and lexical items (i.e., Greenslade et al., 1999; Mackey et al., 2000; Smith, 2012), it seems prudent to establish separate criteria.

Regarding the actual coding of the think-aloud protocols, Leow’s (forthcoming) descriptions were used for each level and his criteria were used as the basis for the criteria of this particular study. Table 10
depicts the current study’s criteria and provides specific examples of each.

Think-aloud protocols from the Grammatical Noticed+ TA Group and the Lexical Noticed+ TA Group were transcribed and then the target lexical items and the instances of the target grammatical form for each participant were found within the transcripts. All were then coded according to the criteria in Table 10. There are three possible depths of processing of a grammatical form: low, intermediate, and high. A low depth of processing is described as a verbalization that shows no potential for processing the target form grammatically. An intermediate depth of processing is characterized by comments on the target item in relation to the grammatical features; a high depth of processing requires that the learner evidences arriving at either an accurate or inaccurate underlying grammatical rule. Likewise, there are also three levels (low, intermediate, and high) of depth of processing of a lexical item. While a low depth of processing of a lexical item involves an utterance that shows no potential for any form-meaning connection, an intermediate depth of processing is characterized by an utterance that provides some evidence of processing. An utterance with a high depth of processing provides evidence of making a form-meaning connection, whether it be accurate or inaccurate.

Since the analyses were performed at an item level (see Chapter Three for more details), no total scores were used.
Table 10: Criteria for Grammatical and Lexical Depths of Processing

<table>
<thead>
<tr>
<th>Depth of Processing of a Grammatical Form</th>
<th>Low</th>
<th>Intermediate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Shows no potential for processing target form grammatically</td>
<td>Comments on target item in relation to grammatical features</td>
<td>Arrives at an accurate or inaccurate underlying grammatical rule</td>
</tr>
<tr>
<td>Descriptors with Specific Examples</td>
<td>• reads target: <em>y pensó en el caso de la señora</em> (participant 021)</td>
<td>• correctly translates the target into the past, but gets the person wrong: <em>naturally I went to bed</em> (went target is <em>se acostó</em>, participant 023)</td>
<td>• translates the target into the past: <em>naturally went he went to bed immediately</em> (when target is <em>se acostó</em>, participant 034)</td>
</tr>
<tr>
<td></td>
<td>• repeats target: <em>A las seis de la tarde, lavó su lavó su ropa favorita</em> (participant 017)</td>
<td></td>
<td>• correctly translates target, even if lexical meaning of the verb is incorrect: <em>Laura was happy and exchanged</em> (when target verb is <em>caminó</em>, participant 037)</td>
</tr>
<tr>
<td></td>
<td>• translates only lexical meaning of target, not grammatical meaning: <em>he’s eating breakfast</em> (when target is <em>desayunó</em>, participant 038)</td>
<td></td>
<td>• translates target, even if it is said as one of several options: <em>no se quedó en como casa como siempre he stayed in the house on Saturdays or he doesn’t stay</em> (when target is <em>se quedó</em>, participant 032)</td>
</tr>
<tr>
<td></td>
<td>• says s/he isn’t sure what it is: <em>llegó una señora muy elegante y obviamente muy rica. Um I’m not sure what that means</em> (participant 033)</td>
<td></td>
<td>• shows that a hypothesis has been made by means of using a tense other than the present indicative: <em>hmm I will talk no more</em> (when target is <em>habló</em>, participant 060)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Depth of Processing of a Lexical Item</th>
<th>Low</th>
<th>Intermediate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>Shows no potential for emerging form-meaning connection</td>
<td>Provides some evidence of processing target item</td>
<td>Provides evidence of making some form-meaning connection (accurate or inaccurate)</td>
</tr>
<tr>
<td>Descriptors</td>
<td>• simply reads target: <em>y tiene un negocio que...</em> (participant 017)</td>
<td>• processes deeper by referring to a past event or activity when they saw the form: <em>y mi apellido I’ve seen that word apellido I can’t remember right now</em> (participant 012)</td>
<td>• evidences that s/he knows the meaning of the word, even if it is inaccurate <em>El paseo más. El paseo más prestigioso. Her office must be located in a prestigious part of Miami.</em> (participant 042)</td>
</tr>
<tr>
<td></td>
<td>• says s/he doesn’t know what it is: <em>um fichero I don’t know what that means</em> (participant 051)</td>
<td></td>
<td>• achieves a very similar meaning to the target that is almost interchangeable: <em>slept a little bit more</em> (when target is <em>cabezada</em>, participant 034)</td>
</tr>
<tr>
<td></td>
<td>• repeats target: <em>un traje formal y zapatos negros</em> <em>un traje</em> (participant 072)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

27 The target has been bolded.
The researcher coded the think-aloud protocols and a colleague separately coded all protocols in order to determine an interrater agreement index. An interrater reliability analysis using the Kappa statistics was performed to determine consistently among the two raters. The interrater reliability for the raters was found to be Kappa = 0.84 (p < 0.01), 95% CI (0.77, 0.91). The total number of all cases was n = 233. According to Landis and Koch (1977), this score is considered almost perfect agreement. The researcher and colleague discussed the cases of disagreement (20) and resolved the differences so that total agreement was reached on every coding from every think-aloud protocol (Kappa = 1.00, p = 0.00, 95% CI (0.00, 0.00).

**Fixation measures.**

First fixation duration, gaze duration, second pass time, and total fixation time are all dependent variables in the current study for the specific purpose of determining if there is a relationship between any of the measures and recognition for the eye-tracking groups (attended, noticed, and detected intake groups). Using Data Viewer, by S-R Research, the value of each of these measures was determined for each instance of the target grammatical structure and target lexical items. The interest area for the target grammatical form was the entire verb form, not just the inflection. The interest area of the target lexical items was also at the word level. For participants who were placed via random assignment into the attended intake experimental condition, the fixation times recorded were those during which participants were looking at the masked target word (i.e., Xs).
Chapter Four: Results

This chapter reveals the results of the seven research questions stated in chapter one, in addition to statistical information necessary to ensure the validity of the methodology. The first part of this chapter reviews some short notes regarding the statistics, followed by the additional statistical information to ensure validity. The rest of the chapter is dedicated to the results of each research question, presented in numerical order.

Statistical Notes

Why percentages are used for lexical items.

All raw data of posttest measures for the Lexical Groups were first converted to a percent and then submitted to the respective statistical analyses. This was necessary due to the fact that several participants demonstrated prior knowledge (on the pretest or the debriefing questionnaire, see the chapter on methodology for details) of one or more of the target lexical items. Any previously-known target lexical item was removed from that participant’s analyses, so although each participant was exposed to ten target lexical items, some participants’ scores were based on only, for example, nine target lexical items (in the case that they demonstrated prior knowledge of one target lexical item). Therefore, it was determined that the best way to remedy this was to use a percent value consisting of the average score of all previously-unknown target lexical items divided by that participant’s number of previously-unknown target lexical items (See Figure 3). Percent lexical values were computed for each participant for the following pretest and posttest measures: Controlled Production Pretest, Items Pretest, Controlled Production Posttest, Recognition Posttest, and Lexical Comprehension Posttest. Therefore, the values that were actually submitted to the statistical analyses were, respectively, Percent Controlled Production
Pretest, Percent Items Pretest, Percent Controlled Production Posttest, Percent Recognition Posttest, and Percent Lexical Comprehension Posttest.

Figure 3: Calculation of subtracted fixation values.

Participant X’s = Average score of instances of (grammatical or lexical) target – Average scores Subtracted Y Test Score of instances of (grammatical or lexical) control

**Why subtracted fixation values are used for eye-tracking measures.**

In order to achieve a more valid measure of the amount of attention that learners paid to the grammatical and lexical targets, subtracted fixation measures were calculated for eye-tracking measures. This technique was also employed in Godfroid and Uggen (2013). These percent values were calculated by taking each participant’s average value across instances of the target (first the instances of the target grammatical form and then the target lexical items) for each eye-tracking measure and subtracting the respective average value for instances of the control. The resulting value is positive if the participant fixated longer on the target and negative if s/he fixated longer on the control. If such a procedure had not been used in the present study, comparing target fixation measures among participations would have been problematic because there would have been no baseline against which to compare fixations on the targets.

Figure 4: Calculation of Percent Lexical Values

Participant X’s Percent Y test score = \[
\frac{\text{Average score of all previously-unknown target lexical items on Y test}}{\text{Participant X’s number of previously-unknown target lexical items (maximum value=10)}}
\]

Therefore, the following eye-tracking measures reflect this subtraction procedure:

Subtracted Grammatical Total Duration, Subtracted Grammatical First Fixation Duration,
Subtracted Grammatical Gaze Duration, Subtracted Grammatical Second Pass Time, Subtracted Lexical Total Duration, Subtracted Lexical First Fixation Duration, Subtracted Lexical Gaze Duration, and Subtracted Lexical Second Pass Time.

Tests of Validity

Text order.

Since approximately half the participants saw Text A followed by Text B and the other half saw Text B followed by Text A, it is necessary to ensure that text order did not affect test scores on the posttest measures. Grammatical and lexical scores from eight posttest measures (Controlled Production Posttest, Recognition Posttest, Comprehension Posttest, Generalized Production Posttest, Subtracted Total Duration, Subtracted First Fixation Duration, Subtracted Second Pass Time, Subtracted Gaze Duration) were separately submitted to a multivariate ANOVA with Text Order as the independent variable and the eight posttest measures as the dependent variables. This ANOVA was run ten times: once for each of the grammatical intake groups and once for each of the lexical intake groups. Each time the ANOVA was run for the grammatical intake groups, only the grammatical scores of the eight posttest measures were used; likewise, each time the ANOVA was run for the lexical intake groups, only the lexical scores of the eight posttest measures were used. Unless otherwise noted, all analyses were run on the Statistical Package for the Social Sciences (SPSS, Version 22) and the alpha value was set at .05 for statistical significance. All results show no statistically significant effect of text order on neither grammatical nor lexical posttest scores.

Were posttest scores higher for target words than for control words?

As described in the methodology chapter, control words were included for both the target grammatical form and target lexical items. One benefit to this is the opportunity to compare
posttest scores of target versus control words to determine whether scores for targets (grammatical and lexical) were in fact higher than scores for controls (grammatical and lexical).

Paired samples t-tests were run for each intake group for both the target grammatical form and the target lexical item. The scores compared were the target versus the control score for the following posttest measures: Recognition Posttest, Total Duration, First Fixation Duration, Second Pass Time, and Gaze Duration. First presented are the grammatical target results for each intake group for each of the posttest measures above, followed by the lexical target results, also in the same order. It is important to note that due to a mistake made by the researcher, no control lexical items were included in the Recognition Posttest, so lexical results from that posttest are missing.

**Grammatical intake groups.**

Paired-samples t-tests comparing scores on instances of the target grammatical form and instances of the control grammatical form on the Recognition Posttest returned statistically significant for all intake groups except the Grammatical Detected Intake Group; this group was approaching statistical significance (see Table 11). This means that in all grammatical intake groups except the Grammatical Detected Intake Group, participants scored significantly higher on instances of the target grammatical form than they did on instances of the control grammatical form on the Recognition Posttest. In regards to the eye-tracking measures, the paired-samples t-test to compare target scores (M = 155.38, SD = 138.18) and control scores (M = 112.96, SD = 108.95) for instances of the grammatical target form on Second Pass Time returned statistically significant; t(19) = 2.11, p = .05. In other words, the Grammatical Detected Intake Group had

---

28 Note that the percent eye-tracking measures were not used for the analyses whose objective was to determine whether scores and values for targets were significantly different than those for controls because the percent eye-tracking measures are actually a composite of the target and control.
significantly longer Second Pass Times on instances of the target grammatical form than they did on instances of the control grammatical form. No other comparisons for eye-tracking measures for any grammatical intake group returned significant. This means that with the exception of the Grammatical Detected Intake Group, all grammatical intake groups fixated on the target grammatical form for a similar amount of time as they did on the control grammatical form. This provides evidence that there were no inherent aspects of the target grammatical form that caused participants to fixate longer on them than on the control grammatical form. If there had been, it would have been possible that other aspects of the target grammatical form, such as the suffix –ó, was causing the participants to fixate longer and therefore would have affected any results.

Table 11: Summary Results for Paired-Samples T-Tests for Instances of Target vs. Control Grammatical Forms

<table>
<thead>
<tr>
<th>Grammatical Intake Group</th>
<th>Measure</th>
<th>95% CI</th>
<th>Mean Value Target (SD Target)</th>
<th>Mean Value Control (SD Control)</th>
<th>N1/N2</th>
<th>t</th>
<th>p</th>
<th>effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical Noticed+ TA Intake Group</td>
<td>Recognition Posttest</td>
<td>5.06, 7.14</td>
<td>8.50 (.97)</td>
<td>2.40 (2.01)</td>
<td>10</td>
<td>13.31</td>
<td>.00*</td>
<td>.95</td>
</tr>
<tr>
<td>Grammatical Noticed+ Silent Intake Group</td>
<td>Recognition Posttest</td>
<td>1.75, 6.65</td>
<td>7.00 (2.16)</td>
<td>2.80 (2.66)</td>
<td>10</td>
<td>3.88</td>
<td>.00*</td>
<td>.63</td>
</tr>
<tr>
<td>Grammatical Noticed Intake Group</td>
<td>Recognition Posttest</td>
<td>2.27, 7.51</td>
<td>7.33 (1.87)</td>
<td>2.44 (3.13)</td>
<td>9</td>
<td>4.30</td>
<td>.00*</td>
<td>.70</td>
</tr>
<tr>
<td></td>
<td>Total Duration</td>
<td>-71.78, 48.04</td>
<td>407.78 (151.32)</td>
<td>419.65 (135.58)</td>
<td>19</td>
<td>-1.16</td>
<td>.26</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>First Fixation Duration</td>
<td>30.61, 8.84</td>
<td>226.21 (57.96)</td>
<td>237.10 (44.54)</td>
<td>19</td>
<td>-1.16</td>
<td>.26</td>
<td>.07</td>
</tr>
<tr>
<td></td>
<td>Second Pass Time</td>
<td>-32.71, 74.95</td>
<td>153.22 (133.07)</td>
<td>132.10 (124.32)</td>
<td>19</td>
<td>.82</td>
<td>.42</td>
<td>.04</td>
</tr>
<tr>
<td></td>
<td>Gaze Duration</td>
<td>-42.47, 36.56</td>
<td>323.61 (108.37)</td>
<td>326.57 (81.27)</td>
<td>19</td>
<td>-1.16</td>
<td>.88</td>
<td>.01</td>
</tr>
<tr>
<td>Grammatical Detected Intake Group</td>
<td>Recognition Posttest</td>
<td>-.14, 3.83</td>
<td>6.23 (1.69)</td>
<td>4.38 (3.31)</td>
<td>13</td>
<td>2.03</td>
<td>.07</td>
<td>.26</td>
</tr>
<tr>
<td></td>
<td>Total Duration</td>
<td>-17.77, 82.38</td>
<td>344.75 (199.59)</td>
<td>312.45 (179.02)</td>
<td>20</td>
<td>1.35</td>
<td>.19</td>
<td>.09</td>
</tr>
<tr>
<td></td>
<td>First Fixation Duration</td>
<td>Second Pass Time</td>
<td>Gaze Duration</td>
<td>Grammatical Attended Intake Group</td>
<td>Recognition Posttest</td>
<td>Total Duration</td>
<td>First Fixation Duration</td>
<td>Second Pass Time</td>
</tr>
<tr>
<td>----------------</td>
<td>--------------------------</td>
<td>------------------</td>
<td>--------------</td>
<td>----------------------------------</td>
<td>----------------------</td>
<td>---------------</td>
<td>--------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td></td>
<td>-16.99, 28.04</td>
<td>200.54, 195.01</td>
<td>28.46, 40.25</td>
<td></td>
<td>4.73, 3.23</td>
<td>-33.70, 59.65</td>
<td>-6.32, 51.96</td>
<td>-112.92, 30.43</td>
</tr>
<tr>
<td></td>
<td>(45.14)</td>
<td>(49.96)</td>
<td>(99.13)</td>
<td></td>
<td>(2.76)</td>
<td>(101.88)</td>
<td>(60.86)</td>
<td>(103.04)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>20</td>
<td>20</td>
<td></td>
<td>15</td>
<td>19</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>.51</td>
<td>.36</td>
<td>2.11</td>
<td></td>
<td>.29</td>
<td>.58</td>
<td>1.65</td>
<td>-1.21</td>
</tr>
<tr>
<td></td>
<td>.61</td>
<td>.05*</td>
<td>.72</td>
<td></td>
<td>.04*</td>
<td>.57</td>
<td>.12</td>
<td>.24</td>
</tr>
<tr>
<td></td>
<td>.01</td>
<td>.19</td>
<td>.01</td>
<td></td>
<td>.27</td>
<td>.02</td>
<td>.13</td>
<td>.08</td>
</tr>
</tbody>
</table>

*p < .01

**Lexical intake groups.**

Regarding the Lexical Noticed Intake Group, participants only had significantly different values for the eye-tracking measure of Second Pass Time ($t(20) = 2.18$, $p = .04$). In other words, the values for the lexical targets were significantly higher than the values for the lexical controls. For the Lexical Detected Intake Groups, participants showed significantly higher First Fixation Durations ($t(21) = 38.14$, $p < .01$) and Second Pass Time ($t(21) = 2.24$, $p = .04$) values for target lexical items than they did for control lexical items. In regards to the Lexical Attended Intake Group, there were significantly higher Total Duration ($t(14) = -2.75$, $p = .02$) and First Fixation Duration ($t(14) = -2.24$, $p = .04$) values for target lexical items than for control lexical items. There were no other statistically significant differences in eye-tracking values in the lexical intake groups (see Table 12 for more detail).
Table 12: Summary Results for Paired-Samples T-Tests for Instances of Target vs. Control Lexical Forms

<table>
<thead>
<tr>
<th></th>
<th>95% CI</th>
<th>Mean Value</th>
<th>Mean Value</th>
<th>N1/N2</th>
<th>t</th>
<th>p</th>
<th>effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Target (SD Target)</td>
<td>Control (SD Control)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical Noticed Intake Group</td>
<td>Total Duration</td>
<td>-64.35, 117.66</td>
<td>444.83 (104.87)</td>
<td>418.18 (199.41)</td>
<td>20</td>
<td>.61</td>
<td>.55</td>
</tr>
<tr>
<td></td>
<td>First Fixation Duration</td>
<td>-17.97, 29.46</td>
<td>224.25 (47.38)</td>
<td>218.51 (56.80)</td>
<td>20</td>
<td>.51</td>
<td>.62</td>
</tr>
<tr>
<td></td>
<td>Second Pass Time</td>
<td>3.76, 167.40</td>
<td>233.66 (235.04)</td>
<td>148.08 (165.51)</td>
<td>20</td>
<td>2.18</td>
<td>.04*</td>
</tr>
<tr>
<td></td>
<td>Gaze Duration</td>
<td>-42.77, 30.98</td>
<td>299.17 (65.70)</td>
<td>305.06 (100.38)</td>
<td>20</td>
<td>-.33</td>
<td>.74</td>
</tr>
<tr>
<td>Lexical Detected Intake Group</td>
<td>Total Duration</td>
<td>-7.71, 140.23</td>
<td>393.41 (269.21)</td>
<td>327.16 (217.08)</td>
<td>21</td>
<td>1.86</td>
<td>.08</td>
</tr>
<tr>
<td></td>
<td>First Fixation Duration</td>
<td>8.29, 38.14</td>
<td>214.04 (44.11)</td>
<td>190.83 (57.85)</td>
<td>21</td>
<td>3.23</td>
<td>.00*</td>
</tr>
<tr>
<td></td>
<td>Second Pass Time</td>
<td>4.03, 109.79</td>
<td>193.29 (151.84)</td>
<td>136.38 (132.64)</td>
<td>21</td>
<td>2.24</td>
<td>.04*</td>
</tr>
<tr>
<td></td>
<td>Gaze Duration</td>
<td>-26.22, 30.98</td>
<td>262.90 (71.32)</td>
<td>251.41 (132.40)</td>
<td>21</td>
<td>.62</td>
<td>.53</td>
</tr>
<tr>
<td>Lexical Attended Intake Group</td>
<td>Total Duration</td>
<td>217.87 (135.57)</td>
<td>281.54 (132.00)</td>
<td>14</td>
<td>-2.75</td>
<td>.02*</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>First Fixation Duration</td>
<td>181.10 (69.88)</td>
<td>214.09 (43.66)</td>
<td>14</td>
<td>-2.24</td>
<td>.04*</td>
<td>.26</td>
</tr>
<tr>
<td></td>
<td>Second Pass Time</td>
<td>83.17 (125.23)</td>
<td>80.11 (115.86)</td>
<td>14</td>
<td>.23</td>
<td>.10</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>Gaze Duration</td>
<td>212.59 (101.91)</td>
<td>346.41 (53.43)</td>
<td>14</td>
<td>-1.75</td>
<td>.82</td>
<td>.18</td>
</tr>
</tbody>
</table>

*p < .01

Summary.

As shown by paired-samples t-tests, all grammatical intake groups except for the Grammatical Detected Intake Group scored significantly higher on instances of the target...
grammatical form than they did on instances of the control grammatical form. The only eye-tracking measure that showed a statistically significant difference between instances of the target and control grammatical forms was the Second Pass Times in the Grammatical Detected Intake Group; this eye-tracking measure in this group showed that participants had longer Second Pass Times on instances of the target grammatical form than they did for the control grammatical form.

Regarding the lexical intake groups, the Lexical Noticed Intake Group showed significantly higher target lexical values only for Second Pass Time. For the Lexical Detected Intake Groups, participants showed higher First Fixation Duration and Second Pass Time values for target lexical items than they did for control lexical items. The Lexical Attended Intake Group returned significantly higher values for lexical target items than for lexical control items for both Total Duration and First Fixation Duration.

**Were posttest scores significantly higher than pretest scores?**

In order to determine whether posttest scores were significantly higher than pretest scores, scores from pretest/posttest sets were submitted to paired-samples t-tests. More specifically, the pairs examined were the Controlled Production Pretest/Posttest and the Items Pretest/Recognition Posttest. Results from the grammatical groups are presented first, followed by lexical groups.

**Grammatical intake groups.**

Results showed statistically significant differences between the Items Pretest and the Recognition Posttest for all grammatical intake groups (see Table 13 for details): the Grammatical Noticed Intake Group ($t(17) = -7.01, p = .00$), the Grammatical Detected Intake Group ($t(19) = -9.74, p = .00$), the Grammatical Attended Intake Group ($t(18) = -5.63, p = .00$),
the Grammatical Noticed+ TA Group ($t(18) = -12.01, p = .00$), and the Grammatical Noticed+ Silent Group ($t(18) = -9.66, p = .00$). In other words, all grammatical intake groups scored significantly higher on the Recognition Posttest than on the Items Pretest. The same pattern was not found for the Controlled Production Pretest and Posttest. There were no statistically significant differences in scores between these two tests for any of the grammatical intake groups; this means that no grammatical group showed statistically significant differences in scores between the Controlled Production Pretest and the Controlled Production Posttest.

Table 13: Summary Results for Paired-Samples T-Tests for Differences in Grammatical Scores between Pretests and Posttests

<table>
<thead>
<tr>
<th>Grammatical Intake Group</th>
<th>Controlled Production Tests</th>
<th>Items/Recognition Tests</th>
<th>N1/N2</th>
<th>t</th>
<th>p</th>
<th>effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical Noticed+ TA Group</td>
<td>-21.7 (2.43)</td>
<td>.00 (.00)</td>
<td>18</td>
<td>-1.80</td>
<td>.09</td>
<td>.15</td>
</tr>
<tr>
<td>Grammatical Noticed+ Silent Group</td>
<td>-.65 (.83)</td>
<td>.00 (.00)</td>
<td>19/19</td>
<td>-1.42</td>
<td>.17</td>
<td>.10</td>
</tr>
<tr>
<td>Grammatical Noticed Intake Group</td>
<td>-1.17 (3.20)</td>
<td>.00 (.00)</td>
<td>17</td>
<td>-1.00</td>
<td>.33</td>
<td>.06</td>
</tr>
<tr>
<td>Grammatical Detected Intake Group</td>
<td>-1.55 (2.24)</td>
<td>.00 (.00)</td>
<td>19</td>
<td>-1.00</td>
<td>.33</td>
<td>.05</td>
</tr>
<tr>
<td>Grammatical Attended Intake Group</td>
<td>-7.63 (2.83)</td>
<td>.00 (.00)</td>
<td>18</td>
<td>-9.66</td>
<td>.00</td>
<td>.84</td>
</tr>
</tbody>
</table>
Lexical intake groups.

The results from the lexical intake groups mirrored those of the grammatical intake groups in terms of significant and non-significant differences on pre- and posttests. There were statistically significant differences between the Items Pretest and the Recognition Posttest for all lexical intake groups. The Lexical Noticed Intake Group analysis returned $t(20) = -5.91, p = .00$, the Lexical Detected Intake Group returned $t(21) = -6.09, p = .00$, the Lexical Attended Intake Group showed $t(14) = -2.98, p = .01$, the Lexical Noticed+ TA Group showed $t(18) = -8.90, p = .00$, and the Lexical Noticed+ Silent Group, $t(18) = -5.82, p = .00$. In other words, all lexical intake groups scored significantly higher on the Recognition Posttest than on the Items Pretest. Additional details can be found in Table 14.

Again, as was the case with the grammatical intake groups, these results were not repeated for the Controlled Production Pretest and Posttest. There were no statistically significant differences in scores between these two tests for any of the lexical intake groups (see Table 14). In other words, none of the lexical intake groups showed statistically significant differences in scores between the Controlled Production Pretest and the Controlled Production Posttest.

Table 14: Summary Results for Paired-Samples T-Tests for Differences in Lexical Scores between Pretests and Posttests

<table>
<thead>
<tr>
<th>Lexical Noticed+ TA Intake Group</th>
<th>Controlled Production Tests</th>
<th>Items/Recognition Tests</th>
<th>95% CI</th>
<th>Mean Value Pretest (SD)</th>
<th>Mean Value Posttest (SD)</th>
<th>N1/N2</th>
<th>t</th>
<th>p</th>
<th>effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.00 (.00)</td>
<td>.03 (.07)</td>
<td>18</td>
<td>-1.97</td>
<td>.06</td>
<td>.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.06, .00</td>
<td>.00 (.00)</td>
<td>18</td>
<td>-1.97</td>
<td>.06</td>
<td>.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.67, -.41</td>
<td>.00 (.00)</td>
<td>18</td>
<td>-8.90</td>
<td>.00*</td>
<td>.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.54 (.26)</td>
<td>.54 (.26)</td>
<td>18</td>
<td>-8.90</td>
<td>.00*</td>
<td>.83</td>
</tr>
</tbody>
</table>

* $p < .01$
Summary.

In summary, paired-samples t-tests showed that participants in all grammatical and lexical intake groups improved significantly from pretest to posttest on the Items Pretest/Recognition Posttest. There were no significant differences between pretest and posttest for the Controlled Production tests for any grammatical or lexical intake group.

Results of verb stem meaning pretest.

The verb stem meaning pretest was administered in order to assure that all participants possessed accurate prior knowledge of the lexical meaning of the infinitives of the instances of the target grammatical form. Since participants had already had much exposure to the infinitives of the instances of the target grammatical form in class, through activities, and in the textbook, scores on the first administration of the verb stem meaning pretest were very high (see Table 15.
for descriptive statistics). All participants \((n = 14)\) who did not score 100% on the first administration retook the test; all obtained perfect scores on the second administration.

Table 15: Descriptive Statistics for the First Administration of the Verb Stem Meaning Pretest\(^a\)

<table>
<thead>
<tr>
<th></th>
<th>1(^{st}) administration</th>
<th>2(^{nd}) administration</th>
<th>(n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical Noticed+ TA Intake Group</td>
<td>19.84 (.69)</td>
<td>20.00 (0)</td>
<td>19</td>
</tr>
<tr>
<td>Grammatical Noticed+ Silent Intake Group</td>
<td>19.72 (.75)</td>
<td>20.00 (0)</td>
<td>18</td>
</tr>
<tr>
<td>Grammatical Noticed Intake Group</td>
<td>19.68 (.82)</td>
<td>20.00 (0)</td>
<td>19</td>
</tr>
<tr>
<td>Grammatical Detected Intake Group</td>
<td>19.35 (1.50)</td>
<td>20.00 (0)</td>
<td>20</td>
</tr>
<tr>
<td>Grammatical Attended Intake Group</td>
<td>19.63 (1.01)</td>
<td>20.00 (0)</td>
<td>19</td>
</tr>
<tr>
<td>Total</td>
<td>19.64 (1.00)</td>
<td>20.00 (0)</td>
<td>95</td>
</tr>
</tbody>
</table>

\(^a\) Highest possible score = 20

**Summary**

This section revealed results of several analyses performed for validity reasons. The overwhelming majority of paired-samples t-tests performed to determine if participants exposed to one text order scored higher on posttests than those exposed to a second text order showed that text order did not play a role in posttest scores; only comprehension scores of two intake groups showed a significant difference between scores of one text order versus the second text order. Overall, these results indicate that the text order variable can be collapsed for the analyses of the research questions.

Paired-samples t-tests also showed that all grammatical intake groups except for the Grammatical Detected Intake Group scored significantly higher on instances of the target grammatical form on the Recognition Posttest than they did on instances of the control grammatical form. For the most part, the eye-tracking measures did not show this same
difference. Regarding the lexical intake groups, the Lexical Noticed Intake Group showed significantly higher target lexical values only for Second Pass Time, while the Lexical Detected Intake Groups showed higher values for both First Fixation Duration and Second Pass Time; the Lexical Attended Intake Group returned significantly lower values for Total Duration and First Fixation Duration.

In other words, it appears that there were no inherent aspects of the target grammatical form that caused participants to fixate longer on them than on the control grammatical form; had there been, perhaps other aspects of the target grammatical form were causing participants to fixate longer and this would have affected results.

Regarding improvement from pretest to posttest scores, paired-samples t-tests showed that participants in all grammatical and lexical intake groups improved significantly from the Items Pretest to the Recognition Posttest. However, there were no significant differences between pretest and posttest for the Controlled Production tests for any grammatical or lexical intake group.

Research Questions

Research question one.

In order to determine if there was reactivity present in the current study, it is necessary to investigate the differences between the Noticed+ TA Intake Groups and the Noticed+ Silent Intake Groups. The three assessment measures that are examined for reactivity are the Recognition Posttest, the Controlled Written Production Posttest, and the Comprehension Posttest. The grammatical results will be presented first, followed by the lexical results.
Does thinking aloud while performing a reading task have any effect on beginning Spanish readers’ a) recognition, b), production, and c) comprehension of a grammatical form when compared to readers not thinking aloud?

In order to determine if there was reactivity in the grammatical aspect of the current study, the performance of the Grammatical Noticed+ TA Intake Group was compared with the performance of the Grammatical Noticed+ Silent Intake Group on the grammatical scores of three assessment measures: the recognition posttest, the controlled production posttest, and the comprehension posttest.

As will be presented in RQ2, although an ANOVA showed that Grammatical Intake Group did affect performance on the recognition posttest (grammatical score) \( F(4, 90) = 4.92, p < .01 \), a Scheffé’s method post hoc (see Table 20) revealed no statistically significant difference between the Grammatical Noticed+ TA Intake Group and the Grammatical Noticed+ Silent Intake Group on the recognition posttest (grammatical score, Mean difference = 7.08, 95% CI = 6.14, 8.02). The Scheffé post hoc comparison showed that the mean difference between the Grammatical Noticed+ TA Intake Group and the Grammatical Noticed+ Silent Intake Group was 1.63, \( p = .55 \), and 95% CI = -1.31, 4.57. While effect size was .18, power was .95. As will be presented in RQ3, a one-way ANOVA also showed no significant difference in grammatical production on the controlled production posttest (grammar score) between the Grammatical Noticed+ TA Intake Group and the Grammatical Noticed+ Silent Intake Group \( F(4, 90) = .89, p = .48 \), refer to Table 22 and Table 23 for details). Effect size was .04 and power was .27. Furthermore, a one-way ANOVA with Grammatical Intake
Group as the independent variable and Grammatical Comprehension as the dependent variable showed no significant differences between the Grammatical Noticed+ TA Intake Group and the Grammatical Noticed+ Silent Intake Group \((F(4, 90) = 1.63, p = .17)\), Effect size was .07 and power was .48. The pertinent information from Table 20, Table 22, and Table 23, in addition to the grammatical comprehension results, is compiled in Table 16 and Table 17.

Table 16: Descriptive Statistics for RQ1

<table>
<thead>
<tr>
<th></th>
<th>For Grammatical Intake Groups*</th>
<th>For Lexical Intake Groups*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Recognition, Noticed+ TA Intake Group</td>
<td>7.89 (2.87)</td>
<td>.54 (.26)</td>
</tr>
<tr>
<td>Recognition, Noticed+ Silent Intake Group</td>
<td>6.26 (2.83)</td>
<td>.35 (.26)</td>
</tr>
<tr>
<td>Production, Noticed+ TA Intake Group</td>
<td>1.00 (2.43)</td>
<td>.26 (.56)</td>
</tr>
<tr>
<td>Production, Noticed+ Silent Intake Group</td>
<td>.26 (.81)</td>
<td>.37 (1.17)</td>
</tr>
<tr>
<td>Comprehension, Noticed+ TA Intake Group</td>
<td>5.53 (6.66)</td>
<td>4.93 (6.69)</td>
</tr>
<tr>
<td>Comprehension, Noticed+ Silent Intake Group</td>
<td>9.50 (6.41)</td>
<td>6.61 (5.80)</td>
</tr>
</tbody>
</table>

\*n = 19 for all groups

Table 17: ANOVA results of RQ1

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>sum of square</th>
<th>mean square</th>
<th>f</th>
<th>p</th>
<th>effect size</th>
<th>power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition</td>
<td>1</td>
<td>163.61</td>
<td>40.90</td>
<td>4.92</td>
<td>.00</td>
<td>.18</td>
<td>.95</td>
</tr>
<tr>
<td>Production</td>
<td>4</td>
<td>9.38</td>
<td>2.35</td>
<td>.89</td>
<td>.48</td>
<td>.04</td>
<td>.27</td>
</tr>
<tr>
<td>Comprehension</td>
<td>4</td>
<td>151.47</td>
<td>37.87</td>
<td>1.63</td>
<td>.17</td>
<td>.07</td>
<td>.48</td>
</tr>
<tr>
<td>Lexical</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition</td>
<td>1</td>
<td>.36</td>
<td>.36</td>
<td>5.31</td>
<td>.03</td>
<td>.13</td>
<td>.61</td>
</tr>
<tr>
<td>Production</td>
<td>4</td>
<td>1.10</td>
<td>.28</td>
<td>.62</td>
<td>.65</td>
<td>.03</td>
<td>.20</td>
</tr>
<tr>
<td>Comprehension</td>
<td>4</td>
<td>63.00</td>
<td>15.75</td>
<td>.77</td>
<td>.55</td>
<td>.04</td>
<td>.24</td>
</tr>
</tbody>
</table>
Does thinking aloud while performing a reading task have any effect on beginning Spanish readers’ a) recognition, b), production, and c) comprehension of lexical items when compared to readers not thinking aloud?

The lexical performance of the Lexical Noticed+ TA Intake Group was compared with that of the Lexical Noticed+ Silent Intake Group to determine if there was any reactivity in the lexical aspect of this study. Scores from the recognition posttest, the controlled production posttest, and the comprehension posttest were analyzed. Regarding the recognition posttest, the Lexical Intake Group did affect performance on the Percent Recognition Posttest (lexical score): $F(4, 91) = 5.13, p < .01$, but the Scheffé’s method post hoc did not find any statistical differences between the Lexical Noticed+ TA Intake Group and the Lexical Noticed+ Silent Intake Group (see Table 17). Effect size was .13 and power was .61. In RQ3, results will be presented regarding the effects of intake condition on lexical production. Lexical Intake Group did not affect lexical performance on the controlled production posttest ($F(4, 91) = .62, p = .65$), so there was no significant difference between the Lexical Noticed+ TA Intake Group and the Lexical Noticed+ Silent Intake Group on the production measure (see Table 22 and Table 23). Effect size was .65 and power was .20. Regarding comprehension of the target lexical items, another one-way ANOVA with Lexical Intake Group as the independent variable and Lexical Comprehension as the dependent variable showed no significant differences between the Lexical Noticed+ TA Intake Group and the Lexical Noticed+ Silent Intake Group ($F(4, 90) = .77, p = .55$). Effect size was .04 and power was .24.
In summary, no reactivity for grammatical or lexical intake groups was found in the current study.

**Research question two.**

The second research question asked whether level of intake has an effect on adult L2 learner’s recognition of a) a grammatical form and/or b) lexical items from a written text as measured by recognition tests. The statistical report on this research question is divided into two sections: a report on the grammatical aspect and a report on the lexical aspect.

**Does level of intake have an effect on adult L2 learner’s recognition of a grammatical form?**

In order to answer this research question, a one-way ANOVA was run with Grammatical Intake Group as the independent variable and Recognition Posttest Grammatical Score as the dependent variable. The descriptive statistics for the groups were: Grammatical Noticed+ TA Group, $M = 7.89$, SD = 2.83, $n = 19$; Grammatical Noticed+ Silent Group, $M = 6.26$, SD = 3.11, $n = 19$; Grammatical Noticed Intake Group, $M = 5.28$, SD = 3.20, $n = 18$; Grammatical Detected Intake Group, $M = 5.45$, SD = 2.50, $n = 20$; and Grammatical Attended Intake Group, $M = 3.89$, SD = 3.02, $n = 19$ (see Table 18). Analysis results showed that Grammatical Intake Group did affect performance on the Recognition Posttest (grammatical score): $F(4, 90) = 4.92$, $p < .01$. Comparisons using Scheffé’s method post hoc (see Table 20) found a statistical difference between the Grammatical Attended Intake Group and the Grammatical Noticed+ TA Group (Mean difference = -4.00, CI = -6.94, -1.06, $p < .01$). However, there were no other statistical differences between any other group (see Table 20). Although the effect size was small ($\eta^2_p = .18$), the power was large (observed power = .95). This means that level of intake may indeed have an effect on adult L2 learner’s recognition of a grammatical target form. More specifically,
these results suggest that learners can recognize more instances of a grammatical form at the highest level of intake (Grammatical Noticed+ TA Intake Group) than at the lowest one (Grammatical Attended Intake Group).

Table 18: Descriptive Statistics for RQ2

<table>
<thead>
<tr>
<th></th>
<th>For Grammatical Intake Groups</th>
<th>For Lexical Intake Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>n</td>
</tr>
<tr>
<td>Noticed+ TA Intake Group</td>
<td>7.89 (2.87)</td>
<td>19</td>
</tr>
<tr>
<td>Noticed+ Silent Intake Group</td>
<td>6.26 (2.83)</td>
<td>19</td>
</tr>
<tr>
<td>Noticed Intake Group</td>
<td>5.28 (3.20)</td>
<td>18</td>
</tr>
<tr>
<td>Detected Intake Group</td>
<td>5.45 (2.50)</td>
<td>20</td>
</tr>
<tr>
<td>Attended Intake Group</td>
<td>3.89 (3.02)</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 19: ANOVA results of RQ2

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>sum of square</th>
<th>mean square</th>
<th>f</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical Intake Groups</td>
<td>4</td>
<td>163.61</td>
<td>40.90</td>
<td>4.92</td>
<td>.00*</td>
</tr>
<tr>
<td>Lexical Intake Groups</td>
<td>4</td>
<td>96.14</td>
<td>24.03</td>
<td>5.13</td>
<td>.00*</td>
</tr>
</tbody>
</table>

*p < .01

Table 20: Homogenous subsets from Scheffé post hoc on Grammatical Level of Intake and Recognition Posttest

<table>
<thead>
<tr>
<th>Grammatical Level of Intake Group for Scheffé post hoc a,b,c</th>
<th>n</th>
<th>Subset</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Noticed+ Silent Intake Group</td>
<td>19</td>
<td>6.26</td>
</tr>
<tr>
<td>Noticed+ TA Intake Group</td>
<td>19</td>
<td>7.89</td>
</tr>
<tr>
<td>Attended Intake Group</td>
<td>19</td>
<td>3.89</td>
</tr>
<tr>
<td>Noticed Intake Group</td>
<td>18</td>
<td>5.28</td>
</tr>
<tr>
<td>Detected Intake Group</td>
<td>20</td>
<td>5.45</td>
</tr>
<tr>
<td>Sig.</td>
<td>.18</td>
<td>.10</td>
</tr>
</tbody>
</table>

a Uses Harmonic Mean Sample Size = 18.98

b The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.

c Alpha = .05.

29 Grammatical recognition data from one participant was missing.
Does level of intake have an effect on adult L2 learner’s recognition of a lexical form?

A second one-way ANOVA was run to answer the second part of this research question. Lexical Intake Group was the independent variable and Percent Recognition Posttest Lexical Score was the dependent variable. The descriptive statistics for the groups were: Grammatical Noticed+ TA Group, $M = 4.79$, $SD = 2.25$, $n = 19$; Grammatical Noticed+ Silent Group, $M = 3.32$, $SD = 2.43$, $n = 19$; Grammatical Noticed Intake Group, $M = 2.43$, $SD = 1.89$, $n = 21$; Grammatical Detected Intake Group, $M = 2.27$, $SD = 1.78$, $n = 22$; and Grammatical Attended Intake Group, $M = 2.00$, $SD = 2.56$, $n = 15$ (see Table 18). Results demonstrated that Lexical Intake Group did affect performance on the Percent Recognition Posttest (lexical score): $F(4, 91) = 5.13$, $p < .01$. Scheffé’s method post hoc found statistical differences between the Lexical Noticed Intake Group and the Lexical Noticed+ TA Group (Mean difference = -2.36, CI = -4.52, - .20, $p = .02$), between the Lexical Detected Intake Group and the Lexical Noticed+ TA Group (Mean difference = -2.52, CI = -4.65, - .38, $p = .01$), and finally between the Lexical Attended Intake Group and the Lexical Noticed+ TA Group (Mean difference = - - .79, CI = -5.14, - .44, $p = .01$, see Table 19). Although the effect size was small ($\eta_p^2 = .18$), the power was large (observed power = .96). This means that level of intake may indeed have an effect on adult L2 learner’s recognition of a lexical target form. Taken together, these results suggest that learners can recognize more instances of a lexical form at a higher level of intake than at a lower one. This effect is visible when comparing the Lexical Noticed+ TA Intake Group with three other groups: the Lexical Noticed Intake Group, the Lexical Detected Intake Group, and the Lexical Attended Intake Group.
Research question three.

The third research question of this study examined whether level of intake has an effect on adult L2 learners’ controlled written production of a) a grammatical form and/or b) lexical items from a written text as measured by controlled written production tests. The grammatical results will be presented first, followed by the lexical results.

Does level of intake have an effect on controlled written production of a grammatical form?

The analysis run to answer this research question was a one-way ANOVA with Grammatical Intake Group as the independent variable and Controlled Production Posttest (grammatical score) as the dependent variable. The descriptive statistics, as documented in Table 22, were: Grammatical Noticed+ TA Group, $M=1.00$, $SD=2.43$, $n=19$; Grammatical Noticed+ Silent Group, $M=.26$, $SD=.81$, $n=19$; Grammatical Noticed Intake Group, $M=.06$, $SD=.24$, $n=18$; Grammatical Detected Intake Group, $M=.50$, $SD=2.24$, $n=20$; and Grammatical Attended Intake Group, $M=.37$, $SD=1.17$, $n=19$. Analysis results showed that Grammatical Intake Group did not affect performance on the Controlled Production Posttest (grammatical score): $F(4, 90) = .89$, $p = .48$ (see Table 23). Both the effect size ($\eta^2 = .04$) and
the power were small (observed power = .27). In other words, level of intake did not appear to significantly affect controlled written production of a grammatical form.

Table 22: Descriptive Statistics for RQ3

<table>
<thead>
<tr>
<th>Intake Group</th>
<th>M (SD)</th>
<th>n</th>
<th>M (SD)</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noticed+ TA Intake Group</td>
<td>1.00 (.24)</td>
<td>19</td>
<td>.26 (.56)</td>
<td>19</td>
</tr>
<tr>
<td>Noticed+ Silent Intake Group</td>
<td>.26 (.81)</td>
<td>19</td>
<td>.37 (1.17)</td>
<td>19</td>
</tr>
<tr>
<td>Noticed Intake Group</td>
<td>.06 (.24)</td>
<td>18</td>
<td>.05 (.22)</td>
<td>21</td>
</tr>
<tr>
<td>Detected Intake Group</td>
<td>.50 (2.24)</td>
<td>20</td>
<td>.18 (.50)</td>
<td>22</td>
</tr>
<tr>
<td>Attended Intake Group</td>
<td>.37 (1.17)</td>
<td>19</td>
<td>.20 (.56)</td>
<td>15</td>
</tr>
</tbody>
</table>

Table 23: ANOVA results of RQ3

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>sum of square</th>
<th>mean square</th>
<th>f</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical Intake Groups</td>
<td>4</td>
<td>9.38</td>
<td>2.35</td>
<td>.89</td>
<td>.48</td>
</tr>
<tr>
<td>Lexical Intake Groups</td>
<td>4</td>
<td>1.10</td>
<td>.28</td>
<td>.62</td>
<td>.65</td>
</tr>
</tbody>
</table>

*p < .01

**Does level of intake have an effect on controlled written production of a lexical form?**

A one-way ANOVA was performed to answer this part of the research question; Lexical Intake Group was the independent variable and Percent Controlled Production Posttest was the dependent variable. The descriptive statistics are documented in Table 22 and were: Lexical Noticed+ TA Group, M = .26, SD = .56, n = 19; Lexical Noticed+ Silent Group, M = .37, SD = 1.17, n = 19; Lexical Noticed Intake Group, M = .05, SD = .22, n = 15; Lexical Detected Intake Group, M = .18, SD = .50, n = 22; and Lexical Attended Intake Group, M = .20, SD = .56, n = 15. Results mirrored those of the grammatical form: Lexical Intake Group did not affect performance on the Controlled Production Posttest (lexical score): \( F(4, 91) = .62, p = .65 \) (see
Both the effect size ($\eta^2 = .03$) and the power were small (observed power = .20). In other words, level of intake did not appear to significantly affect controlled written production of lexical items.

**Research question four.**

The fourth research question aims to determine whether there are any relationships between L2 learners’ eye fixation time measures (first fixation duration, gaze duration, second pass time, total fixation time) and their abilities to recognize a) a grammatical form and/or b) lexical items in written L2 input as measured on recognition posttests. The analysis information and results concerning the grammatical form will be presented first, followed by those of the lexical items.

**Are there any significant relationships between fixation measures and recognition of a grammatical form?**

First, in order to obtain a wider range of values, the four eye-tracking measures (Subtracted Grammatical Total Duration, Subtracted Grammatical First Fixation Duration, Subtracted Grammatical Gaze Duration, Subtracted Grammatical Second Pass Time) were used as summed variables for this part of the research question. As an example, each participant was exposed to ten instances of the target grammatical form and ten instances of the control grammatical form and thus had ten subtracted values for each eye fixation measure. The ten subtracted values were summed for each participant so that each participant had four values: Summed Grammatical Total Duration, Summed Grammatical First Fixation Duration, Summed Grammatical Gaze Duration, and Summed Grammatical Second Pass Time. These summed values were used in order to have a wider range of values.
Before performing the main analysis of the first part of RQ4, the researcher ran tests of skewness on these summed variables to determine whether the data was normally distributed. It was important to check for this due to the fact that participants had only a perceived, not actual, time limit for completing the reading passages, so the range of eye-tracking measures may have greatly varied. The skewness tests showed that the only variable that was normally distributed was Summed Grammatical Second Pass Time; the other three variables (Summed Grammatical First Fixation Duration, Summed Grammatical Gaze Duration) were transformed using a log transformation to obtain normal distribution\(^{30}\).

In order to answer this part of RQ4, bivariate correlations for each grammatical intake group were performed with the following variables: Recognition Posttest Score\(^{31}\), Summed Grammatical Total Duration, Summed Grammatical First Fixation Duration, Summed Grammatical Gaze Duration, and Summed Grammatical Second Pass Time. Descriptive statistics are listed in Table 24. The analysis returned no significant correlations between Recognition Posttest Scores and a) Summed Grammatical Total Time, b) Summed Grammatical First Fixation Duration, c) Summed Grammatical Gaze Duration, and d) Subtracted Grammatical Second Pass Time for any grammatical intake group (see Table 25).

Table 24: Descriptive Statistics for Fixation Measures and Recognition of a Grammatical Form

<table>
<thead>
<tr>
<th>Grammatical Noticed Intake Group</th>
<th>Grammatical Detected Intake Group</th>
<th>Grammatical Attended Intake Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD) n</td>
<td>M (SD) n</td>
</tr>
<tr>
<td>Grammatical Recognition Posttest</td>
<td>5.28 (3.20) 18</td>
<td>5.45 (2.50) 20</td>
</tr>
<tr>
<td>Subtracted Grammatical Total Duration</td>
<td>6.20 (1.02) 12</td>
<td>6.31 (.83) 14</td>
</tr>
</tbody>
</table>

\(^{30}\) Eye fixation measure values do not use decimals because the unit of measure (milliseconds) is already small.

\(^{31}\) The maximum possible score per participant for the Grammatical Recognition Posttest was 10.
Subtracted Grammatical First Fixation Duration 5.33 (.81) 3 5.81 (.76) 14 5.88 (1.40) 13
Subtracted Grammatical Gaze Duration 6.23 (1.03) 7 6.18 (.79) 13 6.07 (1.35) 12
Subtracted Grammatical Second Pass Time -7.69 (737.19) 19 42.64 (407.52) 20 -143.74 (512.45) 19

Table 25: Correlations between Fixation Measures and Recognition of a Grammatical Form

<table>
<thead>
<tr>
<th>Grammatical Recognition Posttest</th>
<th>Grammatical Noticed Intake Group</th>
<th>Grammatical Detected Intake Group</th>
<th>Grammatical Attended Intake Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subtracted Grammatical Total Duration</td>
<td>.37</td>
<td>.21</td>
<td>.45</td>
</tr>
<tr>
<td>Subtracted Grammatical First Fixation Duration</td>
<td>-.87</td>
<td>-.12</td>
<td>-.01</td>
</tr>
<tr>
<td>Subtracted Grammatical Gaze Duration</td>
<td>-.68</td>
<td>.10</td>
<td>-.34</td>
</tr>
<tr>
<td>Subtracted Grammatical Second Pass Time</td>
<td>.36</td>
<td>.33</td>
<td>-.06</td>
</tr>
</tbody>
</table>

Note. *p < .05

For the Grammatical Noticed Intake Group (n = 19), correlation results were as follows: Grammatical Recognition Posttest and Subtracted Grammatical Total Duration ($r = .37, p = .26$; shared variance = 14%), Grammatical Recognition Posttest and Subtracted Grammatical First
Fixation Duration ($r = -0.87, p = 0.32$; shared variance = 76%), Grammatical Recognition Posttest and Subtracted Grammatical Gaze Duration ($r = -0.68, p = 0.09$; shared variance = 46%), and Grammatical Recognition Posttest and Subtracted Grammatical Second Pass Time ($r = 0.36, p = 0.14$; shared variance = 1%).

Results for the Grammatical Detected Intake Group ($n = 20$) were also non-significant: Grammatical Recognition Posttest and Subtracted Grammatical Total Duration ($r = 0.21, p = 0.46$; shared variance = 4%), Grammatical Recognition Posttest and Subtracted Grammatical First Fixation Duration ($r = -0.12, p = 0.69$; shared variance = 1%), Grammatical Recognition Posttest and Subtracted Grammatical Gaze Duration ($r = 0.09, p = 0.77$; shared variance = 1%), and Grammatical Recognition Posttest and Subtracted Grammatical Second Pass Time ($r = 0.33, p = 0.15$; shared variance = 11%).

There were also no significant correlations between any pair of variables in the Grammatical Attended Intake Group ($n = 19$): Grammatical Recognition Posttest and Subtracted Grammatical Total Duration ($r = 0.45, p = 0.16$; shared variance = 20%), Grammatical Recognition Posttest and Subtracted Grammatical First Fixation Duration ($r = -0.01, p = 0.99$; shared variance = %), Grammatical Recognition Posttest and Subtracted Grammatical Gaze Duration ($r = -0.34, p = 0.28$; shared variance < 1 %), and Grammatical Recognition Posttest and Subtracted Grammatical Second Pass Time ($r = -0.06, p = 0.80$; shared variance < 1%).

*Are these any significant relationships between fixation measures and recognition of a lexical form?*

This second part of RQ4 was set up as a binary approach due to the uniqueness of each target lexical item in comparison with the instances of the target grammatical form, which functioned more as a group. Therefore, the fixation measures of each of the ten target lexical
items that each participant was exposed to was paired with the correctness of their response to that item on the recognition posttest. Since each participant contributed more than one score and the dependent variable (recognition) was binary, a Generalized Linear Mixed model analysis of the logistic regression type was used. The dependent variable was Lexical Recognition Binary and the independent variables were Subtracted Lexical Total Duration Binary, Subtracted Lexical First Fixation Duration Binary, Subtracted Lexical Gaze Duration Binary, and Subtracted Lexical Second Pass Time Binary. Given that the linear mixed model is more appropriate when the dependent variable is continuous and the Generalized Linear Mixed model is the better choice when the dependent variable is binary, the Generalized Linear Mixed model was used for the analysis of this research question; Generalized Linear Mixed models were fit to the subtracted fixation binary durations. Furthermore, Stata 13, a data analysis and statistical software, was used in lieu of SPSS due to its increased power to perform more complicated statistical procedures such as mixed models.

32 As mentioned earlier, any target item for which a participant demonstrated accurate prior knowledge was removed from the analysis for that participant.
33 Eye fixation measure values do not use decimals because the unit of measure (milliseconds) is already small.
34 The linear mixed model was also considered as a possible analysis for RQ5, RQ6, and the second half of RQ4 that ultimately used the Generalized Linear Mixed model (second half of RQ4, RQ5, RQ6). The reasoning behind this is that although the dependent variables in all of these research questions are binary, there is arguably an underlying scale. For example, in the case of the dependent variable of production, the coding is either 1 (produced) or 0 (not produced). But actual production is not so clear-cut: although strict grading criteria were establish and used, two participants who earned a score of 0 (not produced) could still have performed very differently: one may have just missed the cutoff for production because they produced a target but had three (instead of only two) letters incorrect, while another participant also earning a score of 0 produced nothing at all. There is obviously an arguable difference between these two participants who both earned a 0 (not produced); the same type of argument can be made for participants who earned a score of 1 (produced) for production. Therefore, even when the dependent variable is scored in a binary way, it is still possible to consider the use of the linear mixed model instead of the Generalized Linear Mixed model due to the arguable existence of an underlying scale. However, in the current study, a more conservative approach was taken and the Generalized Linear Mixed model was used; binary dependent variables were not considered to have underlying scales.
Table 26: Descriptive Statistics for Fixation Measures and Recognition of a Lexical Form

<table>
<thead>
<tr>
<th></th>
<th>Lexical Recognition Binary</th>
<th>Subtracted Lexical Total Duration Binary</th>
<th>Subtracted Lexical First Fixation Duration Binary</th>
<th>Subtracted Lexical Gaze Duration Binary</th>
<th>Subtracted Lexical Second Pass Time Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD) n</td>
<td>M (SD) n</td>
<td>M (SD) n</td>
<td>M (SD) n</td>
<td>M (SD) n</td>
</tr>
<tr>
<td>Lexical Noticed Intake Group</td>
<td>.26 (.44) 220</td>
<td>6.05 (.61) 192</td>
<td>5.40 (.42) 192</td>
<td>5.64 (.50) 186</td>
<td>5.61 (.58) 76</td>
</tr>
<tr>
<td>Lexical Detected Intake Group</td>
<td>.24 (.43) 220</td>
<td>5.79 (.67) 193</td>
<td>5.37 (.38) 193</td>
<td>5.52 (.47) 196</td>
<td>5.61 (.72) 50</td>
</tr>
<tr>
<td>Lexical Attended Intake Group</td>
<td>.21 (.41) 140</td>
<td>5.68 (.58) 106</td>
<td>5.43 (.54) 106</td>
<td>5.63 (.51) 115</td>
<td>5.68 (.91) 29</td>
</tr>
</tbody>
</table>

Table 27: Results of Generalized Linear Mixed model for Lexical Noticed Intake Group

<table>
<thead>
<tr>
<th></th>
<th>Lexical Recognition Binary</th>
<th>Coefficient (b)</th>
<th>Standard error</th>
<th>p</th>
<th>Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical Noticed Intake Group</td>
<td>Subtracted Lexical Total Duration Binary</td>
<td>-.19</td>
<td>.31</td>
<td>.54</td>
<td>-.79, .41</td>
</tr>
<tr>
<td></td>
<td>Subtracted Lexical First Fixation Duration Binary</td>
<td>.03</td>
<td>.42</td>
<td>.95</td>
<td>-.79, .85</td>
</tr>
<tr>
<td></td>
<td>Subtracted Lexical Gaze Duration Binary</td>
<td>-.13</td>
<td>.37</td>
<td>.73</td>
<td>-.85, .60</td>
</tr>
<tr>
<td></td>
<td>Subtracted Lexical Second Pass Time Binary</td>
<td>.24</td>
<td>.57</td>
<td>.68</td>
<td>-.88, 1.35</td>
</tr>
<tr>
<td>Lexical Detected Intake Group</td>
<td>Subtracted Lexical Total Duration Binary</td>
<td>-.16</td>
<td>.28</td>
<td>.56</td>
<td>-.71, .38</td>
</tr>
<tr>
<td></td>
<td>Subtracted Lexical First Fixation Duration Binary</td>
<td>-.01</td>
<td>.46</td>
<td>.98</td>
<td>-.92, .90</td>
</tr>
</tbody>
</table>
For the Lexical Noticed Intake Group (n = 21), no analyses returned significant relationships: Lexical Recognition Binary and Subtracted Lexical Total Duration Binary ($b = -.19$, $p = .54$, 95% confidence interval = -.79 to .41), Lexical Recognition Binary and Subtracted Lexical First Fixation Duration Binary ($b = .03$, $p = .95$, 95% confidence interval = -.79 to .85), Lexical Recognition Binary and Subtracted Lexical Gaze Duration Binary ($b = -.13$, $p = .73$, 95% confidence interval = -.85 to .60), and Lexical Recognition Binary and Subtracted Lexical Second Pass Time Binary ($b = -.24$, $p = .68$, 95% confidence interval = -.88 to 1.35). See Table 26 and Table 27 for details.

There were also no analyses that returned statistically significant relations for any of the eye-tracking fixation measures and Lexical Recognition Binary for the Lexical Detected Intake Group (n = 22, see Table 27): Lexical Recognition Binary and Subtracted

<table>
<thead>
<tr>
<th></th>
<th>Subtracted Lexical Gaze Duration Binary</th>
<th>Subtracted Lexical Second Pass Time Binary</th>
<th>Subtracted Lexical Total Duration Binary</th>
<th>Subtracted Lexical First Fixation Duration Binary</th>
<th>Subtracted Lexical Gaze Duration Binary</th>
<th>Subtracted Lexical Second Pass Time Binary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical Attended Intake Group</td>
<td>-.34</td>
<td>.39</td>
<td>.39</td>
<td>-1.11, .43</td>
<td>-.33</td>
<td>.57</td>
</tr>
<tr>
<td></td>
<td>-.76</td>
<td>.67</td>
<td>.26</td>
<td>-2.07, .55</td>
<td>-.54</td>
<td>.69</td>
</tr>
<tr>
<td></td>
<td>.19</td>
<td>.96</td>
<td>.84</td>
<td>-1.69, 2.07</td>
<td>-.23</td>
<td>.64</td>
</tr>
</tbody>
</table>
Lexical Total Duration Binary \( (b = -0.16, p = 0.56, 95\% \text{ confidence interval} = -0.71 \text{ to } 0.38) \),
Lexical Recognition Binary and Subtracted Lexical First Fixation Duration Binary \( (b = -0.01, p = 0.98, 95\% \text{ confidence interval} = -0.92 \text{ to } 0.90) \), Lexical Recognition Binary and Subtracted Lexical Gaze Duration Binary \( (b = -0.34, p = 0.39, 95\% \text{ confidence interval} = -1.11 \text{ to } 0.43) \), and Lexical Recognition Binary and Subtracted Lexical Second Pass Time Binary \( (b = -0.76, p = 0.67, 95\% \text{ confidence interval} = -1.44 \text{ to } 0.79) \). See Table 26 for descriptive statistics.

For the Lexical Attended Intake Group \( (n = 15) \), the Generalized Linear Mixed models for Lexical Recognition Binary and the eye-fixation measures all returned non-significant (see Table 27) relation: Lexical Recognition Binary and Subtracted Lexical Total Duration Binary \( (b = -0.33, p = 0.56, 95\% \text{ confidence interval} = -1.44 \text{ to } 0.79) \);
Lexical Recognition Binary and Subtracted Lexical First Fixation Duration Binary \( (b = -0.54, p = 0.49, 95\% \text{ confidence interval} = -1.91 \text{ to } 0.83) \). Lexical Recognition Binary and Subtracted Lexical Gaze Duration Binary \( (b = -0.23, p = 0.72, 95\% \text{ confidence interval} = -1.49 \text{ to } 1.03) \), and Lexical Recognition Binary and Subtracted Lexical Second Pass Time Binary \( (b = 0.19, p = 0.84, 95\% \text{ confidence interval} = -1.69 \text{ to } 2.07) \). See Table 26 for descriptive statistics. In other words, there were no significant relationships between Lexical Recognition Binary and fixation measure.

In summary, the only statistically significant relationship between fixation measures and recognition of a grammatical form was for the Grammatical Noticed Intake Group: there was a positive relationship between Subtracted Grammatical Second Pass Time and Grammatical Recognition Posttest. Regarding lexical items, there were no significant relationships between lexical recognition and any fixation measure. However overall, it appears that there are very few
significant relationships between fixation measures and recognition of a grammatical and/or lexical form.

**Research question five.**

The fifth research question aims to determine whether there are any relationships between L2 learners’ depth of processing during a reading task, and subsequent recognition and written production of a) a grammatical form and/or b) lexical items embedded in the reading task. First presented will be the results concerning the grammatical form, followed by the results of the lexical items.

*Is there a relationship between depth of processing, as measured by form-meaning connections in think-aloud protocols during a reading task, and subsequent recognition and written production of a grammatical form embedded in the reading task?*

In order to answer this part of RQ5, a binary analysis was performed: each instance of the target grammatical form for every participant in the Grammatical Noticed+ TA Group was coded as either ‘recognized’ or ‘not recognized’ (variable name: “Grammatical Recognition Item-by-Item”) and ‘produced’ or ‘not produced’ (variable name: “Grammatical Production Binary”). For each of these instances, the codings of the TA protocols were consulted to determine whether the participant achieved a depth of processing of low, intermediate, or high on that particular instance (variable name: Grammatical Depth of Processing Binary). Due to the non-independent nature of the data and the binary coding of dependent variables, these data were then submitted to a Generalized Linear Mixed model analysis in Stata (see RQ4, part 2).

Descriptive statistics were as follows: Grammatical Depth of Processing Binary (M = 1.37, SD = .97, n = 190), Grammatical Recognition Binary (M = .71, SD = .46, n = 190),
Grammatical Production Binary (M = .05, SD = .22, n = 190). The analysis (see Table 28) revealed a non-significant correlation between Grammatical Depth of Processing Binary and Grammatical Recognition Binary (b = -.03, p = .92, 95% confidence interval = -.72 to .65) and a significant correlation between Grammatical Production Binary and Grammatical Depth of Processing Binary (b = 1.86, p < .01, 95% confidence interval = .44 to 3.27). In other words, as Grammatical Depth of Processing Binary increased, it was more likely that participants in the Grammatical Noticed+ TA Group correctly produced the corresponding items on the lexical controlled production posttest. However, the model applied cannot predict the grammatical recognition in a statistically significant manner.

Table 28: Correlations between Depth of Processing and a) Recognition and b) Production of a Grammatical Form and Lexical Items

<table>
<thead>
<tr>
<th></th>
<th>Coefficient (b)</th>
<th>Standard error</th>
<th>p</th>
<th>Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical Depth of Processing IBI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammatical Recognition Binary</td>
<td>-.03</td>
<td>.35</td>
<td>.92</td>
<td>-.72, .65</td>
</tr>
<tr>
<td>Grammatical Production Binary</td>
<td>1.86</td>
<td>.72</td>
<td>&lt;.01**</td>
<td>.44, 3.27</td>
</tr>
<tr>
<td>Lexical Depth of Processing IBI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical Recognition Binary</td>
<td>.60</td>
<td>.18</td>
<td>&lt;.01**</td>
<td>.24, .96</td>
</tr>
<tr>
<td>Lexical Production Binary</td>
<td>.48</td>
<td>.56</td>
<td>.39</td>
<td>-.61, 1.57</td>
</tr>
</tbody>
</table>

Note. ** p < .01
Is there a relationship between depth of processing, as measured by form-meaning connections in think-aloud protocols during a reading task, and subsequent recognition and written production of lexical items embedded in the reading task?

In order to answer this second part of RQ5, the same type of binary analysis was performed as in the first part of this research question: every target lexical item for every participant in the Lexical Noticed+ TA Group was coded as either ‘recognized’ or ‘not recognized’ (variable name: “Lexical Recognition Binary”) and ‘produced’ or ‘not produced’ (variable name: “Lexical Production Binary”). For every instance, codings from the TA protocols were used to determine if the participant had achieved a depth of processing of low, intermediate, or high on that particular instance (variable name: Lexical Depth of Processing Binary). Therefore, Generalized Linear Mixed models were fit to the depth of processing measures; Stata was used for the analyses of this research question.

Regarding the descriptive statistics, Lexical Depth of Processing Binary revealed $M = 1.94$, $SD = 1.18$, $n = 126$, Lexical Recognition Binary was $M = .43$, $SD = .50$, $n = 179$, and Lexical Production Binary showed $M = .03$, $SD = .17$, $n = 179$). Results (see Table 28) showed a significant relationship between Lexical Recognition Binary and Lexical Depth of Processing Binary ($b = .60$, $p < .01$; 95% confidence interval = .24 to .96). In other words, as Lexical Depth of Processing Binary increased, participants had a significantly higher probability of recognizing the corresponding lexical item. However, there was no significant relationship between Lexical Production Binary and Lexical Depth of Processing Binary ($b = .48$, $p = .39$, 95% confidence interval = -.61 to 1.57).
In review, as Grammatical Depth of Processing Binary increased, so did the probability that participants were able to produce the form on the controlled written posttest, however the relationship between Grammatical Depth of Processing Binary and Grammatical Recognition Binary was non-significant. The opposite was true for lexical items: as Lexical Depth of Processing Binary increased, participants were significantly more likely to correctly recognize that lexical item, but there was no significant relationship between Lexical Depth of Processing Binary and Lexical Production Binary.

**Research question six.**

The sixth research question investigates any relationship between depth of processing during a reading task and comprehension of a reading passage, as measured by a multiple-choice comprehension test. The grammatical results are presented first, followed by the lexical results.

The analysis performed to obtain the results on this research question was a Generalized Linear Mixed model in Stata and the data were organized in a binary manner, as in RQ5. In other words, each question of the comprehension posttest that focused on an instance of the target grammatical form was assigned either “correct” or “incorrect” (variable name: Grammatical Comprehension Binary) for each participant in the Grammatical Noticed+ TA Group, depending on the correctness of his/her answer. Then, for each of the same instances of the target grammatical form for every participant, the TA codings were consulted and each instance was entered as either a low, intermediate, or high depth of processing (variable name: Grammatical Depth of Processing Binary). The same data preparation was performed for the lexical items
Is there a relationship between depth of processing, as measured by form-meaning connections in think-aloud protocols during a reading task, and comprehension of a reading passage, as measured by a multiple-choice comprehension test?

Descriptive statistics were as follows: Grammatical Depth of Processing Binary (M = 1.37, SD = .97, n = 130), Lexical Depth of Processing Binary (M = 1.99, SD = 1.17, n = 126) Grammatical Comprehension Binary (M = .59, SD = .49, n = 140), Lexical Comprehension Binary (M = .53, SD = .50, n = 139). As seen in Table 29, results revealed a positive correlation between Grammatical Depth of Processing Binary and Grammatical Comprehension Binary (b = .68, p = .02, 95% confidence interval = .13 to 1.22), however there was not a significant correlation between Lexical Depth of Processing Binary and Lexical Comprehension Binary (b = .28, p = .13, 95% confidence interval = -.09 to .65). In other words, as Grammatical Depth of Processing Binary increased, it was more likely that participants correctly answered the corresponding questions on the comprehension posttest.

Table 29: Relationships between Depth of Processing and Comprehension of a Grammatical Form and Lexical Items

<table>
<thead>
<tr>
<th></th>
<th>Coefficient (b)</th>
<th>Standard error</th>
<th>p</th>
<th>Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td>.68</td>
<td>.28</td>
<td>.02*</td>
<td>.13, 1.22</td>
</tr>
<tr>
<td>Binary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lexical</td>
<td>.28</td>
<td>.19</td>
<td>1.3</td>
<td>-.09, .65</td>
</tr>
<tr>
<td>Comprehension</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Binary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. * p < .05
In summary, as Grammatical Depth of Processing Binary increased, it became more likely that the participants correctly answered the corresponding comprehension questions. However, there was no significant relationship between Lexical Depth of Processing Binary and Lexical Comprehension Binary.

**Research question seven.**

The seventh research question aims to determine whether there are any differences between recognition of a grammatical form and lexical items in beginning Spanish learners. Before the main statistical analysis could be started, however, it was necessary to convert the grammatical recognition pretest measurement to a percent value in order to be able to compare it with the percent lexical recognition pretest measurement. This was done by creating a new variable, “Percent Grammatical Recognition Posttest,” which was simply the Grammatical Recognition Posttest score divided by ten (total instances of target grammatical form that any participant was exposed to on the recognition posttest).

This research question required further manipulation of the data due to the intake groups: since each participant qualified for membership into two independent groups based on overall time spent on the reading passages, it would be problematic to directly compare the grammatical recognition of each grammatical intake group to the lexical recognition of each lexical intake group because each group consists of different participants. However, it was noted that the majority of participants (74%) had the same level of group membership for the grammatical intake group as for the lexical intake group. Therefore, the remedy to this problematic comparison is to, for this research question only, remove from the analyses all participants who pertain to different levels of grammatical intake groups and lexical intake groups (26%, or 25 participants out of 96 total participants). Each of the remaining 69 participants, then, pertains to
the same level of grammatical intake group as lexical intake group, and thus enables a less problematic analysis. Paired-samples t-tests were used in all analyses for this research question; the paired variables were Percent Grammatical Recognition Posttest and Percent Lexical Recognition Posttest. Results are presented by intake level. Summary results for all intake levels appear in Table 30.

*Are there any differences between recognition of a) a grammatical form and b) lexical items in beginning Spanish learners from a written text as measured by recognition tests?*

For the noticed+ TA intake level \((n = 19)\), the mean score on the Percent Grammatical Recognition Posttest \((M = .79, SD = .29)\) was larger than the Percent Lexical Recognition Posttest \((M = .54, SD = .26)\). The Percent Grammatical Recognition Posttest scores were significantly higher than the Percent Lexical Recognition Posttest \((t(18) = 3.10, p = .01)\). Power was high (observed power = .83) and effect size was small \((\eta_p^2 = .35)\).

Regarding the noticed+ silent intake level \((n = 19)\), the mean score for Percent Grammatical Recognition Pretest \((M = .63, SD = .28)\) was higher than that of Percent Lexical Recognition Pretest \((M = .35, SD = .26)\). The paired-samples t-test returned significant \((t(18) = 4.62, p = .00)\), meaning that for the noticed+ silent intake level, the mean score for Percent Grammatical Recognition Posttest was significantly higher than that of Percent Lexical Recognition Posttest. Power was high (observed power = .99) and effect size was medium \((\eta_p^2 = .54)\).

For the grammatical and lexical noticed intake groups, there was a total group membership of 14 participants; their mean score on the Percent Grammatical Recognition Posttest was \(M = .56 (SD = .32)\), while their mean scores on the Percent Lexical Recognition
Posttest was $M = .22$ (SD = .17). A paired-samples t-test between these two variables showed a significance of $p = .00$; this, combined with the mean scores, reveal that the noticed intake level had significantly higher Percent Grammatical Recognition Posttest scores than Percent Lexical Recognition Posttest scores (see Table 30). Both power and effect size were large (observed power = .99 and $\eta^2_p = .67$).

Table 30: Summary Results for Paired-Samples T-Tests for Difference between Grammatical and Lexical Recognition

<table>
<thead>
<tr>
<th></th>
<th>95% CI</th>
<th>Percent Grammatical Mean (SD)</th>
<th>Percent Lexical Mean (SD)</th>
<th>N</th>
<th>t-value</th>
<th>p-value</th>
<th>effect size</th>
<th>power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grammatical and Lexical Noticed+ TA Intake Groups</td>
<td>.08, .41</td>
<td>.79 (.29)</td>
<td>.54 (.26)</td>
<td>19</td>
<td>3.10</td>
<td>.01</td>
<td>.35</td>
<td>.83</td>
</tr>
<tr>
<td>Grammatical and Lexical Noticed+ Silent Intake Groups</td>
<td>.15, .41</td>
<td>.62 (.28)</td>
<td>.35 (.26)</td>
<td>19</td>
<td>4.62</td>
<td>.00</td>
<td>.54</td>
<td>.99</td>
</tr>
<tr>
<td>Grammatical and Lexical Noticed Intake Groups</td>
<td>.20, .48</td>
<td>.56 (.32)</td>
<td>.22 (.17)</td>
<td>14</td>
<td>5.12</td>
<td>.00</td>
<td>.67</td>
<td>.99</td>
</tr>
<tr>
<td>Grammatical and Lexical Detected Intake Groups</td>
<td>.21, .56</td>
<td>.62 (.19)</td>
<td>.24 (.18)</td>
<td>9</td>
<td>4.98</td>
<td>.01</td>
<td>.76</td>
<td>.99</td>
</tr>
<tr>
<td>Grammatical and Lexical Attended Intake Groups</td>
<td>-.09, .31</td>
<td>.29 (.27)</td>
<td>.18 (.29)</td>
<td>9</td>
<td>1.28</td>
<td>.24</td>
<td>.17</td>
<td>.20</td>
</tr>
</tbody>
</table>

Regarding the detected intake groups, the Grammatical Detected Intake Group had a mean score of $M = .62$ (SD = .19) on the Percent Grammatical Recognition Posttest while the
Lexical Intake Group showed a mean score of $M = .24$ (SD = .18) on the Percent Lexical Recognition Posttest. ($n = 9$). There was a statistically significant difference between the Percent Grammatical Recognition Posttest score and the Percent Lexical Recognition Posttest score ($t(8) = 4.98, p < .01$). Both power and effect size were large (observed power = .99 and $\eta_p^2 = .76$). In other words, the Grammatical Detected Intake Group scores on the Percent Grammatical Recognition Posttest were significantly higher than the Lexical Detected Intake Group’s Percent Lexical Recognition Posttest scores.

The Grammatical Attended Intake Group and Lexical Attended Intake Group did not show similar results. The nine participants who belonged to these two groups had a mean Percent Grammatical Recognition Posttest scores of $M = .29$ (SD = .27) and a mean Percent Lexical Recognition Posttest score of $M = .18$ (SD = .29). There was no significant difference between their performance on the Percent Grammatical Recognition Posttest and the Percent Lexical Recognition Posttest ($t(8) = 1.28, p = .24$). Both power and effect size were small (observed power = .20, $\eta_p^2 = .17$).

In summary, the seventh research question aimed to determine the differences between recognition of a grammatical form and lexical items in beginning Spanish learners. Due to the nature of the experimental design (each participant belonged to both a grammatical intake group and a lexical intake group), a direct comparison of the recognition scores was not possible. Therefore, the analyses for the seventh research question were based on data only from those participants who belonged to the same level of intake group for both grammatical and lexical. In addition, the variables compared in the paired-samples t-tests were Percent Grammatical Recognition Posttest and Percent Lexical Recognition Posttest. Results showed that participants
scored significantly higher on the Percent Grammatical Recognition Posttest than on the Percent Lexical Recognition Posttest in all intake levels except for the attended intake level.

**Overall Summary.**

For RQ1, no reactivity for grammatical or lexical intake groups was found in the current study. The second research question found that the Grammatical Noticed+ TA Group performed significantly better on the recognition posttest than did the Grammatical Attended Intake Group; regarding lexical items, the Lexical Noticed+ TA Group performed significantly better than the a) the Lexical Noticed Intake Group, b) the Lexical Detected Intake Group, and c) the Lexical Attended Intake Group. Results from the third research question showed that level of intake did not appear to significantly affect written production of neither instances of the target grammatical form nor target lexical items. As stated in the results of RQ4, the only significant relationship between eye fixation measure and recognition was between Subtracted Grammatical Second Pass Time and Grammatical Recognition Posttest. RQ5 and RQ6 revealed that as depth of processing of the target grammatical form increased, participants were significantly more likely to be able to accurately produce this form and answer comprehension questions that corresponded to it. The opposite was true of depth of processing of lexical items: as depth of processing of lexical items increased, so did the ability of the participants to accurately recognize these lexical items. In regard to RQ7, results showed that the noticed, detected, noticed+ TA, and noticed+ silent intake levels all had significantly higher grammatical recognition scores than lexical recognition scores.
Chapter Five: Discussion and Conclusion

Discussion

**Effect of Thinking Aloud on Recognition, Production, and Comprehension.**

The results of this study revealed no reactivity for the recognition posttest, controlled written production posttest, and comprehension posttest. There were no significant differences between a) the Grammatical Noticed+ TA Intake Group and the Grammatical Noticed+ Silent Intake Group and b) the Lexical Noticed+ TA Intake Group and the Lexical Noticed+ Silent Intake Group for any of these assessment measures.

It is interesting that no reactivity, neither positive nor negative, was found in the current study. After all, Bowles’ (2010) meta-analysis reported that thinking aloud during a reading passage seems to facilitate form recognition, thinking aloud during a production task has a small, negative effect on productive form learning, and thinking aloud during a reading task slightly improves comprehension. However, several previous studies (Bowles, 2008; Bowles & Leow, 2005; Leow & Morgan-Short, 2004) did show that think aloud protocols did not significantly affect L2 learners’ processing in reading tasks, except in latency. One reason that perhaps contributed more to the absence of reactivity may be the amount of time allowed to process the L2 input, that is, it is possible that the lower levels of processing and attention that accompanied thinking aloud during this study’s reading task may have a played a role in keeping reactivity at bay.

Regardless, the fact that no reactivity was found in this current study is a definite advantage because it strengthens the internal validity of the study. In addition, since reactivity was not documented, it can be assumed that the cognitive processes were not heavily altered by
thinking aloud and that the think-aloud groups do, in fact, represent a true reflection of normal
cognitive processing (Bowles, 2010).

**Effect of Level of Intake on Recognition.**

The second research question aimed to determine whether level of intake has an effect on adult L2 learners’ recognition of a) a grammatical form and/or b) lexical items from a written text as measured by recognition tests.

In order to answer the first part of this research question, a one-way ANOVA with Grammatical Intake Group as the independent variable and Recognition Posttest Grammatical Score as the dependent variable was conducted. Scheffé’s method post hoc comparisons revealed a statistical difference only between the Grammatical Attended Intake Group and the Grammatical Noticed+ TA Group. This difference can be explained by several factors. First, given that participants from all experimental groups did in fact recognize at least some instances of the target grammatical form (see Table 18), it appears that all the experimental conditions of this study incited participants to process the input enough to take in at least some parts of the reading passage. In other words, all intake groups had participants who completed the input processing stage of the learning process in SLA, and were thus able to take in some input. The lack of a significant difference between the Grammatical Detected Intake Group, one of whose characteristics is unawareness, and both Grammatical Noticed Intake Group and the Grammatical Noticed+ TA Intake Group, one of whose characteristics is awareness, it appears that intake in the form of recognition can be achieved in the absence of awareness, thus supporting Leow (forthcoming).

The significantly better grammatical recognition scores of the Grammatical Noticed+ TA Intake Group in comparison with those of the Grammatical Attended Intake Group exemplify
that the amount of input converted into intake is differential based on several variables. Indeed, as manipulated in the research design, the amount of time spent processing the L2 data appeared to have played an important role at this early stage; the mean times as measured in seconds for the Grammatical Attended Intake Group were 99.06 for the target grammatical passage and 95.35 for the target lexical passage when compared to the Grammatical Noticed+ TA Group’s means of 461.54 for the target grammatical passage and 461.54 for the target lexical passage. The Grammatical Noticed+ TA Group had a considerably longer amount of time to pay attention to the reading passages, which also appeared to have allowed them to reach higher levels of depth of processing, awareness, and cognitive effort, all of which could have facilitated greater recognition (cf. discussion on depth of processing below).

Overall, it does seem logical that the noticed+ TA intake level should achieve higher recognition scores than the lower intake groups. The significant difference in grammatical recognition scores between the Grammatical Noticed+ TA Group and the Grammatical Attended Intake Group shows the highest and lowest ends of the attentional and temporal spectrum and if the experimental conditions were able to reveal any differences between grammatical intake levels in grammatical recognition, minimally it should occur between these two intake levels. It very well may be that the current study’s operationalizations of level of intake need to be further refined in order to reveal recognition differences between the other grammatical intake levels although the results do appear to support the Model’s postulation that intake may be derived from any of the three phases but what is crucial is whether such intake is further processed by the learner.

The type of experimental task may have also played a role in the lack of significant differences among the lower grammatical intake levels. Reading tasks, in comparison to
problem-solving tasks (i.e., Leow 2000, Rosa & Leow, 2004 Rosa & O’Neill, 1999), intuitively are less conducive to learner involvement with the task.

The fact that participants from all grammatical intake groups were able to recognize some instances of the target grammatical form on the Recognition Posttest also may signal a grammatical target form like third-person singular form of regular –AR verbs in the preterit is not an exceedingly difficult one for beginning L2 learners of Spanish to take in, even under attentional and time constraints.

A second one-way ANOVA was performed to answer the second part of this research question: whether level of intake has an effect on adult L2 learner’s recognition of a lexical form. The results demonstrated that Lexical Intake Group did affect the Percent Recognition Posttest (lexical score); Scheffé’s post hoc revealed statistical differences between the Lexical Noticed+TA Group and a) the Lexical Noticed Intake Group, b) the Lexical Detected Intake Group, and c) the Lexical Attended Intake Group.

These results underscore the role of type of linguistic item at the early stage of input processing due to the lexical item being either more or less salient. In other words, since more differentiation between groups occurred in the lexical intake groups than in the grammatical intake groups, it may be that the increased saliency (in terms of newness) of lexical items contributed to this difference.

Similarly, there may have been increased differentiation between lexical intake levels due to participants having more difficulty recognizing lexical items in comparison to recognizing the instances of the target grammatical items. If in fact the target lexical items were more difficult
than the target grammatical form, then this increased difficulty might have caused the higher number of distinctions between intake levels for lexical items.

Overall, these findings appear to refute the postulations of researchers (e.g., Gass, 1988; Schmidt, 1990 and elsewhere) who place some role of awareness at this early input to intake stage of the learning process. Indeed, the results from the second research question appear to provide empirical support for the early stage of Leow’s (forthcoming) model in which he postulates that candidacy for intake may originate from different levels of intake that may or may not be accompanied by some level of awareness. As the results indicate, even the attended intake group that paid attention peripherally to the target L2 data in the input demonstrated some ability to take in (as measured on the post-exposure recognition assessment task) some target items, both grammatical and lexical.

The results also appear to suggest that with an increased amount of time to process the L2 data, participants may take in more L2 data when compared to a minimal amount of time to do so (cf. noticed+ TA intake level vs. attended intake level); in other words, deeper processing of intake leads to more intake. Regarding the difference in differentiation between intake levels of grammatical intake groups and lexical intake groups, one explanation may have to do with the nature of a grammatical form: after seeing ten instances of a grammatical form in a reading passage, a learner may start to make the connections and process in a way that is necessary for further processing.

**Effect of Level of Intake on Production.**

The third research question of this study examined whether level of intake has an effect on adult L2 learner’s controlled written production of a) a grammatical form and/or b) lexical items from a written text as measured by controlled written production tests. Results revealed
that level of intake has no significant effect on neither grammatical nor lexical production. However, it is important to remember that production scores showed a near-floor effect, thus limiting the investigation of the influence of other factors on this effect.

The results of this research question are not surprising considering Leow’s (forthcoming) stages of the learning process in SLA. According to his model, intake processing and output processing (resulting in production) must both occur before the L2 learner is hypothetically able to produce a given form. Although other theoretical postulations (Chaudron, 1985; Gass, 1988; McLaughlin, 1990; Schmidt, 1990; Tomlin & Villa, 1994; Robinson, 1995; VanPatten, 1996, 2004) are not as fine-grained as Leow’s, they all do predict that production (output) can only occur after intake, and some give more detail about the intermediary phases that need to occur between intake and output (Gass, 1988; VanPatten, 1996, 2004). It appears to be theoretically improbable that L2 learners should be able to produce before intake processing and output processing. Given that RQ2 showed that the Grammatical Noticed+ TA Intake Group recognized significantly more than the Grammatical Attended Intake Group and that the Lexical Noticed+ TA Intake Group recognized significantly more than each of the three lower lexical intake groups, it seems that the noticed+ intake groups were the only ones who took in a large amount of targets, but not even these noticed+ intake groups were able to produce the targets at a rate any different from the lower groups. Therefore, it seems unlikely that even the highest intake group would have gone beyond the intake processing stage to the internal system, two necessary steps before being able to theoretically achieve output (production). Leow’s model predicts that input may be taken in, but that does not necessarily imply further processing. In order for further processing to occur in the case of the participants in the current study, they likely would have needed more time to interact with the reading passages and additional practice.
with the targets to push the intake through intake processing; further effort would then be required to push through output processing to finally achieve production of the output.

*The characteristics of the three levels of intake (see Table 5)*
Table 5) discussed in the current study also do not seem conducive to production of the target grammatical form and target lexical items. Detected intake and attended intake are theorized to have minimal depth of processing and either no or very little cognitive effort; only noticed intake is hypothesized to have maximally a very low depth of processing and low cognitive effort. Looking back to the previous studies on depth of processing reviewed in Chapter Two of the current study, preliminary evidence does suggest that depth of processing facilitates several aspects of L2 learning in the early stages (Leow et al., 2008; Morgan-Short et al., 2012; Qi & Lapkin, 2001; Shook, 1994; Tulving, 1975). Therefore, it appears that some depth of processing needs to be present in order for aspects of L2 learning such as production to occur (cf. the discussion of research question 5 that addressed the relationships between depth of processing and subsequent performances). Since these three levels of intake (noticed, detected, attended) have little to no depth of processing, it is not surprising that the participants in the current study had very low production scores and that no group had significantly better production scores than another group.

**Relationship between Eye Fixation Time Measures and Recognition.**

The fourth research question aimed to determine whether there were any relationships between L2 learners’ eye fixation time measures (subtracted variants of first fixation duration, gaze duration, second pass time, total fixation time) and their abilities to recognize a) a grammatical form and/or b) lexical items in written L2 input as measured on recognition posttests.

The results indicated that there nearly no statistically significant relationships between fixation measures and recognition of a grammatical and/or lexical form. In order to understand the non-significant relationships between fixation measures and
recognition in this research question, it is important to consider the task performed during
eye-tracking. Although there were differences in the average amount of time taken to
complete the reading passage according to intake group, the amount of time overall was
very small: the average times for the grammatical reading passage varied from 99.06
seconds to 57.95 (Grammatical Noticed Intake Group and Grammatical Attended Intake
Group, respectively) and the average times for the lexical reading passage varied from
95.35 seconds to 54.71 (Grammatical Noticed Intake Group and Grammatical Attended
Intake Group, respectively). It is possible that had participants spent longer on the
reading passages and read without feeling the pressure of a perceived time limit, then
there would have been more significant relationships between fixation measures and
recognition. However, had participants been encouraged to take their time reading, then
it is also unlikely that the current study would have been able to address the proposed low
levels of intake that are characterized by little to no processing, among other
characteristics. Following this strand, it is also a possibility that recognition is only
sensitive to eye-tracking measures when participants read at a slower pace and/or process
at a higher level, both of which did not occur in the current study. Gass (1988 and
elsewhere), Schmidt (1990, 1993), and Tomlin and Villa (1994) all appear to support the
role of attention in intake (operationalized here as recognition), however it is unclear
whether this attention can be peripheral or selective. Given the results of the current
study, the arguably small amount of attention exerted by the noticed, detected, and
attended intake groups may not be enough to facilitate enough recognition to show any
relationships between fixation measures and recognition. The E-Z Reader (Reichle et al.,
1998; Reichle et al, 2006; Reichle et al., 1999, Reichle et al, 2009) processing model
adopted in the current study to explain how word identification, visual processing, attention, and oculomotor control jointly influence when and where the eyes move during processing posits that attention exerted during reading is similar to a spotlight shifting serially from word to word; only when lexical access of the currently fixated word is complete does attention shift to the next word. Since the participants in the current study had very little time to read the reading passages, it is possible that lexical access of one word was not able to be successfully completed before the participants had to move on to the next word; this could explain the lack of significant relationships between fixation measures and recognition. This compliments Leow’s (forthcoming) model if complete lexical access is considered to be processing in the developing system because without this processing, learning cannot occur.

**Relationship between Depth of Processing and a) Recognition and b) Production.**

This research question aimed to determine whether there existed relationships between depth of processing during a reading task and subsequent recognition and written production of a) a grammatical form and/or b) lexical items embedded in the reading task. In summary, analyses revealed that as depth of processing of the instances of the grammatical form increased, participants were more likely to correctly produce those instances on the controlled written posttest; the same was not true for the recognition of instances of the grammatical form. Regarding the lexical items, as depth of processing of lexical items increased, it was more probable that participants would be able to recognize those items on the recognition posttest; the same was not the case for the production of lexical items.

The results of this research question seem to clearly indicate the important role that type of linguistic item plays in relation to depth of processing. Given Goldschneider and DeKeyser’s
(2001) definition of salience as the ease of hearing or perceiving a given structure, the newness of the target lexical items in comparison to the instances of the target grammatical form creates an increased salience. While each target lexical item was a discrete item of which participants possessed no prior knowledge, participants did have accurate prior knowledge of the stem of the instances of the target grammatical form. In other words, the target lexical items can be considered +new and the instances of the target grammatical form can be considered, in a sense, -new; only the morpheme –ó was +new. Also, the morpheme of the target grammatical form of this study may have quickly grown repetitive as participants were exposed to ten instances of verbs with this morpheme within a very short amount of time, thus decreasing it salience. This situation resulted in instances of the target grammatical form that lacked salience in the form of newness in both the stem and the morpheme. Therefore, it is possible that depth of processing of lexical items was significantly more likely to co-occur with accurate lexical recognition because at the recognition level, the lexical item remains in working memory due to its salience as a new, previously-unseen item. It is logical that the more salient item, which in this case are lexical items, should be recognized at a higher rate. However, production of the target lexical items was constrained by the amount of discrete items: participants were exposed to ten previously-unseen lexical items, each functioning as a separate chunk with no underlying rule to aid in deciphering meaning. It appears that the salience of the target lexical items allowed them to enter working memory and thus facilitated recognition, but there was no further processing due to the constraint of amount of items.

In regard to the target grammatical form, the relationship between grammatical depth of processing and recognition may be due to all participants, whether using a lower or higher depth of processing, being able to recognize instances of the target grammatical form. After all, the
instances of the target grammatical form were not ten discrete linguistic chunks, as in the case of the lexical items, but rather ten instances with one common, underlying rule. Furthermore, instances of the target grammatical form are not salient in the same way that the target lexical items are salient because participants already knew the meaning of the stem; it was only the morpheme that was new to them. Therefore, this decreases the saliency of the instances of the target grammatical form. However, given that the participants already possessed accurate prior knowledge of the stem of the instances of the target grammatical form, the morpheme was salient to participants because it was new to them. It appears that in order for grammatical production to occur, participants had to process the morpheme more deeply, thus causing it to go beyond working memory to actual integration into the developing system, which then facilitated production.

Additional evidence to support these results also comes from VanPatten (1996, 2004): he stated that the availability of processing resources is provided in part by the learner’s ability to access lexical items (or, in this case, lexical meaning of the stems of the instance of the target grammatical form) that have already been incorporated into their developing system. Therefore, in addition to the explanation of saliency above, it may be that the prior knowledge of the stems allowed processing resources to be devoted to the previously-unknown morpheme –ó, thus resulting in the significant relationship between depth of processing of a grammatical form and production. In other words, the lexical items had enough salience due to their newness to be significantly associated with accurate recognition, but they were not processed enough to move beyond working memory and into the developing system to allow the possibility of production.

The current study validates the role that depth of processing may play in L2 development such as recognition and production and also supports Leow’s model’s postulation of the crucial
role depth of processing plays in the L2 learning process in addition to previous empirical studies that found links between depth of processing and a) retention (Craik & Tulving, 1975; Hsieh, Moreno, & Leow, forthcoming), b) resolving and reforming former written production errors (Qi & Lapkin, 2001), and c) recognition and written production (Hsieh, Moreno, & Leow, forthcoming; Shook, 1994), and d) oral production (Hsieh, Moreno, & Leow, forthcoming).

The results of this research question highlight the importance of the consideration of type of linguistic item in any discussion of depth of processing. Salience in terms of newness may facilitate short-term storage of lexical items in working memory, however, deeper processing appears to be necessary in order for a higher-level skill like production of lexical items to be able to occur. On the contrary, deeper processing of a +newness morpheme of a grammatical form appears to facilitate production.

**Relationship between Depth of Processing and Comprehension.**

The sixth research question investigated possible relationships between depth of processing and comprehension of a reading passage, as measured by a multiple-choice comprehension posttest. In review, as Grammatical Depth of Processing Binary increased, the likelihood that participants correctly answered the corresponding comprehension questions also increased. There was no significant relationship between Lexical Depth of Processing Binary and Lexical Comprehension Binary.

In order to fully explain the non-significant relationship between lexical depth of processing and comprehension, it is necessary to look back to the discussion of RQ5: it was posited that in the presence of deeper processing, the lexical item may remain in working memory due to its salience in terms of newness, thus explaining the relationship between lexical depth of processing and recognition. The non-significant relationship between lexical depth of
processing and comprehension is similar to the non-significant relationship between lexical depth of processing and production in RQ5: the ability to recognize the lexical items may not have led to internalization of the items, thus rendering them incapable of being further processed, produced, and comprehended. According to Leow’s (forthcoming) model, output in the form of production cannot logically occur without intake; if intake is considered to be recognition and a participant has not demonstrated accurate recognition, then it seems improbable that any higher-level skill such as production or comprehension can occur.

On the other hand, the significant relationship between grammatical depth of processing and comprehension is related to the significant relationship between grammatical depth of processing and production in RQ5: a higher depth of processing may have allowed the internalization that is necessary for both production and comprehension. The presence of a higher depth of processing did not significantly relate to more accurate recognition perhaps because even participants showing a low depth of processing were able to recognize instances of the target grammatical form.

The current study sheds light on the issue of depth of processing and comprehension of a reading passage because it includes the added element of type of linguistic item. Although Leow et al. (2008) and Morgan-Short et al. (2012) both included a lexical item and a verb form, there were many differences between the two targets that make a direct comparison rather difficult. The fact that higher grammatical depth of processing was significantly related to comprehension but higher lexical depth of processing was not related to comprehension yet again highlights the differences between a grammatical form and lexical items. A framework proposed by Craik (2002) and Craik and Lockhart (1972) in cognitive psychology may be able to explain the significant relationship between grammatical depth of processing and comprehension. They
predicted that memory for items depends on the depth of processing during encoding. If memory for items is considered to be comprehension posttest scores, then the prediction indicates that depth of processing and comprehension may be positively related. Furthermore, these two sources posit that a high depth processing is more likely than a shallow or low depth of processing to lead to memory recall, which can be seen as processing meaning. In the current study, the significantly increased likelihood of higher depth of processing of instances of the target grammatical form to co-occur with accurate comprehension lends empirical evidence to Craik and Craik and Lockhart’s framework.

These results are important in light of two previous studies investigating depth of processing and comprehension. While Leow et al. (2008) noted that deeper processing of form (a lexical item, a verb form, and two articles) appeared not to play a detrimental role on participants’ comprehension of a written text, a replication by Morgan-Short et al. (2012) revealed a small but reliable positive relationship between level of processing and comprehension. However, Leow et al. had low participant numbers per cell and qualitative data showing mostly very low levels of processing; this may explain the results from Leow et al. Therefore, the fact that this current study did find a significant relationship between depth of processing and grammatical recognition may be considered to validate results from Morgan-Short et al.

**Difference between Recognition of Grammatical Form vs. Lexical Item.**

The seventh research question aims to determine whether there are any differences between recognition of a grammatical form and lexical items in beginning Spanish learners. Paired-samples t-tests show that all intake levels (noticed, detected, noticed+ TA, noticed+
silent) except the attended intake level had significantly higher Percent Grammatical Recognition Posttest scores than Percent Lexical Recognition Posttest scores.

Intuitively, it seems logical that it would be easier to recognize instances of a grammatical form versus lexical items on a post-reading test because all instances of the grammatical form are subsumed by one underlying rule. The L2 learner does not even have to learn the underlying rule to be able to conclude that all verbs with a certain morpheme (in this case, -ó) are from a certain grammatical form and have something in common. On the contrary, a set of unrelated lexical items have no one rule that a learner can use to make meaning of them. Each lexical item is distinct and functions separately; they share no inherent similarities that may help an L2 learner remember them. This sense of grouping in the case of the grammatical form may have assisted the participants in being able to recognize more instances of the grammatical form when compared with recognition of the lexical items.

A possible contributing factor to the differences revealed between instances of the target grammatical form and target lexical items is prior semantic knowledge of the verb stems. As shown on the verb stem meaning pretest, all participants demonstrated accurate knowledge of the meaning of the infinitives of the instances of the target grammatical form. Therefore, while completing the reading passages, participants already had some ‘bootstrap’ knowledge of the instances of the target grammatical form. This was not necessarily the case with the target lexical items. Participants may have possessed some prior knowledge of the root of one of the target lexical items (e.g. cabez- of cabezada) but this is a different scenario than that of the target grammatical form. Even if a participant recognized the root of one of the target lexical items, these roots are not so transparent as to facilitate easy discernment of the meaning of the target lexical item. On the other hand, knowledge of the meaning of the infinitive of an instance
of the target lexical item is much more transparent in that it does have the same lexical meaning as the instance of the target lexical item. Therefore, the demonstrated accurate prior knowledge of the infinitive of the target grammatical form may have acted as a bridge in facilitating better recognition than in lexical items.

There are at least two possible explanations as to the lack of significant results at the attended intake level. First, it is very possible that the only reason that the attended intake level did not have significantly higher Percent Grammatical Recognition Posttest scores than Percent Lexical Recognition Posttest Scores is because of the low effect size and weak power. In contrast, all other levels had high power and all other levels except the noticed+ TA level had at least medium effect size. Therefore, it seems very possible that the attended intake level could also have significantly higher Percent Grammatical Recognition Posttest scores than Percent Lexical Recognition Posttest scores if the power and effect size were at higher levels.

Another possible explanation as to why the attended intake level showed no significant differences between scores on the Percent Grammatical Recognition Posttest and the Percent Lexical Recognition Posttest is that this level, characterized by only a very low level of selective or peripheral attention, did not reach a certain attentional threshold necessary to even begin to demonstrate differences between the recognition of a grammatical form versus that of lexical items.

As noted earlier in Chapter Two, previous studies on type of linguistic item have shown no clear-cut answer regarding the role that type of linguistic item plays in L2 development, but then again, only a handful compared a grammatical form versus that of lexical items (Greenslade et al., 1999; Mackey et al., 2000; Smith, 2012; VanPatten, 1990); none of those looked at
recognition. The results concerning type of linguistic item from this current study provide a first step in propelling future research into this overlooked strand.

**Limitations and Directions for Future Research**

Like any empirical investigation, this current study has some limitations that serve to propel methodological refinement and new topics for future studies.

First of all, the current study represents a first attempt at separating criteria for grammatical versus lexical depths of processing and therefore the descriptors for both grammatical and lexical depth of processing need to be refined in future studies. Although the criteria for depth of processing of a grammatical form are based on the potential for processing the target form grammatically, the descriptors at the intermediate and high levels need to be further modified to best operationalize the depth of processing descriptions.

The current operationalization of detected intake needs a cleaner operationalization in future studies. Leow’s (forthcoming) intake characteristics (see Table 5
Table 5) posit that detected intake lacks awareness, however the current study has no way of demonstrating that this group was, in fact, unaware. The invisible boundary paradigm used in the attended intake condition is arguably enough to ensure unawareness at that level, however no such technique was used at the detected level and future operationalization of detected intake will need to fine-tune the current study’s attempt to do so.

The type of experimental task used in the current study served as a limitation because intuitively, reading passages do not incite the same levels of processing as do other tasks, such as a problem-solving task (i.e., Leow, 2000). Reading tasks are certainly commonplace in the L2 classroom, and therefore it is important to use them as tasks in empirical research, however a problem-solving task may provide better differentiation between intake levels and may be better able to induce more processing. Future research clearly needs to address type of task.

As elaborated upon in the discussion of RQ7, one major difference between the instances of the target grammatical form and the target lexical items is that the instances of the target grammatical form all share one underlying rule. Since this is not the case for the target lexical items, this creates a difference that was not controlled for in the current study although it is a normal feature of type of linguistic item. It would be of interest to further examine this concept of an underlying rule for both grammatical forms and lexical items. For example, a future study could use as targets both a grammatical form (with an underlying rule) and several carefully-chosen lexical items with an underlying rule, such as transformations like moneda – monedero and lápiz – lapicero or sets of words with a common root like pelo, peluca, peluquero, peluquería, peluquín, etc.

As evidenced in the current study, participants in the Noticed+ TA Intake Groups did achieve a high level of depth of processing, characterized by hypothesis formation. This relates
to Schmidt’s (1990) awareness at the level of understanding that involves restructuring and system learning because it involves the ability to test hypotheses. Therefore, it can be hypothesized that participants in the current study achieving a high level of depth of processing also underwent restructuring and system learning (cf. Leow, 2012, for a correlation between levels of awareness and depth of processing). Future studies should test this by including a generalization posttest of new items.

Finally, this study only tested the first stage, namely, the input-to-intake stage, of the L2 learning process as postulated by Leow’s model. Future research needs to also test the remaining stages of the model.

Conclusions and Implications

This study set out to test Leow’s (forthcoming) Model of the L2 learning process in relation to his postulations (levels of intake and the role of depth of processing) of the early stage of the L2 learning process while also addressing type of linguistic item and the role of reactivity. The current study found that the noticed+ TA intake level performed significantly better on a recognition posttest than did the Grammatical Attended Intake Group (in the case of the Grammatical Noticed+ TA Intake Group) and the Lexical Noticed, Detected, and Attended Intake Groups (in the case of the Lexical Noticed+ TA Intake Group). Level of intake had no significant effects on production and there were no significant relationships between fixation measures and recognition. Regarding depth of processing, as grammatical depth of processing increased, so did accurate production and comprehension. The reverse was true of lexical depth of processing: as depth of processing lexical items increased, recognition also increased. Lastly, significantly higher recognition score were found for the grammatical form than for lexical items.
in the noticed, detected, noticed+ TA, and noticed+ silent intake levels. Notably, reactivity was not found in any of the tests.

The results and discussion of the results of the seven research questions highlight several reoccurring themes that serve as an overarching summary of the information gained from this current study.

First, as postulated by the model, different levels of intake do appear to exist, as documented in the superior recognition scores of the noticed+ TA intake groups in comparison to the attended intake group (in the case of the target grammatical form) and noticed, detected, and attended intake groups (in the case of the target lexical items). Level of intake did not significantly affect production for neither the target grammatical form nor the target lexical items; however deeper depth of processing was significantly related to accurate production. Therefore, it appears that level of intake can explain for lower-level skills, such as recognition, but not for higher-level skills that require integration into the developing system. This is not surprising, given that the intake levels defined in the current study addressed low-to-nonexistent levels of attention, processing, and cognitive awareness, among other indicators. Leow’s model (forthcoming) predicts that while the lower levels of intake may result in linguistic data being taken in and potentially recognized subsequently, they are not likely to enter the developing system unless further processing occurs. In other words, according to the model, all levels of intake have the ability to enter working memory, however the results of this study confirm the model’s predictions that only intake in working memory that receives further processing has the potential of moving on to the developing system. It comes as little surprise that there are no significant differences among the three lower intake groups (Grammatical/Lexical Noticed Intake Group, Grammatical/Lexical Detected Intake Group, Grammatical/Lexical Attended
Intake Group) because the depth of processing characterizing each level is essentially the same: very low to nonexistent. Given the superior performance of the noticed+ TA groups in recognition, it appears that the very low levels of depth of processing and low attention that characterizes the noticed intake groups is not enough; the threshold for further processing may occur between the noticed intake level and the noticed+ intake level. It seems that some characteristic(s) of the noticed, detected, and attended intake levels was somehow fundamentally different from that of the noticed+ TA intake level in terms of recognition. The noticed+ intake level seems to allow participants to process further, underscoring the model’s postulation that while intake may be derived from several phases, the potential for internalization depends upon subsequent processing of such intake.

Another underlying theme visible in the results and discussion sections is the differences between types of linguistic item. As detailed in the discussion of RQ7, instances of grammatical form share an underlying rule: once this rule is accurately acquired, it can simply be applied to old and new instances. On the other hand, each lexical item is unique in and of itself; there are no underlying rules to assist the L2 learner. While it is true that some groups of lexical items may share a common root, the lexical meaning of each item is not always transparent, even if the lexical meaning of the root is already known. This consideration is an important factor in explaining the overwhelmingly superior recognition scores of instances of the target grammatical form in comparison to the target lexical items at almost all levels of intake.

The differences in types of linguistic item also extend to the frameworks that may be necessary to explain the development of each. Regarding a grammatical form, Leow’s model (forthcoming), Craik (2002) and Craik and Lockhart’s (1972) framework that claim that memory for items depends on the depth of processing occurring during encoding may be useful in
explaining development of a grammatical form: the internalization necessary to produce and comprehend may be facilitated by depth of processing.

The differences between types of linguistic items also support different levels of processing. The results and discussions of the relationships between grammatical/lexical depth of processing and a) recognition, b) production, and c) comprehension reveal that deeper lexical processing was more likely to occur in the presence of accurate recognition, while deeper grammatical processing was more likely to occur in the presence of accurate production and comprehension. While the lexical items were more salient in terms of newness, the lack of significant relationships between higher processing and a) production and b) comprehension suggests that the newness of the lexical terms was not enough to incite the higher processing that seems to be essential for production and comprehension. On the other hand, the significant relationships between grammatical depth of processing and a) production and b) comprehension indicate that the less-salient (in terms of newness) verbal morpheme –ő facilitated higher processing and thus created ideal conditions for accurate production and comprehension. After all, participants already possessed accurate prior knowledge of the verb stem (cf. role of old prior knowledge in the model) and cannot be considered +new. Therefore, it becomes clear that the instances of the target grammatical form progressed from working memory into the developing system and were subsequently internalized, thus enabling production. On the other hand, the lexical items, though accurately recognized as predicted by Leow (forthcoming), did not move beyond working memory and therefore were unable to be produced or comprehended. Overall, it appears that type of grammatical item influences depth of processing due to factors such as newness and prior knowledge.
The last overlying theme identified in the results and discussion is that of time spent on the reading passages. Overall, the noticed+ TA intake levels spent much more time on the reading passages than did the noticed, detected, and attended intake levels. This large amount of additional time may very well have allowed participants at the noticed+ TA intake levels to process more and process deeper, expend more cognitive effort, and focus attention on more linguistic data than did the participants in the lower groups. According to Leow’s (forthcoming) model, all of these processes facilitate the integration of the linguistic data from working memory to the developing system; if intake is not further processed then the model predicts that such data will not integrate into the developing system and thus cannot be learned. Furthermore, although the detected and noticed intake levels did spend progressively more time on the reading passages than did the attended intake levels, these differences were small and overall, the time spent was very little. This may account for a lack of differentiation between intake levels in recognition. In addition, the small amount of time that participants in the noticed, detected, and attended intake levels spent on the reading passages may be the cause of non-significant results of the research question investigating the relationship between eye-fixation time and recognition. It is plausible that participants did not have enough time to process the targets: they may have either simply stayed in working memory and the insufficient amount of time prohibited them from being fully processed. This possibility is compatible with the lexical access step of the E-Z Reader processing model (Reichle et al., 1998; Reichle et al, 2006; Reichle et al., 1999, Reichle et al, 2009), which states that complete lexical access of the target should ideally occur before the reader moves on to the next word. If one assumes that complete lexical access involves the developing system, then participants limited by time would likely not have been able to process the target to the extent necessary for learning to take place. However, had more time been
provided for participants at all levels to complete the reading passages, it is unlikely that the study would have been able to address lower levels of intake and low-to-nonexistent processing.

The results of this study have direct implications for the L2 classroom and adult L2 development. First, simple recognition at the word level is certainly possible for learners who read a text and have only a limited amount of time, but it appears unlikely that written production and comprehension are possible for learners who do not internalize the intake and thus process it further. An increased amount of time seems to facilitate conditions in which the L2 learner can employ a variety of learning constructs, such as processing, attention, cognitive effort, all of which may lead to an increased probability of integration into the developing system. L2 learners should be given sufficient time to read and interact with a text at their own pace in order to ensure maximum benefit. L2 instructors need to take into consideration that recognition activities (e.g., having learners search for and circle a target form in a text or indicate which words from a list appeared in a previous reading) start the L2 development process and are an appropriate first step, but that they should precede production and comprehension tasks.

This study also indicates that L2 instructors should choose exposure tasks that facilitate and encourage higher depths of processing. Given that a considerable amount of high depth of processing was documented in the current study, it appears that under the right conditions, readers can indeed process deeply during a reading task. Instructors should remind learners to reading for meaning and at their own pace, in addition to encouraging them to use additional techniques to increase the probability for processing, such as thinking their thoughts aloud.

L2 instructors should also be mindful of the differences between teaching grammatical forms and lexical items. As the results of this study indicate, the morpheme of conjugated verbs whose stems learners already know may incite greater processing; this highlights the importance
of ensuring that learners know the meanings of verb stems before learning new conjugations.

This study also contributes to the attentional strand of SLA in that it addresses several key issues that had previously been in need of further investigation. First of all, previous studies had reported results that hinted at the existence of different levels of intake (see Leow, 1997, 1998b, 2001a, Martínez-Fernández, 2008, Rosa & Leow, 2004, among others for noticed intake; Leung & Williams, 2011, Williams, 2004, 2005 for detected intake; Godfroid et al., 2010 for attended intake), all differing in characteristics such as attention, processing, cognitive effort, and awareness, among others. The current study, which is the first to empirically test these levels of intake postulated in Leow’s (forthcoming) model, contributes to this new strand of SLA research and raises the potential for a better understanding of the early stage of the L2 learning process from a more fine-tuned perspective.

The current study is also among the first to attempt to fill in the gap of a working definition and operationalization of depth of processing. By investigating the relationships between depth of processing and several variables (recognition, production, comprehension), in addition to taking into consideration type of linguistic item, this study provides substantial empirical evidence regarding depth of processing in order to guide both future theoretical discussion and empirical investigations. Furthermore, this study has established the use of separate criteria for grammatical versus lexical depth of processing.

Eye-tracking has been established as a useful measure of attention allocation within the field of psycholinguistics (i.e., Godfroid et al, 2010; Rehder & Hoffman, 2005a, 2005b; Theeuwes & Van de Stigchel, 2009); however this online technique is still new to the SLA field. The current study is innovative in that it triangulated concurrent data from both eye-tracking and think-aloud protocols in order to investigate the research questions. Findings revealed that
fixation measures were not significantly related to recognition at the established intake levels, thus providing findings different from cognitive science studies that successfully used eye-tracking to measure attention (Rehder & Hoffman, 2005a, 2005b; Theeuwes and Van der Stigchel, 2009). However, it is important to keep in mind that this lack of significant findings in regards to fixation measures and recognition may be due to the short amount of time participants took to read.

This study also contributes to the reactivity strand in that it provides evidence that thinking aloud while reading may not cause reactivity on subsequent recognition, production, and comprehension tests; this is of interest considering that Bowles’ (2010) meta-analysis found that thinking aloud during a reading passage had a positive effect on receptive learning, a negative effect on written production, and a small positive effect for comprehension.

In summary, this study serves as a first step in identifying and empirically testing the effects of Leow’s (forthcoming) phases of intake on and the role of depth of processing in adult L2 development within his Model of the L2 learning process. By employing both eye-tracking and think-aloud protocols, it counts on two forms of concurrent data to operationalize difficult constructs such as attention and depth of processing, respectively. The inclusion of a silent condition allowed the study to conclude that thinking aloud did not cause reactivity on the subsequent recognition, production, and comprehension posttest. Furthermore, the incorporation of type of linguistic item as an independent variable is a first in the attentional/processing/comprehension strand of SLA and has provided valuable information regarding the differences in L2 processing between a grammatical form and lexical items. This first attempt at investigating level of intake, depth of processing, and type of linguistic item in L2 development has provided a foundation for future studies to further investigate and refine these
constructs. Needless to say, the other postulations made in Leow’s model need to be empirically tested.
Appendix A: Language Background Questionnaire

Participant # _____ / Initials ______

1. Gender: ___ Male     ___ Female

2. Age: ___

3. Country of birth: __________

4. What is your native language? ___ English     ___ Spanish     ___ Other

5. What languages do you speak at home? ____________________________
   a. If more than one, with whom do you speak each of these languages?
      ____________________________________________________________

6. In what language(s) did you receive the majority of your precollege education?
   _____________________________________________________________
   a. If more than one, please give the approximate number of years for each
      language. ____________________________________________

7. Have you ever been to a Spanish-speaking region for the purpose of studying
   Spanish? ___ yes      ___ no
   a. If yes, when? Where? ______________________________________
   b. For how long? __________________________________________

8. Other than the experience mentioned in Question 7, have you ever lived in a
   situation where you were exposed to a language other than your native language
   (e.g., by living in a multilingual community; visiting a community for purposes of
   study abroad or work; exposure through family members, etc.)? ___ Yes     ___ No
   If yes, please give details below.
<table>
<thead>
<tr>
<th>Experience</th>
<th>Experience</th>
<th>Experience</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Country/region</th>
<th>Language</th>
<th>purpose</th>
<th>From when to when</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

9. In the boxes below, rate your language ability in each of the languages that you know. Use the following ratings: 0: Poor, 1: Good, 2: Very good, 3: Native/nativelike

<table>
<thead>
<tr>
<th>Language</th>
<th>Listening</th>
<th>Speaking</th>
<th>Reading</th>
<th>Writing</th>
<th>Number of years of study</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10. Have you studied Spanish in school in the past at each of the levels listed below?

   If yes, for how long?

   a. Elementary school: __ No __ Yes: __ less than 1 year __ 1-2 years __ more than 2 years

   b. Junior high/middle school: __ No __ Yes: __ less than 1 year __ 1-2 years __ more than 2 years
c. Senior high school: __ No __ Yes: __ less than 1 year __ 1-2 years __ more than 2 years
d. University/college: __ No __ Yes: __ less than 1 year __ 1-2 years __ more than 2 years
e. Other (please specify: ____________): __ No __ Yes: __ less than 1 year __ 1-2 years __ more than 2 years

11. What year are you in school? __ Freshman __ Sophomore __ Junior __ Senior __ Graduate student __ Other

12. What is your major? ____________________________

13. On average, how often did you communicate with native or fluent speakers of Spanish in Spanish in the year prior to the start of this semester?
__ never __ a few times a month __ monthly __ weekly __ daily
Appendix B: Experimental Reading Passages

Version

Juanito: Grammatical target, Lexical control
Laura: Grammatical control, Lexical target

Juanito

35 Although textual enhancement occurred in some experimental conditions, that textual enhancement is not shown in this Appendix because the focus is showing the information conveyed in the reading passages.
carro a la universidad pero no vio a ningún otro estudiante. “¿Dónde están los otros estudiantes y por qué no hay nadie aquí?” Después Juanito miró el calendario de su teléfono. Hoy es sábado, el ocho de marzo y es el fin de semana. “¡No hay clases! Necesito una siesta porque tengo sueño.” Juan vuelve a casa a dormir un poco más. ¡Esta noche quiere salir otra vez y pasarlo bien!

Laura

Laura Jiménez tiene treinta y cuatro años y es detective privada en Miami. Es muy profesional y tiene un negocio que se llama ‘Investigaciones Confidenciales Jiménez.’ También es organizada y en su oficina tiene un fichero con muchos papeles. Su oficina está en el paseo más prestigioso de Miami: Collins Avenue. El lunes por la mañana Laura se levanta a las seis. Después de limpiar su oficina, navega la Red. A las tres llega una señora muy elegante y obviamente muy rica. “Esta señora seguramente es muy rica,” le dice Laura a la secretaria. “Hola, mi nombre es Victoria y mi apellido es Rodríguez Garzón” dice la señora. También dice que su parentela es muy prestigiosa; su padre es muy famoso. Laura está muy interesada y piensa en el caso de la señora. La señora dice que tiene una gran duda: no sabe dónde está su bebé. Dice, “mi bebé siempre desayuna muy bien y después nosotros salimos. Por la tarde nada en la piscina pero ahora no sé dónde está!” Laura escribe unos apuntes en su cuaderno de investigaciones. Laura, muy pensativa, pregunta, “¿Y no sabes dónde está? ¿Tu bebé no está en su cuna donde siempre duerme?” La señora dice que no; no habla más porque obviamente está muy ansiosa y triste. La secretaria le da una taza grande de té con limón. “Muchas gracias” dice la señora, y descansa por unos minutos. Después Laura dice, “Señora debes darme los cien dólares o un cheque.” La señora abre su cartera grande de Coach y ¡sale un pequeño perro! “¡Mi bebé, Fufú, allí estás!
dice la señora. Laura está feliz y camina con la señora a la puerta. Laura siempre está feliz cuando sus casos detectivos tienen una resolución. ¡Este caso es el relato más cómico de todos sus años como detective! ¡Ahora Laura quiere almorzar y salir con sus compañeros!

*Version*

Juanito: Grammatical control, Lexical target

Laura: Grammatical target, Lexical control

Juanito

Esta mañana cuando ve su reloj, dice: “¡Es hora de la clase de economía!” Después de vestirse, maneja su carro a la universidad pero no ve a ningún otro estudiante. “¿Dónde están los otros estudiantes y por qué no hay nadie aquí?” Después Juanito mira el calendario de su teléfono. Hoy es feriado, el veinticinco de diciembre y hay vacaciones. “¡No hay clases! Necesito una cabezada porque tengo sueño.” Juan vuelve a casa a dormir un poco más. ¡Esta noche quiere salir otra vez y pasarlo bien!

Laura

Laura Jiménez tiene treinta y cuatro años y es detective privada en Miami. Es muy profesional y tiene un periódico que se llama ‘Investigaciones Confidenciales Jiménez.’ También es organizada y tiene su tarea de la universidad en un cuaderno. Su oficina está en la escuela más prestigiosa de Miami. El lunes por la mañana Laura se levantó a las seis. Después de limpiar su oficina, navegó la Red. A las tres llegó una señora muy elegante y obviamente muy rica. “Esta señora seguramente es muy rica,” le dijo Laura a la secretaria. “Hola, mi nombre es Victoria y mi madre es María Garón” dijo la señora. También dijo que su trabajo es muy prestigioso; su padre era muy famoso. Laura estaba muy interesada y pensó en el caso de la señora. La señora dijo que tiene un mensaje importante: no sabe dónde está su bebé. Dijo, “esta mañana mi bebé desayunó como normal y después nosotros salimos. Por la tarde nadó en la piscina pero ahora no sé dónde está!” Laura escribió unos apuntes en su cuaderno de investigaciones. Laura, muy pensativa, preguntó, “¿Y no sabes dónde está? ¿Tu bebé no está en su cama donde siempre duerme?” La señora dijo que no; no habló más porque obviamente estaba muy ansiosa y triste. La secretaria le dio un vaso grande de leche para los nervios. “Muchas gracias” dijo la señora, y descansó por unos minutos. Después Laura dijo, “Señora debes darme los cien dólares o un cheque.” La
señora abrió su libro para escribir el cheque cuando ¡vieron un pequeño perro! “¡Mi bebé, Fufú, allí estás! dijo la señora. Laura estaba feliz y caminó con la señora a la puerta. Laura siempre está feliz cuando sus casos detectivos tienen una resolución. ¡Hoy seguramente es el lunes más cómico de todos sus años como detective! ¡Ahora Laura quiere almorzar y salir con sus compañeros!
Appendix C: Reading Instructions by Condition

Noticed Intake Condition

You are about to read two passages. Be sure to skim for meaning. You will be tested afterward on your OVERALL comprehension, so it is NOT IMPORTANT that you understand EVERY SINGLE word. There is a time limit. You will receive a small prize if you finish before the time limit. Pay attention to the enlarged words in bold print and use the mouse to click these words. Each paragraph is presented on a separate screen. There are 12 screens in total, six for each reading passage. When you are finished with one screen, press OK to advance to the next screen. We will not be able to help you in any way. Don’t forget to finish before the time limit and get your prize so read quickly!

Instructions for Detected Intake Condition

You are about to read two passages. Be sure to read for meaning. You will be tested afterward on your OVERALL comprehension, so it is NOT IMPORTANT that you understand EVERY SINGLE word. There is a time limit so read quickly. You will receive a small prize if you finish before the time limit. Pay attention to the words in bold print. Each paragraph is presented on a separate screen. There are 12 screens in total, six for each reading passage. When you are finished with one screen, press OK to advance to the next screen. We will not be able to help you in any way. Don’t forget to finish before the time limit and get your prize so read quickly!
Instructions for Attended Intake Condition

You are about to read two passages. Be sure to read for meaning. You will be tested afterward on your **OVERALL** comprehension, so it is **NOT IMPORTANT** that you understand **EVERY SINGLE** word. **There is a time limit so read quickly.** You will receive a small prize if you finish before the time limit. Each paragraph is presented on a separate screen. There are 12 screens in total, six for each reading passage. When you are finished with one screen, press the space bar to advance to the next screen. **We will not be able to help you in any way.** **Don’t forget to finish before the time limit and get your prize so read quickly!** Please note that some words may disappear and reappear as you read.

Instructions for Noticed+ Intake Condition

You are about to read two passages. Be sure to read for meaning. You will be tested afterward on your **OVERALL** comprehension. **THERE IS NO TIME LIMIT.** You may take as much time as you need. **Pay attention to the enlarged words in bold print.** In this experiment we are also interested in what you think about as you complete this task. In order to find out, we are going to ask you to **THINK ALOUD** in English from the time you start the task to when you finish the task. **We would like you to talk CONSTANTLY.** We don't want you to try to plan out what you say or try to explain to us what you're saying. Just act as if you are alone in the room speaking to yourself. What's most important is that you keep talking, and talk clearly and loudly enough into your microphone. **We will not be able to help you in any way.**
Instructions for Noticed+ Silent Condition

You are about to read two passages. Be sure to read for meaning. You will be tested afterward on your **OVERALL** comprehension. THERE IS NO TIME LIMIT. You may take as much time as you need. Pay attention to the **enlarged words in bold print**. We will not be able to help you in any way.
Appendix D: Items Pretest

Participant #: _____ / Initials _____

Start time _____

Instructions: On the next page, fill in the table indicating your knowledge of the following words:

- Check the ‘no’ column if you have not seen the word before and do not know its meaning.
- Check the ‘seen’ column if you have seen a word before, but do not know or remember its meaning.
- Write the meaning of the word in the ‘meaning’ column if you know what it means. Please do not guess.
<table>
<thead>
<tr>
<th>Word</th>
<th>No (I have not seen this word before and I do not know its meaning.)</th>
<th>Seen (I have seen this word before but I do not know or remember its meaning.)</th>
<th>Meaning (If you know the meaning, write it below)</th>
</tr>
</thead>
<tbody>
<tr>
<td>me levanto</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>peaje</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chiste</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>luciérnaga</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>maneje</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>se acostó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>entrada</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>taza</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bisturí</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cortina</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mascota</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>se acuesta</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>esclavo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>llegó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>paseo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>verosímil</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pensó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>marroquí</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (I have not seen this word before and I do not know its meaning.)</td>
<td>Seen (I have seen this word before but I do not know or remember its meaning.)</td>
<td>Meaning (If you know the meaning, write it below in English or Spanish.)</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>20. nado</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. portada</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. navego</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. cafetera</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. mimamos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. inglés</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. camino</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. ragú</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. termo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. negocio</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. descansó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31. miró</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32. abrazo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33. desayunó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34. popurrí</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35. canción</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36. manejó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37. traducirá</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38. mire</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No (I have not seen this word before and I do not know its meaning.)</td>
<td>Seen (I have seen this word before but I do not know or remember its meaning.)</td>
<td>Meaning (If you know the meaning, write it below in English or Spanish.)</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>39. comprará</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40. caminó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41. lomo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42. cabezada</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43. charló</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44. feriado</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45. tribu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46. cartera</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47. se quedó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48. pienso</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49. colibrí</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50. té</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51. yegua</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52. tren</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53. regresó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54. traje</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55. apellido</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56. iglú</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57. cuero</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>Seen</td>
<td>Meaning</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>58. transladáramos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>59. convenceremos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60. se levantó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61. desayuno</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62. sopa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63. descanso</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64. tomó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65. escuche</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66. remolacha</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>67. iraní</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68. llego</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69. mecánico</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70. oirán</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>71. sacará</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72. champú</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>73. arbol</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>74. molestan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75. mangú</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76. higo</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(I have not seen this word before and I do not know its meaning.)

(I have seen this word before but I do not know or remember its meaning.)

(If you know the meaning, write it below in English or Spanish.)
<table>
<thead>
<tr>
<th>No (I have not seen this word before and I do not know its meaning.)</th>
<th>Seen (I have seen this word before but I do not know or remember its meaning.)</th>
<th>Meaning (If you know the meaning, write it below in English or Spanish.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>77. habló</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78. doblemos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79. fichero</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80. relato</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81. suenan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>82. ají</td>
<td></td>
<td></td>
</tr>
<tr>
<td>83. navegó</td>
<td></td>
<td></td>
</tr>
<tr>
<td>84. toma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>85. lavó</td>
<td></td>
<td></td>
</tr>
<tr>
<td>86. carnet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>87. logremos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>88. empaque</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89. mesón</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90. abre</td>
<td></td>
<td></td>
</tr>
<tr>
<td>91. mono</td>
<td></td>
<td></td>
</tr>
<tr>
<td>92. interés</td>
<td></td>
<td></td>
</tr>
<tr>
<td>93. hablo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>94. barren</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95. pensamiento</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>(I have not seen this word before and I do not know its meaning.)</td>
<td>Seen</td>
</tr>
<tr>
<td>----</td>
<td>---------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>96.</td>
<td>carmesí</td>
<td></td>
</tr>
<tr>
<td>97.</td>
<td>duda</td>
<td></td>
</tr>
<tr>
<td>98.</td>
<td>generen</td>
<td></td>
</tr>
<tr>
<td>99.</td>
<td>se queda</td>
<td></td>
</tr>
<tr>
<td>100.</td>
<td>parpadearán</td>
<td></td>
</tr>
<tr>
<td>101.</td>
<td>regrese</td>
<td></td>
</tr>
<tr>
<td>102.</td>
<td>baila</td>
<td></td>
</tr>
<tr>
<td>103.</td>
<td>dulce</td>
<td></td>
</tr>
<tr>
<td>104.</td>
<td>fistularán</td>
<td></td>
</tr>
<tr>
<td>105.</td>
<td>bebé</td>
<td></td>
</tr>
<tr>
<td>106.</td>
<td>pondrán</td>
<td></td>
</tr>
<tr>
<td>107.</td>
<td>bailó</td>
<td></td>
</tr>
<tr>
<td>108.</td>
<td>manú</td>
<td></td>
</tr>
<tr>
<td>109.</td>
<td>reloj</td>
<td></td>
</tr>
<tr>
<td>110.</td>
<td>cuna</td>
<td></td>
</tr>
<tr>
<td>111.</td>
<td>lave</td>
<td></td>
</tr>
<tr>
<td>112.</td>
<td>espicháramos</td>
<td></td>
</tr>
<tr>
<td>113.</td>
<td>pregunto</td>
<td></td>
</tr>
<tr>
<td>114.</td>
<td>escuchó</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>Seen</td>
<td>Meaning</td>
</tr>
<tr>
<td>----</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>(I have not seen this word before and I do not know its meaning.)</td>
<td>(I have seen this word before but I do not know or remember its meaning.)</td>
<td>(If you know the meaning, write it below in English or Spanish.)</td>
</tr>
<tr>
<td>115. contempláramos</td>
<td></td>
<td></td>
</tr>
<tr>
<td>116. preguntó</td>
<td></td>
<td></td>
</tr>
<tr>
<td>117. charla</td>
<td></td>
<td></td>
</tr>
<tr>
<td>118. nadó</td>
<td></td>
<td></td>
</tr>
<tr>
<td>119. parentela</td>
<td></td>
<td></td>
</tr>
<tr>
<td>120. medito</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

End time _____
Appendix E: Recognition Posttest

Participant #:_____ / Initials _____

Start time _____

Instructions: Carefully read the list of words on the next pages and mark the ones that you recognize seeing from the text you just read.
<table>
<thead>
<tr>
<th>Word</th>
<th>Word DID appear in text</th>
<th>Word DID NOT appear in text</th>
<th>Not sure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. cuna</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. lave</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. escuchó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. contempláramos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. preguntó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. se acostó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. entrada</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. taza</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. bisturí</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. cine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. cortina</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. mascota</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. se acuesta</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. esclavo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. llegó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. baila</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. dulce</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. fistularán</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. marroquí</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. nadó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. parentela</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word</td>
<td>Word DID appear in text</td>
<td>Word DID NOT appear in text</td>
<td>Not sure</td>
</tr>
<tr>
<td>----------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>22. medito</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. cafetera</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. interés</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. hablo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. barren</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. ragú</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. termo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. negocio</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. descansó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31. miró</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32. abrazo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33. desayunó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34. popurrí</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35. canción</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36. manejó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37. mecánico</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38. paseo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39. verosímil</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40. desayuno</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41. sopa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42. descanso</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word</td>
<td>Word DID appear in text</td>
<td>Word DID NOT appear in text</td>
<td>Not sure</td>
</tr>
<tr>
<td>------</td>
<td>-------------------------</td>
<td>----------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>43. charló</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44. feriado</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45. tribu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46. cartera</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47. se quedó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48. pienso</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49. colibrí</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50. té</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51. yegua</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52. tren</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53. regresó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54. traje</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55. apellido</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56. iglú</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57. cuero</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>58. transladáramos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>59. convenceremos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60. se levantó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61. oirán</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62. sacará</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63. llego</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word</td>
<td>Word DID appear in text</td>
<td>Word DID NOT appear in text</td>
<td>Not sure</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>64. tomó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65. escuche</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66. remolacha</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>67. iraní</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68. lomo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69. cabezada</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70. pensó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>71. empaque</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72. mesón</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>73. arbol</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>74. molestan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75. mangú</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76. higo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>77. habló</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78. doblemos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79. fichero</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80. relato</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81. suenan</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>82. ají</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>83. navegó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>84. toma</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word</td>
<td>Word DID appear in text</td>
<td>Word DID NOT appear in text</td>
<td>Not sure</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>85. lavó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>86. carnet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>87. champú</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>88. caminó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>89. logre mos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90. abre</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>91. mono</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>92. traducirá</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>93. mire</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>94. comprará</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95. me levanto</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>96. peaje</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>97. chiste</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>98. pensamiento</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99. carmesí</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100. duda</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>101. generen</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>102. se queda</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>103. parpadearán</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>104. regrese</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>105. mimamos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word</td>
<td>Word DID appear in text</td>
<td>Word DID NOT appear in text</td>
<td>Not sure</td>
</tr>
<tr>
<td>------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>106. inglés</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>107. camino</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>108. bebé</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>109. pondrán</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110. bailó</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>111. manú</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>112. reloj</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>113. espicháramos</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>114. luciérnaga</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115. maneje</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>116. pregunto</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>117. nado</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>118. portada</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>119. navego</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120. charla</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix F: Comprehension Posttest

Version

Juanito: Grammatical target, Lexical control

Laura: Grammatical control, Lexical target

Instructions: Choose the option that best answers the question, based on what you just read. Once you have turned a page, do not flip back to see your previous answers. There are 22 questions for each text (44 total questions). Please write all answers on your answer sheet; do not write in this test booklet.

Juanito

1. Which change best describes the progression of Juanito’s personality throughout the story?
   a. from extroverted to introverted
   b. from introverted to extroverted
   c. from mature to immature
   d. from immature to mature

2. Which title best fits the reading?
   a. A Crazy Night Out
   b. Juanito’s Failed Attempt to Have Fun
   c. Juanito’s Long Day at the University
   d. none of the above
3. What does Juan always wear with his black shoes?
   a. gym clothes
   b. a sweater
   c. suspenders
   d. none of the above

4. Besides lacking self esteem, what else does Juan not have?
   a. a girlfriend
   b. a social network account
   c. nice clothes
   d. none of the above

5. Where do Juan’s roommates go every day?
   a. the gym
   b. the library
   c. class
   d. none of the above

6. What does Juanito NOT do on the night the story takes place? He does NOT…
   a. stay in
   b. eat dinner
   c. go out
   d. take a shower

7. What does Juanito do to his clothes?
   a. he irons them
   b. he folds them
c. he washes them

d. none of the above

8. Where does Juanito get carded?

   a. at the table
   b. standing at the bar
   c. at the liquor store
   d. none of the above

9. What does the employee tell Juanito he needs to enter the club?

   a. a pre-purchased ticket
   b. nicer clothes
   c. dress shoes
   d. none of the above

10. What does Juanito do with the classmate he sees?

    a. he shakes his hand
    b. he avoids him
    c. he chats with him
    d. none of the above

11. What does Juanito do with the drinks he buys?

    a. drink them
    b. share them
    c. pretend to drink them
    d. throw them out

12. What happens with Juan and the song “Vogue”? 
a. he hears it
b. he requests it
c. he sings it
d. none of the above

13. What does Juanito say about his night?
   a. “This is boring.”
   b. “There aren’t any pretty girls here”
   c. “This is the worst night ever.”
   d. none of the above

14. What does Juanito do with three pretty girls?
   a. kiss them
   b. buy them drinks
   c. talk to them
   d. dance with them

15. What does Juanito give to the prettiest girl?
   a. his phone number
   b. his drink
   c. a kiss
   d. none of the above

16. What does Juanito do at 4:00 AM?
   a. returns home
   b. goes to another bar
   c. passes out
d. eats something

17. What does Juanito do when he gets home that night?
   a. talks to his roommates
   b. studies
   c. goes to sleep
   d. call his mom

18. What does Juanito see in the morning that reminds him that he has to go to class?
   a. his email
   b. his backpack
   c. his textbooks
   d. none of the above

19. How does Juanito get to the university the next morning?
   a. he gets a ride with a friend
   b. he rides the bus
   c. he drives
   d. he walks

20. What does Juanito do with the calendar?
   a. throws it away
   b. disregards it
   c. looks at it
   d. loses it

21. What does Juanito realize that today is?
   a. his birthday
b. the first day of finals

c. Monday morning

d. none of the above

22. What does Juanito say he needs after finding out that there is no class?

   a. aspirin

   b. a study session

   c. a beer

   d. none of the above

Laura

23. Which best describes the rich lady’s case?

   a. difficult to solve but with a happy ending

   b. difficult to solve and with a sad ending

   c. easy to solve with a happy ending

   d. easy to solve but with a sad ending

24. Which title best fits the story about Laura?

   a. Laura’s Most Challenging Case

   b. A Comical Mystery

   c. Laura’s Fist Day on the Job

   d. A Snobby Client Ruins Laura’s Day

25. ‘Investigaciones Confidenciales Jiménez’ is the name of a _________.

   a. book

   b. school
c. police department

d. none of the above

26. Where does Laura keep papers in her office?
   a. in her desk drawer
   b. in a box
   c. all over her desk
   d. none of the above

27. Laura’s office is located in the most prestigious __________.
   a. neighborhood
   b. zip code
   c. building.
   d. none of the above.

28. What does Laura do at six o’clock in the morning?
   a. aerobic exercise
   b. drink coffee
   c. eat breakfast
   d. none of the above

29. What does she do after cleaning her office?
   a. surf the Internet
   b. fix the Internet
   c. turn on the Internet connection
   d. complain about the slowness of the Internet

30. What does the rich lady do at three o’clock in the afternoon?
a. arrives to Laura’s office
b. calls Laura’s office
c. leaves Laura’s office
d. none of the above

31. What does the rich lady say about ‘Rodríguez Garzón’?
   a. it’s her son’s name
   b. it’s her school’s name
   c. it’s her mother’s name
   d. none of the above

32. Who does the rich lady say is very prestigious?
   a. her children
   b. her lawyer
   c. her boss
   d. none of the above

33. After becoming interested in the case, what does Laura do?
   a. think about the case
   b. visit the crime scene
   c. interview the witnesses
   d. call the police

34. What does the rich lady tell Laura that she has?
   a. an unsolvable mystery
   b. connections to the mafia
   c. influence in the government
d. none of the above

35. What does the rich lady say that her baby does in the morning?
   a. wakes up late
   b. wakes up early
   c. cries
   d. eat breakfast

36. What does the rich lady say that her baby does in the afternoon?
   a. swims in the pool
   b. takes a nap
   c. plays with friends
   d. reads a magazine

37. What does Laura do after writing notes in her notebook?
   a. asks the rich lady a question
   b. checks her past cases
   c. makes some phone calls
   d. none of the above

38. Where does the baby sleep?
   a. with the rich lady
   b. on the sofa
   c. in the car
   d. none of the above

39. What does the rich lady do because she is anxious and tired?
   a. cries
b. leaves

c. yells
d. not talk

40. How does the rich lady drink her tea?
   a. out of a thermos
   b. out of a straw
   c. very quickly
   d. none of the above

41. What does the rich lady do after thanking the secretary for the drink?
   a. leaves
   b. talks on the phone
   c. asks for more
   d. rests

42. Where was the rich lady’s dog hiding?
   a. under the table
   b. in the car
   c. in the dog house
   d. none of the above

43. What does Laura do as the rich lady goes to the door?
   a. says goodbye
   b. walks with her
   c. asks her again for the check
   d. pets the dog
44. Laura thinks that this is the funniest __________ in all her years as a detective.

a. dog
b. woman
c. joke
d. none of the above

Version

Juanito: Grammatical control, Lexical target

Laura: Grammatical target, Lexical control

Instructions: Choose the option that best answers the question, based on what you just read. Once you have turned a page, do not flip back to see your previous answers.

There are 22 questions for each text (44 total questions). Please write all answers on your answer sheet; do not write in this test booklet.

Juanito

1. Which change best describes the progression of Juanito’s personality throughout the story?

   a. From extroverted to introverted
   
   b. From introverted to extroverted
   
   c. From mature to immature
d. From immature to mature

2. Which title best fits the reading?
   a. A Crazy Night Out
   b. Juanito’s Failed Attempt to Have Fun
   c. Juanito’s Long Day at the University
   d. none of the above

3. What does Juan always wear with his black shoes?
   a. gym clothes
   b. a sweater
   c. suspenders
   d. none of the above

4. Besides lacking self esteem, what else does Juan not have?
   a. a girlfriend
   b. a social network account
   c. nice clothes
   d. none of the above

5. Where do Juan’s roommates go every day?
   a. the gym
   b. the library
   c. class
   d. none of the above

6. What does Juanito NOT do on the night the story takes place? He does NOT…
   a. stay in
b. eat dinner
c. go out
d. take a shower

7. What does Juanito do to his clothes?
   a. he irons them
   b. he folds them
   c. he washes them
   d. none of the above

8. Where does Juanito get carded?
   a. at the table
   b. standing at the bar
   c. at the liquor store
   d. none of the above

9. What does the employee tell Juanito he needs to enter the club?
   a. a pre-purchased ticket
   b. nicer clothes
   c. dress shoes
   d. none of the above

10. What does Juanito do with the classmate he sees?
    a. he shakes his hand
    b. he avoids him
    c. he chats with him
    d. none of the above
11. What does Juanito do with the drinks he buys?
   a. drink them
   b. share them
   c. pretend to drink them
   d. throw them out

12. What happens with Juan and the song “Vogue”?
   a. he hears it
   b. he requests it
   c. he sings it
   d. none of the above

13. What does Juanito say about the song ‘Vogue’?
   a. “I hate this”
   b. “I love Madonna”
   c. “‘Vogue’ again??”
   d. none of the above

14. What does Juanito do with three pretty girls?
   a. kiss them
   b. buy them drinks
   c. talk to them
   d. dance with them

15. What does Juanito give to the prettiest girl?
   a. his phone number
   b. his drink
c. a love letter
d. none of the above

16. What does Juanito do at 4:00 AM?
   a. returns home
   b. goes to another bar
   c. passes out
   d. eats something

17. What does Juanito do when he gets home that night?
   a. talks to his roommates
   b. studies
   c. goes to sleep
   d. calls his mom

18. What does Juanito see in the morning that reminds him that he has to go to class?
   a. his email
   b. his backpack
   c. his textbooks
   d. none of the above

19. How does Juanito get to the university the next morning?
   a. he gets a ride with a friend
   b. he rides the bus
   c. he drives
   d. he walks

20. What does Juanito do with the calendar?
a. throws it away
b. disregards it
c. looks at it
d. loses it

21. What does Juanito realize that today is?
   a. his birthday
   b. the first day of finals
   c. Monday morning
   d. none of the above

22. What does Juanito say he needs after finding out that there is no class?
   a. aspirin
   b. a study session
   c. a beer
   d. none of the above

Laura

23. Which best describes the rich lady’s case?
   a. difficult to solve but with a happy ending
   b. difficult to solve and with a sad ending
   c. easy to solve with a happy ending
   d. easy to solve but with a sad ending

24. Which title best fits the story about Laura?
   a. Laura’s Most Challenging Case
b. A Comical Mystery

c. Laura’s Fist Day on the Job

d. A Snobby Client Ruins Laura’s Day

25. ‘Investigaciones Confidenciales Jiménez’ is the name of a __________.

   a. book
   
   b. school
   
   c. police department
   
   d. none of the above

26. Being as organized as she is, what does Laura keep in a workbook/notebook?

   a. her financial records
   
   b. her schedule
   
   c. her client forms
   
   d. none of the above

27. Where is Laura’s office located?

   a. in the most prestigious neighborhood
   
   b. in the most prestigious zip code
   
   c. on the most prestigious street
   
   d. none of the above

28. What does Laura do at six o’clock in the morning?

   a. aerobic exercise
   
   b. drink coffee
   
   c. eat breakfast
   
   d. none of the above
29. What does she do after cleaning her office?
   a. surf the Internet
   b. fix the Internet
   c. turn on the Internet connection
   d. complain about the slowness of the Internet

30. What does the rich lady do at three o’clock in the afternoon?
   a. arrives to Laura’s office
   b. calls Laura’s office
   c. leaves Laura’s office
   d. none of the above

31. What does the rich lady say about ‘María Garzón’?
   a. it’s her baby’s name
   b. it’s her great-grandmother’s name
   c. it’s her alias
   d. none of the above

32. Who does the rich lady say is very prestigious?
   a. her children
   b. her lawyer
   c. her boss
   d. none of the above

33. What does the rich lady tell Laura that she has?
   a. an unsolvable mystery
   b. connections to the mafia
c. influence in the government

d. none of the above

34. After becoming interested in the case, what does Laura do?

a. think about the case
b. visit the crime scene
c. interview the witnesses
d. call the police

35. What does the rich lady say that her baby does in the morning?

a. wakes up late
b. wakes up early
c. cries
d. eats breakfast

36. What does the rich lady say that her baby does in the afternoon?

a. swims in the pool
b. takes a nap
c. plays with friends
d. read a magazine

37. What does Laura do after writing notes in her notebook?

a. asks the rich lady a question
b. checks her past cases
c. makes some phone calls
d. none of the above

38. Where does the baby sleep?
39. What does the rich lady do because she is anxious and tired?
   a. cries
   b. leaves
   c. yells
   d. not talk

40. What does the secretary give the rich lady to calm her nerves?
   a. a hard drink
   b. a cup of coffee
   c. a sedative
   d. none of the above

41. What does the rich lady do after thanking the secretary for the drink?
   a. leaves
   b. talks on the phone
   c. asks for more
   d. rests

42. What does the rich lady open to be able to write out the check?
   a. her backpack
   b. her suitcase
   c. her briefcase
43. What does Laura do as the rich lady goes to the door?
   a. says goodbye
   b. walks with her
   c. asks her again for the check
   d. pets the dog

44. Laura thinks that this is the funniest ___________ in all her years as a detective.
   a. dog
   b. woman
   c. joke
   d. none of the above
Appendix G: Comprehension Posttest Answer Sheet

Participant Number:_________ Initials: _______

Start Time: _____

Instructions: Read each question from the comprehension packet and choose the option that best answers the question, based on what you just read. Please write all answers on this answer sheet; do not write in the test booklet. Once you have turned a page, do not flip back to see your previous answers. There are 22 questions for each text (44 total questions).

1. _____  
2. _____  
3. _____  
4. _____  
5. _____  
6. _____  
7. _____  
8. _____  
9. _____  
10. _____  
11. _____  
12. _____  
13. _____  
14. _____  
15. _____  
16. _____  
17. _____  
18. _____  
19. _____  
20. ____  
21. ____  
22. ____  
23. _____  
24. _____  
25. _____  
26. __  
27. ___  
28. ___  
29. ___  
30. ____  
31. ____  
32. ____  
33. ____  
34. ____  
35. ____  
36. ____  
37. ____  
38. ____  
39. ____  
40. ____  
41. ____  
42. ____  
43. ____  
44. ____  

End time _____
Appendix H: Debriefing Questionnaire

Participant # ____ / Initials: _____

You have just participated in a study of a grammatical form and several vocabulary words in Spanish. The target grammatical form was the preterit of 3rd person –ar verbs.

For example:
Ayer Juanito regresó a las tres de la tarde.
“Yesterday Juanito returned at three in the afternoon.”

Please indicate your knowledge/recognition of this form before participating in this study by marking the appropriate answer below:

_____ Yes, I knew/recognized this form before doing the exercises.

OR

_____ No, I did not know/recognize this form before doing the exercises.

There were also 20 vocabulary words. Complete the table below to indicate your knowledge of these words before reading the passage and the pretests associated with this experiment.

<table>
<thead>
<tr>
<th>Word</th>
<th>I had never seen this word before reading the passage.</th>
<th>I had seen this word before reading the passage but I didn’t know what it meant.</th>
<th>I had seen this word before reading the passage and I knew what it meant.</th>
</tr>
</thead>
<tbody>
<tr>
<td>traje</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mascota</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mesón</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>entrada</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>carnet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>canción</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>abrazo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reloj</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word</td>
<td>I had never seen this word before reading the passage.</td>
<td>I had seen this word before reading the passage but I didn’t know what it meant.</td>
<td>I had seen this word before reading the passage and I knew what it meant.</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td>feriado</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>holiday</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cabezada</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nap</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>negocio</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>business</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fichero</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>file</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cabinet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>paseo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>street</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>apellido</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>last</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>name</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>duda</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>doubt</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cuna</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>crib</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>taza</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mug</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cartera</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>purse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>relato</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tale</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Thank you for your participation!
Appendix I: Practice Think-aloud Protocol Task

Think aloud while you complete the following task of calculating the total cost of items on a grocery list. In other words, verbalize aloud or say whatever comes to your mind while you perform this task. You don’t need to explain what you are saying; just talk aloud as you do the task.

You want to buy

4 apples at 25 cents each =
4 lbs. of onions at $1 per pound =
6 cans of soda at 25 cents each =
1 gallons of milk at $3 each =

Total ___________
Appendix J: Practice Eye-tracking Task

Instructions: The following text comprises a short paragraph. Skim, that is, read very quickly, the text for meaning only. **YOU HAVE A TIME LIMIT OF ONLY 6 SECONDS PER SENTENCE!** Each sentence is presented on a separate screen. There are 5 sentences in total. When you are finished with one screen, press OK to advance to the next screen. You cannot backtrack.

Juan es estudiante de medicina. Está muy estresado. No duerme la noche antes de su examen de anatomía. Al fin duerme a las seis de mañana. Saca mala nota en su examen y está muy triste.
References


Calderón, A. (2013). The effects of L2 learner proficiency on depth of processing, levels of


De Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of
word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning, 50*(1), 1-56.


F. I. M. Craik (Eds.), The Oxford handbook of memory (pp. 229-244). New York: Oxford University Press.


Godfroid, A., Boers, F., & Housen, A. (2013). An eye for words: gauging the role of attention in


Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,*


Acquisition, 29(1), 39-66.


Smith, B. (2010). Employing eye-tracking technology in researching the effectiveness of recasts in CMC. In F. M. Hult (Ed.), *Directions and prospects for educational linguistics* (pp. 79-97). New York: Springer.


Theeuwes, J., & Van der Stigchel, S. (2009). Saccade trajectory deviations and inhibition-of-


