

ATTENTION TO IRREGULAR VERBS BY BEGINNING LEARNERS OF GERMAN:
AN EYE-MOVEMENT STUDY

By

Maren Schierloh

A DISSERTATION

Submitted to
Michigan State University
In partial fulfillment of the requirements
For the degree of

DOCTOR OF PHILOSOPHY

Second Language Studies

2011

ABSTRACT

ATTENTION TO IRREGULAR VERBS BY BEGINNING LEARNERS OF GERMAN: AN EYE-MOVEMENT STUDY

By

Maren Schierloh

In this study I examine the noticing of verbs with stem vowel changes by beginning adult learners of L2 German who have not been formally introduced to this linguistic feature. Two research questions (RQs) guided the experimental design and the empirical analyses in this study: (1) Do adult beginning learners of German who are unfamiliar with stem-changing verbs attend to those irregularities during reading? (2) Is increased attention to irregular verbs associated with subsequent learning of them? Regarding RQ1, I hypothesized that during the reading, learners would fixate on vowel-stem-changing verbs significantly longer than regular verbs in which the stem does not change, with fixations measured through eye-tracking technology. Regarding RQ2, I hypothesized that variations in eye fixation durations would not be predictive of the learning of stem-changes in that there would be no statistically significant correlation between pre- and posttest improvements and fixation times. I predicted that learners would improve on the inflections and overall spelling of verbs due to the exposure to the verbs during the reading activity.

The participants, 43 beginning learners enrolled in a first semester German course, took a picture-based sentence production test to estimate their initial command of stem-changing verbs. Then, they read sentences with stem-changing verbs (experimental condition) and regular verbs (control condition) on a computer screen. The sentences appeared in pairs of two, one baseline sentence (no-stem change, regardless of condition) and one critical sentence (stem-change in the

experiment condition). The sentence endings as well as the person (2nd and 3rd person singular) were counterbalanced across 24 trials, which were also randomized for each participant. As they read, participants' eye movements were recorded. Last, participants took a post-test, which was identical to the pre-test with the difference being the order in which the stimuli-pictures were presented.

I used multivariate ANOVA modeling to investigate whether the novel verb irregularities (i.e., vowel stem changes) had an effect on learners' eye fixations (total time, first fixation duration, and gaze duration). Statistical and visual inspection of the data revealed outlying observations, which were cut from the regular data set (truncated) and analyzed separately via Chi-square tests and logistic regressions. Regarding total time, there was a significant effect of verb type, which was also evident in the analysis of the outlier data. For first fixation, statistical significance was only found with the outlier observations, revealing a trend similar to the findings with total time. The findings regarding gaze duration were also not statistically significant. The outliers evidenced the same but weaker pattern. Overall, there was a notable trend showing that irregular verbs were fixated on longer, which was statistically significant in half of all statistical analyses. The correlation analyses between total time gains (i.e., the fixation time critical condition subtracted from the baseline condition) and learning gains revealed that noticing (longer fixations on irregular verbs) was related to holistic learning gains (not necessarily associated with stem-changes).

I interpret the results from multiple perspectives with reference to previous research on noticing and language learning. The results provide for a more nuanced yet more empirically-based understanding of the noticing construct.

To My Parents

ACKNOWLEDGEMENTS

A great many people have been indispensable—not only in the pursuit of this dissertation project, but also throughout my graduate career. The conduct of this dissertation project would not have been possible without the help of my dissertation guidance committee (Drs. Aline Godfroid, Paula Winke, Susan Gass, and Debra Hardison), Dr. Marcus Johnson from SR Research, the volunteers at the Department of Statistics and Probability, Shelee-Ann Flemmings, the German instructors and students, my colleagues, my friends, and my family. It gives me great pleasure to take this opportunity to express my heart-felt gratitude to all.

Aline Godfroid has been an irreplaceable source of inspiration and encouragement; it was her enthusiasm and interest in this project that motivated me to follow through. Aline's exceptional expertise in experimental design, eye-tracking research and statistics, her thorough feedback, in-depth thinking and eye for detail have helped me tremendously. I feel fortunate being her student, colleague, and friend—a privilege I will always treasure.

In uncountable ways, I am indebted to my advisor, Paula Winke, who has become a life-long role model for me. I thank her for motivating me, and for increasing my awareness for seeing the bigger picture, not only in respect to my dissertation. I also owe much gratitude to Susan Gass, whose genuine care on so many levels gave me confidence and enthusiasm to power through. I am honored having been mentored by a Distinguished Professor of her caliber. Last but not least, I am grateful for Debra Hardison's service on my dissertation committee, but even more so, for the great impact she has had on my development as a university-level teacher.

The programming of the eye-tracking project would not have been possible without Marcus Johnson from SR research, whose detailed instructions and many screen-shots helped me

navigate the world of computer programming, a terrain I never thought I would understand. My sincere thanks are also due to Vineet Bansal whose Excel-script-writing skills saved me many hours of cut-and-paste, and, more importantly, from human error. Similarly, I would like to extend my gratitude to the volunteers at the Department of Statistics and Probability, who kindly welcomed my frequent inquiries and showed great excitement for my data. I also would like to acknowledge and express my gratitude for the assistance of Shelee-Ann Flemmings during my data collection time.

While I feel grateful for being among many fine colleagues in the MSU Second Language Studies Program, I have been incredibly blessed with the friendship that grew from my teamwork with Luke Plonsky. Everyone needs a friend like Luke.

I cannot imagine having arrived at this point in my life or academic career without the love and support of my family. My future mother and father in law have given me great support throughout; it is overwhelming to know that people I admire so much are so proud of me. My parents in Germany, Rainer and Ulrike Schierloh, are the ones to whom this dissertation is dedicated. Thank you for instilling in me the values of perseverance, discipline, and love for learning. Thank you for supporting the pursuit of my dreams, and for always being there for me.

And finally, I am thankful beyond measure to my future husband, Dr. Christopher Uggen, for his love and care, encouragement and support, and his genuine interest in my studies - while also reminding me that there is more to life than ANOVAs and Logistic Regressions! Thank you for bringing home Apollo, our dear yellow-lab puppy, whose everlasting happiness and faithful company I cherish every day, especially during long hours of writing. Thank you, Chris, for believing in me—and for encouraging me to never do any less than my very, very best.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xi
INTRODUCTION	1
CHAPTER 1 TAKING A COGNITIVE PERSPECTIVE ON SLA	4
Attention and Awareness	4
Schmidt's Noticing Hypothesis.	10
Tomlin and Villa's Theory of Attention	10
VanPatten's Input Processing Theory	11
Robinson's (2003) Theorizing of Attention.....	14
Implicit versus Explicit Learning.....	25
Intentional versus Incidental Learning.....	29
CHAPTER 2 NOTICING	33
Noticing and Learning	37
The Measurement of Noticing	40
Eye-Tracking as a Measurement of Noticing	44
Word Frequency.....	45
Word Predictability	46
Word Length	47
The E-Z Reader Model	49
CHAPTER 3 GERMAN STEM-CHANGING VERBS	54
L2 Morphology and Its Noticability	54
German Stem Changing Verbs	55
The Difficulty and Noticability of L2 Forms.....	57
CHAPTER 4 METHODOLOGY	61
Research Questions and Hypotheses	61
Materials	64
The Pre- and Post-test	64
The Reading Task	65
The Vocabulary Familiarity Test.....	70
Procedures.....	70
Analyses	73
Analysis of the Eye-tracking Data	74
Analysis of the Test Scores.....	76
CHAPTER 5 RESULTS	78

Results Addressing RQ 1	79
Total Time.....	79
Total time ANOVA.....	81
Post Hoc Tests.....	82
Pair-wise Comparisons	83
Nested ANOVA Assumptions	84
Homogeneity of Residuals Variances.....	84
Normality of the Residuals	85
Analysis of Omitted Outliers	85
First Fixation Duration.....	93
First Fixation Duration ANOVA	93
Nested ANOVA Assumptions	94
Analysis of Outliers	94
Gaze Duration.....	99
Gaze Duration ANOVA.....	100
Assumptions.....	101
Analysis of Outliers	101
Summary of Findings for Research Question 1.....	105
Results Addressing RQ 2	106
Correlation of Fixation Times with Learning Gains.....	110
Partial Correlation of Fixation Times with Learning Gains	111
Summary of Findings for RQ 2	117
CHAPTER 6 DISCUSSION.....	118
Research Question One.....	118
Attention to Irregular Verbs.....	119
“Outlier Cases” with Extreme Fixation Times	121
Research Question Two	124
Non-saliency of Irregular Verb Morphology.....	125
The Developmental Level of the Learner and Attention	131
Rule-based versus Item-based Learning.....	132
Parabolic Relationship	134
The Verbs “Sehen” and “Fahren”	136
Limitations and Directions for Future Research.....	137
Limitations	137
Suggestions for Future Research	142
CHAPTER 7 CONCLUSION.....	148
Pedagogical Suggestions.....	149
The Contributions of This Study to SLA Theory	150
APPENDICES	151

REFERENCES	159
------------------	-----

LIST OF TABLES

Table 1: Descriptive Statistics.....	64
Table 2: Counterbalancing of Sentence Endings and Persons.....	69
Table 3: Variables in the Interest Area Report	74
Table 4: Logistic Regression with Total Time.....	91
Table 5: Logistic Regression with First Fixation Duration.....	99
Table 6: Logistic Regression with Gaze Duration	105
Table 7: Verbs with Stem Changes from e → i/ie	153
Table 8: Verbs with Stem Changes from a → ä	153
Table 9: Sentence Stimuli by Group	154
Table 10: Sentence Stimuli by Group (a).....	155
Table 11: Sentence Stimuli by Group (b)	156
Table 12: Sentence Stimuli by Group (c).....	157
Table 13: Frequency of the Verbs Used in this Study	158

LIST OF FIGURES

Figure 1. Partial Display of Gass's (1997) Model of Apperception	8
Figure 2. A Generic Model of Human Information Processing with Three Memory Systems ...	15
Figure 3. Wickens's (2007) Model of Information Processing	17
Figure 4. Wickens's Proposed Structure of Processing Resources (2007, p. 186)	19
Figure 5. Continuum with Different Levels of Awareness	22
Figure 6. Graphic Representation of the Noticing Hypothesis	33
Figure 7. Simple Illustration of the Experimental Design	66
Figure 8. Example of Eye-tracker Screens.....	68
Figure 9. Nested Design.....	79
Figure 10. Post-Hoc Results.....	83
Figure 11. Verb Type and Total Time Gain (Regular versus Irregular)	87
Figure 12. Verb Type and Total Time Gain (by Stem-change)	89
Figure 13. Verb Type and First Fixation (Regular versus Irregular)	97
Figure 14. Verb Type and First Fixation (by Stem-change)	98
Figure 15. Verb Type and Gaze Duration (Regular versus Irregular)	103
Figure 16. Verb Type and Gaze Duration (by Stem-change)	104
Figure 17. Design of the Study	107
Figure 19. Illustration of a Semi Partial Correlation (Controlling for Reading Speed).....	110
Figure 20. Simplistic Depiction of the Parabolic Relationship.....	114
Figure 21. Verb and Mean Scores Stem-change.....	115

Figure 22. Verb and Mean Scores Holistic	116
---	------------

INTRODUCTION

This Ph.D. dissertation reports on a study investigating whether adult beginning learners of German (mostly native speakers of English) attend to or notice irregular, stem-changing verbs, given that they do not know the feature of verb stem-changes in German. I was interested not only in the learners' attention to irregularities, but also whether increased attention to these features is associated with improvements in learners' production of stem-changing verbs. To measure attention, I used eye-tracking technology, a measurement tool recently utilized to investigate the noticing construct in second language acquisition (SLA).

My interest in language learners' noticing of novel linguistic features and its relationship to learning is grounded in the ongoing and intense debate in SLA centering on the following question: Does incidental attention to a particular language form promote the acquisition of that form? (e.g., Doughty & Williams, 1998; Norris & Ortega, 2000; Skehan & Foster, 2001). This question, and the role of noticing in particular, have played crucial roles in a vast number of SLA studies. As supported by Robinson (1995; 1996; 2003; 2008), Schmidt (1995; 2001; 2010), Skehan (1998) and other scholars in the field (e.g., Adams, 2003; Carr & Curran, 1994; Carrell, 1992; DeKeyser, 1997; R. Ellis, 1994; Fotos, 1993, 1994; Gass & Alvarez-Torres, 2005; Hama & Leow, 2010; Izumi, 2002; Izumi & Bigelow, 2000; Izumi, Bigelow, Fujiwara, & Fearnow, 1999; Izumi & Izumi, 2004; Leow et al., 2008; Leow & Morgan-Short, 2004; Leow, 1997b, 1998, 2000; Long, 1996; Mackey, 2006; Mackey, Gass, & McDonough, 2000; Norris & Ortega, 2000; Roberts, 1995; Rosa & Leow, 2004; Rosa & O'Neill, 1999; Sachs & Suh, 2007; Swain, 1995, 1998, 2005; Swain & Lapkin, 1995; Van Hell & De Groot, 1999), successful language learning by adults seems to be tightly linked to learners' noticing of words, phrases, and grammar structures.

The notion that noticing is essential for the acquisition of new linguistic systems was introduced and developed by Schmidt (1990, 1993a, 1994, 1995, 2001), and it has since been subject to intense study (e.g., Leow, Hsieh, & Moreno, 2008; Leow, 1997, 2000; Robinson, 1995; 2003; 2008) and much scrutiny (e.g., Carroll, 2006a, 2006b; Schachter, 1998; Truscott, 1998; Williams, 2005). Much of the argumentation is grounded in the difficulty of operationalizing and measuring L2 learners' internal cognitive processes (e.g., Egi, 2004; Leow, 2000, 2002; Simard & Wong, 2001). Researchers in SLA and cognitive science have attempted to overcome this challenge by introducing and testing several hypotheses related to the type and amount of attention necessary for language learning (Godfroid, 2010; Godfroid, Housen & Boers, 2010; Simard & Wong, 2001; Tomlin & Villa, 1994), and the specific aspects of language that are most likely to be noticed (Schierloh, in press; Carroll, 2006a, 2006b; N. Ellis & Sagarra, 2010).

In this dissertation, I operationalize noticing as *focal attention* with some awareness (Schmidt, 1995, 2001; Robinson, 2003). I use eye-tracking methodology to monitor where and for how long beginning learners look when they read simple sentences in their second language (L2). As such, eye-tracking allowed me to tap into the learners' attentional/visual foci during reading and to investigate whether increased attention to new a morphological feature yields uptake or learning of that word or feature.

Research on the relationship between noticing and learning is of theoretical and practical importance. As several scholars have observed (see, for example, Godfroid, 2010), there is a need to refine the theoretical foundations of the noticing construct, and investigations into the extent to which eye-tracking methodology can provide an accurate, empirically-based measurement of the noticing construct is of utmost interest in this respect. On a practical level,

there is a need for researchers to investigate whether beginning learners pay attention to new linguistic features when they first come across them while reading. Such information can help, for example, instructors of foreign languages understand how their students process (or do not process) new linguistic information, and whether explicit instruction and interventions are needed to enhance students' attention to and learning of form.

This dissertation is organized in the following way: In the first chapter, relevant areas of the literature pertaining to three areas of SLA research are discussed: (a) attention and awareness; (b) implicit versus explicit learning; and (c) incidental versus intentional learning. The second chapter discusses the concept of noticing, the central issue of this dissertation. After discussing the *Noticing Hypothesis* and the three types of noticing, methodological issues are reviewed and the methodology of this study—eye-tracking—as well as this study's operationalization of noticing are discussed. In the third chapter I describe the linguistic characteristics of German stem-changing verbs, their underlying rules, and their acquisition by foreign language learners. The fourth chapter outlines the methodology of this study. The following chapter reports the empirical findings. In chapter six I discuss the findings in light of the research questions and address the observed relationships among attention and learning.¹ In the final chapter I summarize the findings and outline limitations and suggestions for future research. I conclude by providing pedagogical implications.

¹ In regard to this dissertation study, the term 'learning' corresponds to gains in performance from pre-test to post-test.

CHAPTER 1 TAKING A COGNITIVE PERSPECTIVE ON SLA

In this chapter I present literature on some of the main cognitive issues in SLA and how they have been conceptualized, researched, and debated. In particular, I discuss how the notions of attention and awareness have been theorized, discussed, and operationalized in contexts of SLA and cognitive psychology. In this discussion, I explain the concepts of implicit (i.e., unaware) and explicit (i.e., aware) learning, and the roles these concepts play in my dissertation research. Closely related to the issues surrounding implicit and explicit learning are the notions of incidental and intentional learning; those will be briefly discussed in the last sub-section of Chapter 1.

Attention and Awareness

In the SLA literature, Schmidt was the first scholar who conceptualized and discussed different levels of awareness/consciousness in relation to SLA. Schmidt (1990) used the notions of ‘awareness’ and ‘consciousness’ interchangeably,² and in this paper I use the single term ‘awareness’ to maintain clarity. Although Schmidt did not suggest a dichotomy of attention and awareness, he suggested that there are different *levels* of conscious learning, ranging on the awareness scale from intentional learning—with the highest level of awareness involved—to implicit learning, and to subliminal learning—with the least amount of awareness involved. With this conceptualization, Schmidt is also the founder of the well-known *Noticing Hypothesis*, which will be discussed in Chapter 2. His differentiation of the umbrella term of consciousness or awareness (in this paper termed awareness) and what it means for the study of SLA is

² In Schmidt (1990) awareness is one of the three senses in which the term consciousness is commonly used when discussing the possibility of unconscious learning.

important because it discusses the involvement of different levels of consciousness in the L2 learning process.

Schmidt's investigations into awareness originated as a reaction to Krashen (1981, 1985) and Krashen and Terrell (1983) who hypothesized that adults acquire an L2 unconsciously, just as children learn their first language (L1). Schmidt (1990) argued that Krashen over-simplified the role of consciousness in the acquisition process. According to Schmidt (1990), consciousness plays an essential role in adult L2 acquisition, and there are three senses of consciousness: (1) consciousness as awareness; (2) consciousness as intention; and (3) consciousness as knowledge or understanding (see also Godfroid, 2010). When referring to the first sense, consciousness as awareness, he further distinguished three levels.

1. Perception
2. Noticing (i.e., here: awareness and attention)
3. Understanding

The crucial part of this conceptualization is that noticing does not *require* awareness; it is a *type* or a *quality* of awareness (sometimes also referred to by Schmidt as 'awareness at the level of noticing'). In 2001, Schmidt proposed that a low level of awareness—at the level of noticing (i.e., focal attention)—is necessary for input to become intake for learning. The details and impact of this *Noticing Hypothesis* are discussed in Chapter 2.

Schmidt's (2001) argument that 'perception' is below the level of conscientious noticing has been criticized because it is impossible to investigate the absolute absence of awareness in learners (cf. Hama & Leow, 2010) who cannot report this type of perception verbally (Godfroid, 2010). If language learners merely *perceive* input, without awareness, it is highly unlikely that they will learn from it (Sharwood-Smith & Truscott, 2010; Sharwood-Smith & Truscott, in

press; Truscott, 1998). The question is whether there are categorical distinctions on the awareness scale, and, if so, where do these boundaries, or thresholds, lie in our consciousness that spans from perception (i.e., non-conscious registration) to noticing (i.e., conscious perception or perception that involves some level of awareness)³ to understanding (i.e., full awareness of the learning process and of what is being learned). Sharwood-Smith and Truscott (in press) similarly argued, “the concept of noticing can only be understood if there is a principled means of drawing these boundaries. But no such means have yet been offered” (p. 8). In the field of cognitive science, however, researchers have been occupied with very similar, if not the same, question of how we can determine whether a participant has *consciously* perceived certain visual stimuli (e.g., Dehaene & Changeux, 2004; Dehaene et al., 2001). This marks the difference between subliminal processing, on the one hand, and conscious perception on the other.

In referring to *research* on the involvement of awareness in learning, the cognitive psychologists Merikle, Smilek, and Eastwood (2001) distinguished two different threshold areas within research on unconscious perception: (a) a *subjective threshold*, which is a participant’s failure to report conscious awareness of a stimulus, and (b) an *objective threshold*, which is a

³ In the cognitive science literature, there is agreement that (visual) perception involves awareness, simply because no study has shown clear evidence of zero consciousness/awareness given the difficulty or impossibility of supporting the null hypothesis. It is important to note here that in the field of cognitive psychology, functional neuroimaging studies indicated that visual processing is possible without conscious awareness with blind sight patients (see, for example, Overgaard, Feh, Mouridsen, Bergholt, & Cleermans, 2008). Thus, we cannot state for certain that *all* perception involves awareness.

participant's inability to make accurate forced-choice decisions about a stimulus.⁴ According to Merikle et al. (2001), research participants often demonstrate awareness assessed by the objective threshold, even if there is a lack of evidence of awareness at the subjective threshold. This view corresponds to work by Lamme (2003), who maintains that “attention does not determine whether stimuli reach a conscious state but determines whether a (*conscious*) report about stimuli is possible” (p. 13, highlights are mine). Cognitive psychologists, unlike most SLA researchers, usually favor the reliance on an objective threshold for validating reports of attention compared to presenting possibly biased or inaccurate reports of conscious experience (at the subjective level).

Regardless of the unsolved question as to the involvement of awareness in foreign language processing and learning, SLA researchers agree that Schmidt's distinction of perception, noticing (i.e., focal attention in reading), and understanding relies on the extent to which awareness is involved. Thus, Schmidt made a crucial contribution to the field of SLA. Yet, even more valuable is his emphasis on the *essence* of attention, which he believes to involve some level of awareness, for language learning to occur.⁵ In 2001, he directly addressed the role of attention:

⁴ In an experiment this decision-making could be to guess with above chance level whether a word is a word or not, or, in reference to this dissertation study, whether a word is fixated and re-fixated (i.e., attended to) or not.

⁵ While in the SLA literature researchers widely believe that consciousness goes hand in hand with attention, it is important to acknowledge that recently psychological and neurophysiological researchers put forward that attention and consciousness are *distinct* (not isomorphic, as Schmidt suggested) because they have distinct neuronal mechanisms (B. J. Baars, 1997; B. J. Baars, 2002; Block, 2005; Dehaene, Changeux, Naccache, Sackur, & Sergent, 2006; Koch, 2004; Koch & Tsuchiya, 2007; Lamme, 2003; Naccache, Blandin, & Dehaene, 2002; Woodman & Luck, 2003).

“[...] the concept of attention is necessary in order to understand virtually every aspect of second language acquisition, including the development of interlanguages (IL) over time, variation within IL at particular points in time, the development of L2 fluency, the role of individual differences, such as motivation, aptitude, and learning strategies in L2 learning, and the ways in which negotiation for meaning, and all forms of instruction contribute to language learning” (p. 3).

While admitting that it is nearly impossible to truly separate attention and awareness (p. 4), Schmidt argued that his concept of noticing is equivalent to Gass’s ‘apperception’ of novel input that involves some level of awareness. According to Gass (1997), “apperception is an internal cognitive act in which a linguistic form is related to some bit of existing knowledge (or gap in knowledge)” (p. 4). Figure 1 demonstrates one part of Gass’s (1997) model that illustrates a conceptualization of apperception.

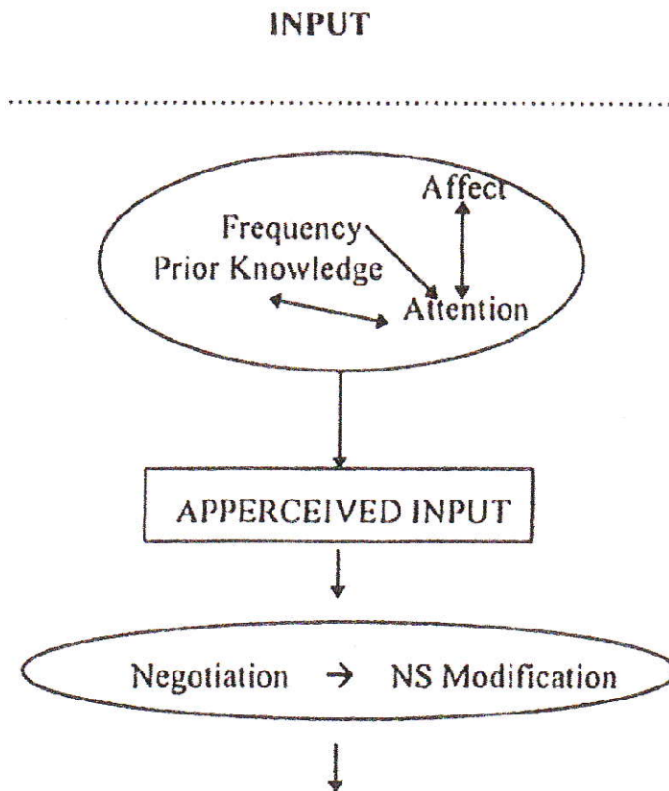


Figure 1. Partial Display of Gass’s (1997) Model of Apperception

According to this model, apperceived input is input that has been attended to with some *level of awareness*, perhaps because the input is recognized due to its frequency, the learner's prior knowledge or experience, or because there is something particular in the input that does not match the current representation of the language in the learner's mind. This concept of noticing as focal attention, or *apperception*, provides an important theoretical foundation for this dissertation study. The important point to keep in mind is that I do not equalize attended input (i.e., 'intake') with acquisition (not all attended input enters long-term memory), but I will use the concept of apperceived input to discuss some findings in this study.

It is important to acknowledge that, with few exceptions (e.g., Hama & Leow, 2010), SLA researchers have moved away from debating the role of awareness in the SLA process, mainly due to the impossibility of investigating the complete absence of awareness in language learners (e.g., Block, 1995; Godfroid, 2010; Rosa & Leow, 2004; Rosa & O'Neill, 1999; but cf. Williams, 2005). More recently, researchers have designed studies to investigate *how* attention and which *types* of attention function in adult language acquisition (e.g., Gass & Alvarez-Torres, 2005; Gass, Svetics, & Lemelin, 2003; Godfroid, Housen, & Boers, 2010b; Leow, 1997, 2000; Robinson, 2003), questions which are also central to this dissertation research. The literature in cognitive psychology largely refers to *selective processing* when referring to one of the main functions of attention and distinguishes two types of attention: *focused attention* and *divided attention*. Individuals employ *focused attention* when they only attend to selected parts of the stimuli input while other stimuli are ignored (therefore, it is also known as *selective attention*). Divided attention is when more than one stimulus or type of information is attended to simultaneously (Eysenck & Keane, 2010). Robinson (2008) used the same conceptual distinction, but he used different terminology. He called *divided attention* perceptual attention

and *focused attention* focal attention. According to Robinson, focal attention is “under some degree of voluntary executive control” and therefore a “precondition for awareness” (p. 133). I adopt the terminology by Robinson in this dissertation. With the distinction of *perceptual attention* and *focused attention* in mind, I will explain how the role of attention has been theorized in SLA. I will elaborate on attention in (1) Schmidt’s *Noticing Hypothesis*, (2) Tomlin and Villa’s Theory of Attention, (3) VanPatten’s Input Processing Theory, and (4) Robinson’s (2003) Theory of Attention.

*Schmidt’s Noticing Hypothesis*⁶

In drawing on work in cognitive psychology, Schmidt (2001) concluded that attention is selective and only partially under voluntary control. Attention controls the access to awareness and is essential for learning, but the amount of attention needed differs depending on the linguistic aspects (syntax, morphology, pragmatics, etc.) at hand. Although he never stated it explicitly, it can be assumed that Schmidt is theorizing on what Robinson (2003) calls *focal attention*.

Tomlin and Villa’s Theory of Attention

Tomlin and Villa’s view of the role of attention in SLA (1994) differed from the mainstream theories in SLA, which attracted opposition (e.g., Schmidt, 2001; Simard & Wong, 2001). They divided attention into three distinct processes: (1) alertness, which deals with the learner’s affective and motivational state, specifically the learner’s overall readiness to pay attention to the input features they encounter either through reading or listening; (2) orientation, which refers to the concentration of attention to a specific feature in the input at the expense of

⁶ Schmidt’s Noticing Hypothesis is the focus of Chapter 3, hence the succinct summary.

others; and (3) detection, the stage at which the cognitive processing of the focused input takes places. Detection is also the stage at which the specific language feature is registered in working memory. Tomlin and Villa claimed that the last component, detection, does not require the first two components to take place. That is, learners can detect a specific part of the incoming stimuli without alertness or orientation. They further claimed that *none* of the three processes of attention requires awareness, but attention is a prerequisite for awareness. These claims have received scrutiny but also interest from others. Most notably, Leow (1998) conducted a study in which he assigned language learners to four conditions, all of which were categorized as + awareness, but differed as to whether they were + or – orientation and + or – detection. According to Leow’s findings, detection, but not orientation, appeared to be a crucial attentional mechanism that allows L2 development to occur.

Simard and Wong (2001) criticized Leow’s research as well as Tomlin and Villa’s theorizing on the basis that alertness, orientation, and detection constitute one process in the mind—they not separable. They criticized Loew’s research because it did not measure what it intended to (i.e., alertness, orientation, and detection), and that a much more fine-grained analysis of attention is needed if one is to tease apart these attentional mechanisms. In addition to criticizing Leow (1998), Simard and Wong challenged Tomlin and Villa (1994) for their claim that alertness, orientation, and detection can take place without awareness. According to Simard and Wong, the involvement or non-involvement of awareness in attentional processes is an issue that has not been resolved to date, so strong claims such as those by Tomlin and Villa as well as Leow, should not be made. The role of awareness will resurface in my discussion of explicit and implicit learning.

VanPatten’s Input Processing Theory

VanPatten's Input Processing (IP) theory makes claims regarding the nature of attention, language processing, and the structure of attentional resources. VanPatten (1990, 1994) was interested in how learners allocate their attention when processing input and asked the following question: "What happens when learners are asked to attend to meaning and form together or just to meaning or form?" His research findings suggested that both dimensions of the linguistic input, meaning and form, compete for learners' attention. Because of limited working memory capacity, VanPatten claims that learners have to prioritize one over the other and they can only attend to form when they can easily understand the meaning of the input. Based on these findings, VanPatten (1996, pp. 24-26) formulated his IP Principles, which pedagogically underpin his theory:

- P1: Learners process input for meaning before they process it for form.
- P1a: Learners process content words in the input before anything else.
- P1b: Learners prefer processing lexical items to grammatical items (for example, morphological markings) for semantic information.
- P1c: Learners prefer processing 'more meaningful' morphology before 'less or non-meaningful' morphology
- P2: For learners to process form that is not meaningful, they must be able to process information or communicate content at no (or little) cost to attention.

Based on these principles, learners might only be able to pay attention to form if the overall meaning of the input is understood effortlessly. If this is not the case, learners focus their attention to the meaning, usually content words in the input, as their working memory capacity is limited. That is, meaning is primary, and form processing is secondary. I will turn to the issue of whether or not language learners can be described as this type of "limited capacity processors" in Chapter 3.

DeKeyser, Salaberry, Robinson, and Harrington (2002) criticized VanPatten's IP Theory in relation to three areas: (1) the validity of the *limited-capacity processing* concept, (2) the consideration of only the *single-resource* model of attention, and (3) the details of the mechanisms involved in L2 processing. Concerning the first two areas, DeKeyser et al. criticized the nature of VanPatten's research (VanPatten, 1994, 1996; VanPatten & Cadierno, 1993; VanPatten & Oikkenon, 1996) where VanPatten and his colleagues manipulate the linguistic input so learners are likely to trade off attention to form to attention to meaning,⁷ which is then explained within the framework of the what DeKeyser et al. consider an outdated *single-resource, limited-capacity* model of attention. In summary, DeKeyser et al. wrote that they believe *attention* in the IP model is too vaguely defined and not in accordance with existing cognitive theory.

In reference to VanPatten's *IP Theory*, information on bottom-up and top-down processing adds to the completeness of the theory. When learners process input, they naturally process the language from the top down (i.e., relying on contextual cues to aid in the comprehension of the information) as well as from the bottom up (i.e., relying on linguistic cues) (Bell, 2006; Kintsch, 2005; Kurby, Britt, & Magliano, 2005). Tsui and Fullilove (1998), for example, examined the processing skills used by skilled and less-skilled L2 readers. They observed that skilled readers are better able to engage in top-down processing. In light of attentional processes, skilled L2 readers—those who favor top-down processing—might pay more attention to linguistic cues, such as morphology and syntax in the L2. Yet, both top-down

⁷ DeKeyser et al. (2002) also find fault with the stimuli (target words) selected in VanPatten's study, so the results do not reflect "a form-meaning processing distinction but rather an easy-difficult scanning distinction" (p. 813).

and bottom-up processes jointly influence the mental representations of linguistic stimuli that are formed during or as a result of comprehension (Kintsch, 2005).⁸

Beginning learners who experience processing difficulties in their L2 are assumed to depend more on contextual cues that are present elsewhere in the sentence or come from other sources, such as pictures, the situational context, or the learners' background knowledge.

Robinson's (2003) Theorizing of Attention

In 2003, Robinson provided a very comprehensive synthesis of the work on attention in the field of SLA. He differentiated three types of attention:

- (1) Attention as selection (focused attention)
- (2) Attention as capacity
- (3) Attention as effort

Regarding the first type, *attention as selection*, which is relevant in light of the present study, Robinson relied on Wickens, Gordon, and Liu's (1997) information processing framework to elucidate where and how attention as selection takes place during information processing.

⁸ In the cognitive science literature, where attention is typically investigated with reference to visual perception rather than language processing, Corbetta and Shulman (2002) and Yantis (2008), for example, referred to top-down ("active") and bottom-up ("passive") modes of attention. Simply put, top-down attention is controlled by the individuals' objectives, while bottom-up attention is driven by external stimuli, or stimuli that are not part of the individual's immediate goals. Similarly, Ruz and Nobre (2008) studied how attention modulates initial stages of visual word processing (with native speakers). They found that selective attention to linguistic attributes (i.e., top-down processing) enhances the initial stages of visual word processing.

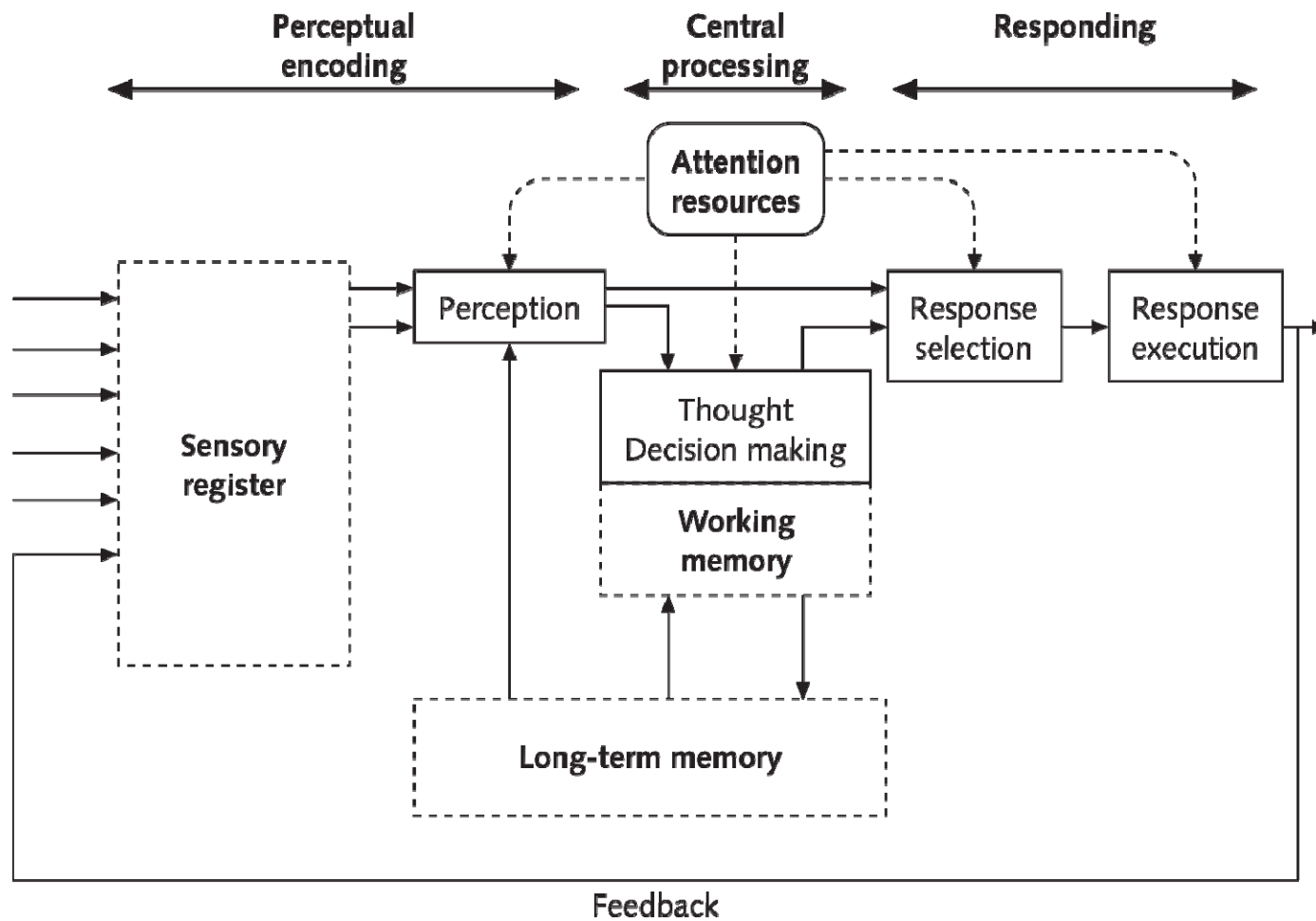


Figure 2. A Generic Model of Human Information Processing with Three Memory Systems

Attention is a major resource during central processing. It influences the perception and the selection of stimuli; that is, it controls what part of the perceived input enters into an individual's working memory. In Robinson's (2003) words, "focal attention and noticing are selective of input" (p. 637).

In 2007, Wickens proposed a modified model of information processing that elucidates the allocation of attentional resources in *every* aspect of information processing. Attentional resources are specifically allocated to the selection and the perception, the latter of which is also influenced by already established representations in long-term memory. In Wickens's 2007 model, attentional selection is hypothesized to happen early in this model. Working memory, which is an intermediary step to long-term memory, is also under attentional control.

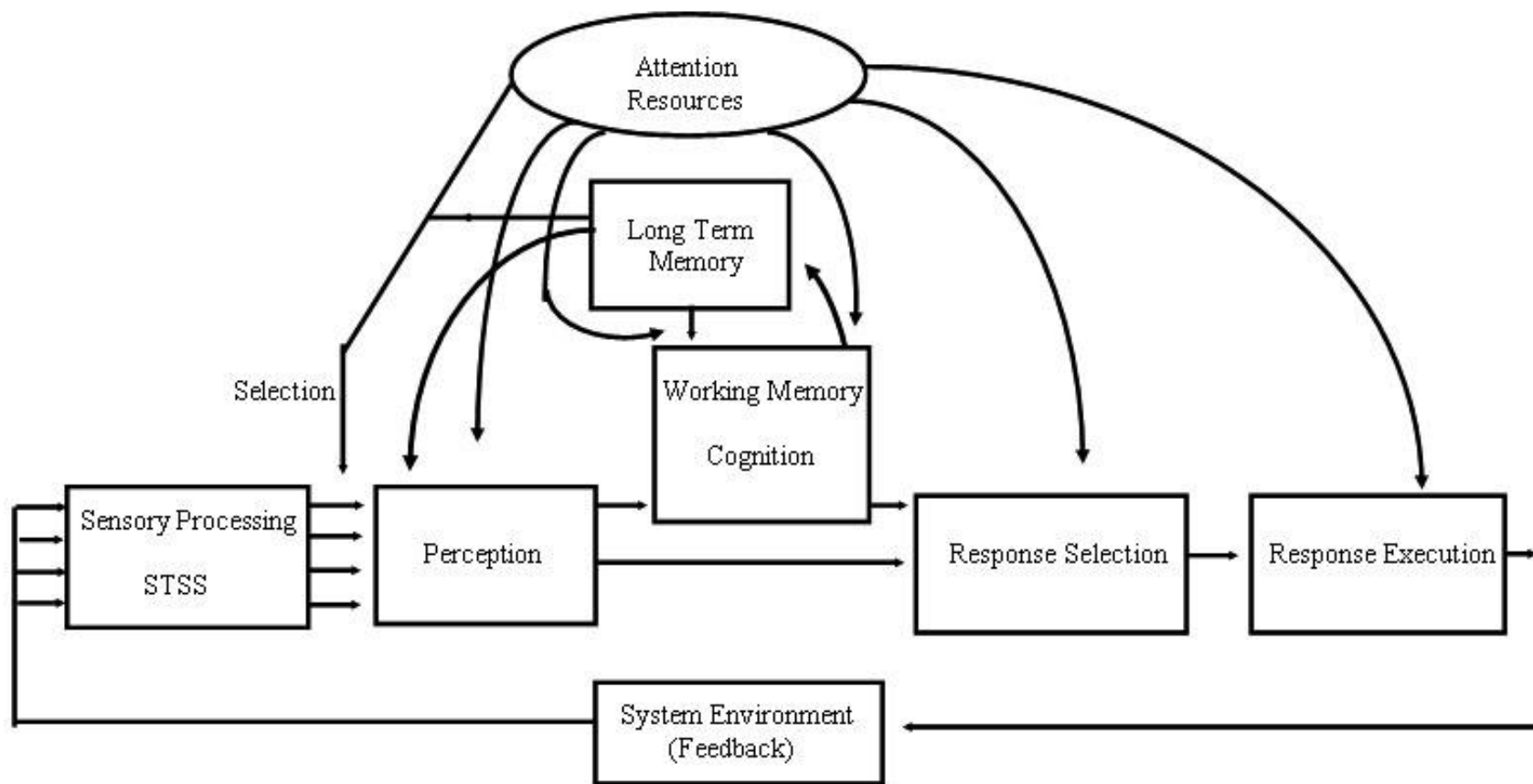


Figure 3. Wickens's (2007) Model of Information Processing

Thus, *attentional selection* is likely to be the result of *focal attention*; it happens when input is attended to during the first stage of the information-processing model shown above. According to Robinson, attention as selection does not require but is likely to involve awareness. Robinson equated *attention as selection* with Schmidt's *noticing*; it happens when input is initially attended and later, after it is further processed and stored in working memory. As such, selection also happens *after* the information has been analyzed. Unlike VanPatten's direct connection between the two, selection is independent of working memory capacity limits. Selection is a consequence of the nature of the task demands and the learner's associated objectives. This conceptualization reflects what Cowan suggested in 1995: "[T]here theoretically could be selective attention only because the subject would rather not encode useless information, and not because selection is needed in order to conserve a limited attentional resource" (p. 203).

Although Robinson does not view *attention as selection* as a result of limits in working memory capacity, he elaborates on working memory limits when referring to *attention as capacity*. With regard to the latter, Robinson discussed three theoretical models. The *single-capacity* model holds that "complex and less automatized tasks consume more attentional capacity, and require greater effort" (p.634). The multiple-resource model, adapted from Wickens (1992), takes the single capacity model a step further, stating that attention can be allocated to different tasks (i.e., divided attention), as long as the tasks require different processing mechanisms (e.g., auditory versus visual). Figure 4 below, which is adapted from Wickens (2007), shows that "resource competition is argued to exist within, but not between, separate attentional pools" (p. 644). In this sense, when language stimuli are processed via one modality only, as in reading (i.e., visually), there supposedly is *no* divide of attentional resources

and hence the model predicts there will be competition. Recall that VanPatten, however, claimed that if learners allocate attention to content words that convey the sentence meaning, they usually do not allocate their attention to form. Thus, the VanPatten's IP Theory and the Resource Competition Model convey differing concepts.

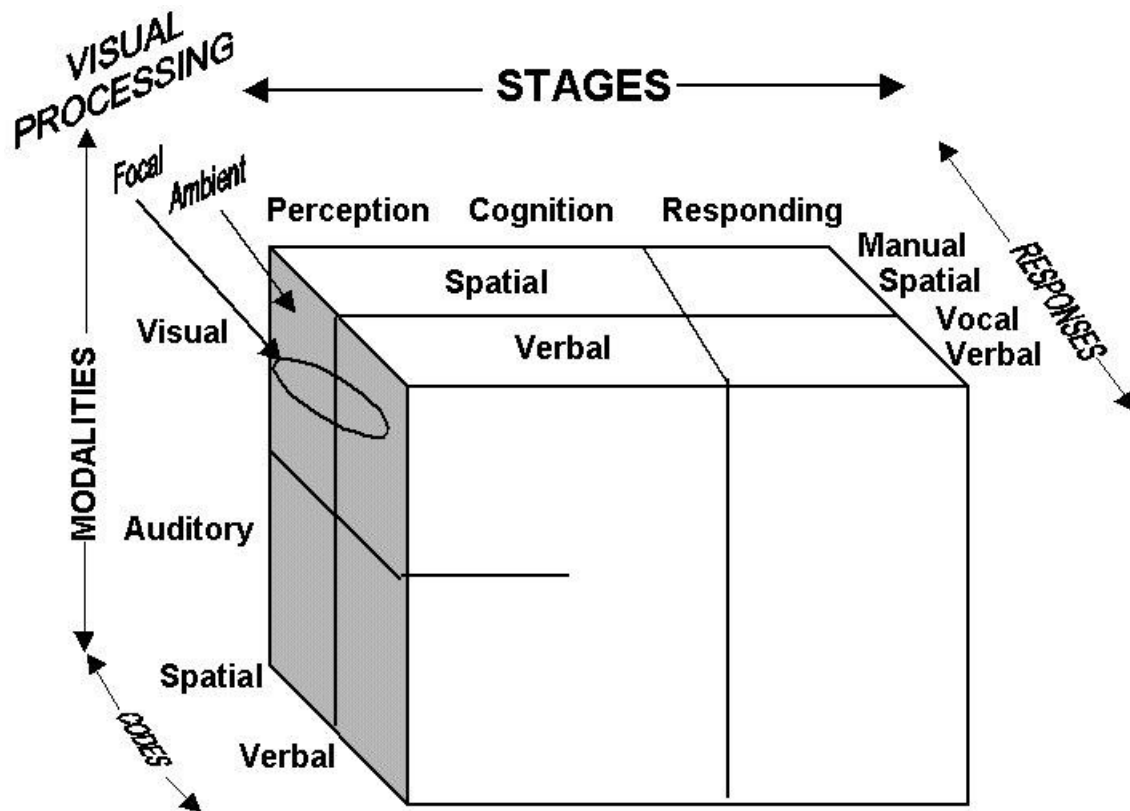


Figure 4. Wickens's Proposed Structure of Processing Resources (2007, p. 186)

The multiple-resource model—as used to theorize attention during L2 learning—received some criticism as researchers were unable to specify where and how attention differentiates between inputs from two language modalities and what happens if one modality requires more attentional resources than others (Rosa & O'Neill, 1999).

The last model that Robinson discusses in his account on *attention as capacity* is the *interference model*. The model demonstrates that learners' low performance on certain language tasks does not result from limited short-term memory capacity (e.g., McLaughlin, 1965, 1990; McLaughlin & Heredia, 1996), but from "involuntary attention shifts" (p. 645) induced by conflicting *codes* (of L1, interlanguage, and L2 syntax, morphology, semantics, and phonology/orthography).

Last, *attention as effort* refers to the extent to which a learner can focus his or her attention on a certain language task over period of time, which related to the notion of *sustained attention*. If *attention as effort*, or *sustained attention*, falls short, a learner's processing of incoming stimuli may lack the concomitant selection of information from input (i.e., a function of focused attention), or, in terms of output, a learner's oral production of language may include fewer self-repairs and less overall monitoring and planning of the output. Attention as effort thereby corresponds to Swain's (2005) notion of pushed output and Dörnyei's (2009) discussion of the influence of motivation on language processing.

Robinson (2003) agrees with Schmidt's (2001) view that a learner's focal attention selects which stimuli gets processed more deeply, or, in Schmidt's (2001) words "which parts of the input gain access to the learner's awareness" (p. 6). According to Robinson, the allocation of focal attention is a prerequisite for the linguistic form (word or structure) to enter into one's working memory, where the attended form will be processed further.⁹ In Robinson's words, "memory processes, such as maintenance and elaborative rehearsal, which allocation of focal attention activates, are co-responsible for noticing and the durability and extent of awareness that

⁹ This is the point where Schmidt assumes awareness to be involved.

noticing is accompanied by” (p. 637). Eventually, but not a guarantee, the linguistic feature will be stored in long-term memory. According to Robinson (1995), “more permanent encoding in long-term memory is a consequence of the level of activation in short-term memory and the result of rehearsal and elaboration.”¹⁰ This process is what Robinson and other scholars have defined as “learning.”

In this study focal attention (i.e., noticing) was measured via eye-movement recordings during reading. Thereby, I could determine whether the stimuli (regular and irregular verbs in German sentences) that received focal attention—stimuli that were noticed—might be processed more deeply. Since this study operationalized noticing as the learners’ *focal attention*, I cannot make strong claims about their awareness. Yet, based on prior eye-movement research, especially disappearing text studies¹¹ (e.g., Rayner, 2009; Rayner, Liversedge, White, & Vergilino-Perez, 2003; Rayner, Liversedge, & White, 2006), it is safe to assume that people are aware of (i.e., consciously perceive) any stimulus that they look at or attend to for more than 60 milliseconds (Godfroid, 2010; Liversedge & Findlay, 2000).¹² In this dissertation, I conceptualize noticing as follows: awareness is inherent in focal attention (i.e., noticing), and constitutes a continuum from perception (no or little awareness) to understanding on the lower end (full awareness) on the upper end. Following Schmidt (1995, 2001) my conceptualization of

¹⁰ According to Cowan (2001), short-term memory needs to be distinguished from working memory which is responsible for processes used for temporarily storing and manipulating information.

¹¹ The disappearing text studies will be explained in more detail in Chapter 2.

¹² The complexity of the relationship between attention and awareness has been explicated in various models in the field of cognitive science (e.g., Dehaene et al., 2006; Koch & Tsuchiya, 2007; Lamme, 2003), with the current consensus being that attention and awareness correspond to different cognitive circuits, whereby awareness influences attention.

noticing does not involve attention to *abstract or complex rules or principles*,¹³ but mostly focal attention to surface structures and lexical items that might be a novelty for the learner. Thus, I do not aim to measure awareness at the level of *understanding* (Schmidt, 1990) where learners are aware of generalizations of rules. Rather, my conceptualization is in line with Schmidt's distinction between noticing and metalinguistic awareness. Whether focal attention leads to deeper processing that gives rise to learning depends on where *attention* on the awareness continuum is located (see Figure 5). This idea is partly in line with Sharwood-Smith and Truscott (in press) who suggested the existence of *noticing thresholds*.

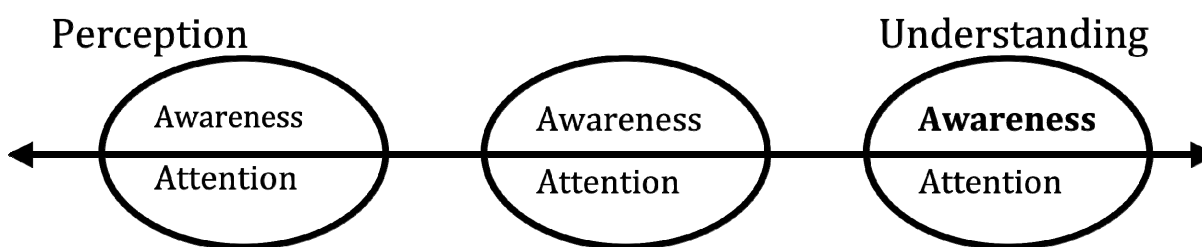


Figure 5. Continuum with Different Levels of Awareness

Leow's research (Leow et al., 2008; Leow, 1997, 2000) collectively suggested that learners who showed higher levels of awareness learned more than those who merely attended and noticed. His studies (1997; 2000) examined the learning of Spanish irregular verbs by adult native speakers of English. The participants were exposed to these forms via crossword puzzle tasks, and they were instructed to think aloud while working through the puzzles. Half of the participants noticed the verbs' stem-vowel changes; that is, they commented on them in their concurrent verbal reports. The learners who verbally reported on the stem-vowel changes

¹³ But I do not deny that there is the possibility of noticing at the level of metalinguistic understanding (Schmidt, 2001, p. 19)

displayed statistically significant gains in accuracy on producing stem-vowel changing verbs on their posttests (compared to a pretest). Those participants who did not comment any noticing of stem changes did not improve on the posttests.

In this study, I assume awareness is a *complex* process, and a quality of noticing.¹⁴ However, it is important to note that Williams (2004, 2005) found evidence for learning *without* awareness. Williams (2005) explored whether adult language learners can implicitly learn associations between grammatical morphemes and the underlying meaning. The study participants, native and non-native speakers of English, read English sentences in which the article ‘the’ or ‘a/an’ were replaced by *gi*, *ul*, *ro*, or *ne*. They were instructed that the choice of the determiner was based on whether the object was “near” or “far” from the subject of the sentence. Yet, in addition to the distance criterion, the choice of the determiner was also dependent on whether the object was animate or not. Williams aimed to find out whether the participants became aware or remained unaware of this extra dimension of the rule, and how their level of awareness was reflected in their learning gains. He found that some participants were able to correctly generalize the rule of animacy to new determiner-noun combinations, while being *unaware* of the actual rule or the relevance of animacy. Williams concluded that learning without awareness is possible.

Hama and Leow (2010) conceptually replicated William’s (2005) study with methodological extensions regarding the measurement of ‘awareness,’ employing what they termed a ‘hybrid research design.’ They qualitatively measured learners’ awareness during the experiment via think-aloud protocols (online) and after the experiment via questionnaires (offline). The online data showed that some participants were aware of the animacy rule during

¹⁴ Note that I do not state that noticing does not require awareness, but is a quality of awareness.

task performance. Those learners improved on posttests, while “unaware participants did not appear to demonstrate significant learning of the animacy feature of the novel words at the production level of trained items or at the level of generalizations to new items” (p. 482). Thus, Hama and Leow concluded that there is “no dissociation between awareness and learning” (p. 482).

To conclude, the *exact* roles of attention and, especially, awareness in language learning remain unresolved to date. While we can assume that some form of attention is necessary for the integration of new linguistic information (whether it may lead to learning or not), SLA researchers have yet to establish models that capture the complex interaction of attention and working memory, as well as attention and awareness,¹⁵ and awareness and learning. Godfroid (2010) provides the probably most useful account to date: “Noticing refers to a sequence of events involving attention and awareness, *the quantity (duration) and quality (nature of mental operations) of which will determine the durability of the memory trace created [...]*” (p. 183, highlights are mine). In this study, I am defining noticing as a process where attention and awareness (two distinct processes in the mind) merge, whereby the attention paid to a new linguistic structure exceeds a threshold (Godfroid, 2010, p. 111) so that this novelty is held in working memory for long enough to increase the (initially low) level of awareness. If this noticing process (attention and awareness) occurs with sufficient *frequency*, the linguistic item will eventually build a representation/become a stored exemplar in long-term memory. While, in my view, attention plays a more prominent role than awareness, the latter also plays an important

¹⁵ However, recall from an earlier footnote that researchers in neurophysiology and cognitive science argue that attention and consciousness are *distinct* (not isomorphic, as Schmidt suggested) because they have distinct functions neuronal mechanisms (B. J. Baars, 1997; B. J. Baars, 2002; Block, 2005; Dehaene et al., 2006; Koch, 2004; Koch & Tsuchiya, 2007; Lamme, 2003; Naccache et al., 2002; Woodman & Luck, 2003 Dehaene et al., 2006).

role as it is likely to reflect the “nature of mental operations” that learners engage in while the linguistic form is in focus of their attention (Godfroid, 2010, p. 183).

Implicit versus Explicit Learning

The constructs of awareness and attention, as discussed above, lay the foundation of a critical understanding of the concepts of explicit and implicit learning in SLA. The notions of awareness and consciousness are intrinsically linked to the implicit/explicit learning distinction. In the field of cognitive linguistics and SLA, the role of explicit and implicit knowledge has evolved as an important area of investigation (e.g., DeKeyser, 2003; Dörnyei, 2009; N. Ellis, 1994a, 1994b, 2002, 2005, 2007; R. Ellis, 2002; Hulstijn, 1989, 2003, 2005; Reber, Allen, & Reber, 1999; Robinson, 1996, 1997a, 1997b, 2005; Schmidt, 1994b; Shanks, 2005). DeKeyser (2003) defined implicit learning as “learning without awareness of what is being learned” (p. 314). His definition goes a step beyond the work of Reber—the pioneer of implicit learning research—who defined implicit learning as “the process whereby knowledge is acquired largely independently of awareness” (1999, p. 475).¹⁶ Reber rejected the possibility that there can be a complete lack of awareness, suggesting that both implicit and explicit processes are involved in “virtually everything interesting that human beings do” (p. 488). Similarly, J. N. Williams (2005) considered the explicit/implicit distinction as a continuum as opposed to a dichotomy. At the *implicit* end of the spectrum, awareness is not involved in the learning process and product, while at the *explicit* end of the spectrum, learners are entirely conscious of

¹⁶ In reviewing empirical research that targets artificial grammars, sequence learning, and control of complex systems, DeKeyser (2003) concluded that “implicit learning is necessarily rather concrete and really abstract learning is necessarily explicit” (p. 333). This points to the importance of clearly defining the objects of learning in implicit learning research; that is, are they concrete items or abstract rules?

their internal input comparisons and hypothesis testing (p. 271). In other words, explicit knowledge can be brought to awareness and can be verbalized, while implicit knowledge cannot (Hulstijn, 2002, 2003, 2005). The current consensus is that although implicit and explicit learning are separable and distinct, humans engage with them in interplay, so implicit and explicit learning mechanisms can influence one another (N. Ellis, 2005, 2011; Segalowitz, 2003; Dörnyei, 2009).

The cognitive science literature (e.g., Anderson, 2005; Cleermans & Destrebecqz, 2005; Hampton, 2005; E. Smith, 2005; Taylor, 2003) is much in line with N. Ellis's (2005, 2011) views. Cognitive psychologists usually contrast explicit and implicit *processing* mechanisms as follows: Explicit processing is conscious rule-based processing, while contextual variations are neglected. Implicit processing is similarity-based processing; the input is matched with properties of existing representations.

To give a practical example, when learners use their L2 fluently, they draw on implicit processes and their attention is focused on meaning rather than on form (R. Ellis, 2005; VanPatten, 1990; 2003). Explicit processes usually take over when learners experience comprehension and production difficulties (N. Ellis, 2005). In such cases, learners deliberately pay attention to language form; that is, they analyze the input they receive and/or the output they produce. In doing so they utilize their explicit knowledge and engage explicit processes.

According to N. Ellis (2011), L2 learners engage in explicit learning when it comes to the "fragile" aspects of grammar (p. 1), which are not acquired in spite of their frequent occurrence in the L2. Irregular verbs are an example of such fragile aspects of grammar. "[L]earners fail to notice cues which are lacking in salience and redundant in cuing meaning, or because of interference where the features need to be processed in a different way from that usual in their

L1” (p. 1). N. Ellis emphasized the crucial role of noticing and attention in the acquisition of non-salient L2 forms and the importance of form-focused, explicit instruction in this respect. He provided the example of the *frequency effect*; in particular, learners might implicitly acquire knowledge of associations between language elements because they encounter these elements together so frequently (albeit *incidentally*) in the input and because learners acquire ‘probabilistic knowledge’ (p. 11). The latter refers to knowledge of the probabilities of occurrence of words that are like to follow or precede other words (transitional probabilities), and form-function mappings. Yet, there is an interface between explicit and implicit knowledge. “Many aspects of language are unlearnable, or at best only very slowly acquirable, from implicit processes alone. Which is why an attentive focus on the form-meaning relation is also necessary in the initial registration of pattern recognizers for constructions” (p. 12). Learners might not attend or register new patterns or forms because they are not salient or because learners falsely process them just as they process their L1, not noticing the relevant particular cues.

In German, for example, morphological verb inflection in the present tense is redundant because verbs are accompanied by the sentence-subject that indicates the person and number. Thus, the actual verb inflection (ending and vowel stem change for irregular verbs) is redundant in the sense that it is not the only source that indicates whether the subject is singular, plural, or whether the person corresponds to *I, you, he/she*, etc. Thus, this structure is likely not salient to learners (especially if their L1 morphology lacks this type of inflection), or they process the inflected verbs just as they process verbs in their L1, English—without paying attention to morphological changes. In cases like these, *explicit* learning plays a crucial role (N. Ellis, 2005). To help the learner to ‘notice’ the cue in the first place, form-focused instruction or consciousness-raising tasks are needed (Fotos, 1993, 1994; Sharwood-Smith & Truscott, 2010). In fact, there is

strong empiric evidence that instruction that guides learners' attentional foci to morphological cues in the input is effective (Doughty and Williams 1998; Ellis and Laporte 1997; Hulstijn and DeKeyser 1997; Lightbown, Spada, and White 1993; Long 1983; Spada 1997). Norris and Ortega (2000) conducted a comprehensive meta-analysis revealing that explicit L2 instruction is generally more effective than implicit instruction.

If learners receive instruction on a particular grammatical structure, they will build up explicit memory of that structure. When learners then encounter a certain structure on which they have already received instruction, it is likely that the learners engage in a conscious process of remembering their prior learning experience of that structure, so they are aware of the knowledge they have. In contrast, learners show evidence of implicit memory “when there is facilitation of the processing of a stimulus as a function of a prior encounter with an identical or related stimulus but where the subject at no point has to consciously recall the prior event” (N. Ellis, 2010, p. 18). For example, learners of L2 German would have *implicit* memory or knowledge when they inflect an irregular verb correctly without being able to consciously recall they had ever seen this verb form before. These facts about implicit *memory* notwithstanding, implicit *learning* still requires some attention to the formal features of the language. While this seems apparent for the learning of content words; implicit learning is also likely to involve some attention to form, even if learners are not conscious or aware of the relevant aspects of the target form in terms of the underlying structure or “rule”.

To date, SLA researchers (e.g., DeKeyser, 2003; Dörnyei, 2009; N. Ellis, 1994a, 1994b, 2002, 2005, 2007, 2010; R. Ellis, 2002; Hulstijn, 1989, 2003, 2005; Perruchet & Pacteau, 1990; 2006; Reber et al., 1999; Robinson, 1996, 1997a, 1997b, 2005; Schmidt, 1994b; Segalowitz, 2003; Williams, 2004, 2009) and cognitive psychologists (Cleermans & Destrebecqz, 2005;

Mathews et al., 1989; Paradis, 1994; Shanks & St. John, 1994; Shanks, 2005) have generally concluded that much more research into the complementary systems of explicit and implicit learning is needed in order to obtain a clearer picture of the contribution of attention and awareness in the establishment of coherent representations of language structures in the mind. In Ellis's (2011) words, more research is needed into the "ways that linguistic constructions are first noticed and registered, and thence figured and tuned into the system" (p. 20). The notion of explicit and implicit learning processes will resurface in the remaining chapters of this dissertation. In particular, I will discuss the noticing construct in regard to specific linguistic structures they may function as targets for noticing. The notion of implicit and explicit language processing is important in regard to the learners' memory representations of the noticed forms. Specifically, different types of L2 forms (e.g., lexical items or morphosyntactic target structures) may be represented differently in the brain—either symbolically or associatively. Consequently, different L2 structures invite different levels of engagement of explicit and implicit processes.

Intentional versus Incidental Learning

The construct of implicit versus explicit learning is intrinsically related to the construct of intentional versus incidental learning. Intentional learning involves a deliberate attempt to learn; and this is very likely to involve awareness. For example, a learner might read a text with the intent to increase his/her vocabulary and is therefore likely to consciously attend to unfamiliar words in the text. In contrast, incidental learning is marked by the absence of a conscious intent to learn, but it might involve unplanned conscious attention to words or formal features of the L2. Precisely, "[t]he term incidental learning is used, in applied linguistics, to refer to the acquisition of a word or expression without the conscious intention to commit the element to

memory, such as “picking up” an unknown word from listening to someone or from reading a text” (Hulstijn, in press, p. 1).

Explicit learning, as explained in the previous section, overlaps with intentional learning to the extent that it is a conscious process and a likely intentional one as well. Yet, explicit attention to some feature can occur *incidentally* without planning, for example, when a learner experiences difficulties or attends to some unknown words or features in the L2 input just because he or she is eager to learn. In such a case, explicit and intentional learning diverge.

There is a paucity of research that investigates incidental/intentional learning of L2 *grammar*. Yet, there is a wealth of SLA literature on incidental L2 vocabulary acquisition (see Hulstijn, 2003, for a review). Incidental focus on L2 grammar structures usually occurs when learners naturally focus on meaning, but then experience difficulty with a certain linguistic structure and incidentally attend to form (e.g., Loewen, 2005). Such an experience might facilitate “the cognitive mapping among forms, meaning, and use that is fundamental to language learning” (Doughty & Williams, 2001, p. 221).

Studies on incidental learning of grammar (not vocabulary) in the spoken modality have largely analyzed L2 learners’ spoken interaction with teachers or native-speaking interlocutors. In contrast, most studies on the learning of a grammatical form during *reading* are input enhancement studies. Although it could be argued that input enhancement studies—due to the very purpose of the input manipulation—do not tap into *incidental* learning, many input enhancements studies have had a control group that received the texts without enhancement, so incidental grammar learning was indirectly examined with those learners. The meta-analysis by Lee and Huang (2008) yielded that “learners exposed to enhanced texts outperformed learners

who read unenhanced texts by a very small-sized effect ($d = 0.22$)” (p. 323).¹⁷ Lee and Huang suggested that the learners who read unenhanced texts might have improved their grammar knowledge due to the input flood, “which in itself is thought to have small but detectable effects on grammar learning” (p. 323). In other words, the learning of L2 grammar structures during reading seems quite feasible if the structure occurs with sufficient frequency. Yet, the input enhancement studies provide limited insight into the cognitive processes that occur during reading, and how the learners process the enhanced or unenhanced forms, as well as what types of learning mechanisms account for improvements on posttests. Research on incidental learning targeting oral interactions were better able to tap into cognitive (especially attentional) processes, as verbal data can provide of incidental attention to form, the effects of which can be tested with post-tests (e.g., Ellis, Basturkmen, & Loewen, 2002; Loewen, 2005; Swain & Lapkin, 2002; J. Williams, 2001; J. N. Williams, 2004). This dissertation study investigates—via eye-movement measures—how L2 structures are processed in real time during reading, so it probes into the incidental learning of L2 verb morphology.

In Robinson’s view (2005), working memory and the capacity to rehearse sub-vocally forms that have been attended to jointly determine the possibility and extent of incidental learning. “[W]orking memory is the strongest individual difference predictor of incidental SLA; [it] forms part of the aptitude complex for learning from incidental exposure to written input” (p. 261). Thus, learners with higher working memory capacity might be more apt to incidentally and implicitly acquire underlying L2 systems than learners with more limited working memory capacity.

¹⁷ In another meta-analysis on input enhancement by Han, Park and Combs (2008), similar results were obtained.

In summary, Schmidt (1990) put forth key arguments on incidental, implicit, and subliminal learning (the first two will be discussed in detail in the following sections). Concerning the first notion, he claims “incidental learning is certainly possible when task demands focus attention on relevant features of the input” (p. 149). The issue of implicit learning is more complex. He stated that implicit learning is possible, but “there is no reason to accept [...] that awareness does not affect language learning, [...] or that most language learning is implicit” (p. 149). In this sense, Schmidt supported his critical notion that “subliminal learning is impossible, and that intake is what learners consciously notice” (p. 149). As such, learners must consciously notice aspects of language (including vocabulary, phonology, grammatical form, and pragmatics) in the input in order for it to be registered in memory. This premise is the heart of the *Noticing Hypothesis*, which is detailed in the following chapter.

CHAPTER 2 NOTICING

The *Noticing Hypothesis*—as proposed in the 1990s by Schmidt (1990; 1993a, 1993b, 1994, 1995)—is that noticing requires awareness and is a *necessary* condition for second language acquisition. The *Noticing Hypothesis* rejects subliminal learning (i.e., learning for acquisition). Figure 6 demonstrates the commonly used visualization of the *Noticing Hypothesis*.

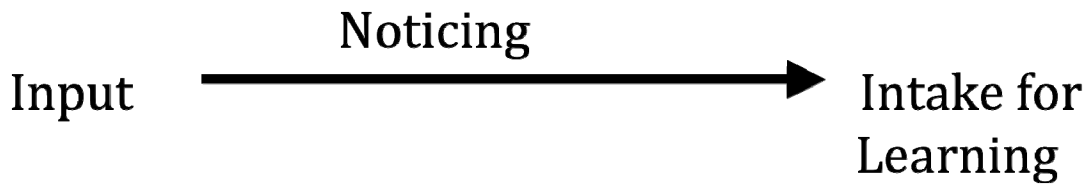


Figure 6. Graphic Representation of the Noticing Hypothesis

(adopted with permission from Godfroid, 2010)

According to Schmidt, only what has been noticed (i.e., consciously registered) in the input can become intake. More precisely, in 1993 Schmidt specified noticing as “the necessary and sufficient condition for the conversion of input to intake” (p. 209). Yet, “this characterization does not specify exactly what intake is, other than that it is the *product of noticing* and an intermediary step in the acquisition process” (Godfroid, Housen, & Boers, 2010, p. 170).

In his latest essay (2010), Schmidt maintained that it is common sense that learning is driven by attention, as we “don’t learn much about things we don’t attend to” (p. 721). Regardless of the intuitive plausibility of this notion, the early version of the noticing hypothesis has been objected to on three major grounds: (1) that noticing is a prerequisite for learning; (2)

that noticing does involve awareness; and (3) noticing is not related to linguistic knowledge, (competence), but appertains to metalinguistic knowledge.¹⁸

Regarding the first objection, Gass (1997) and Leow (1997), among others, disagreed with Schmidt's premise that conscious attention is a *necessary prerequisite* for L2 acquisition. According to their view, attention is not a prerequisite for all learning as learners are able to, for example, generalize a certain grammatical structures to new ones. Schachter (1998) claimed that only language features pertaining to lexical semantics (individual words) and phonetics (novel phonetic inventories) require focused attention and that all other linguistic features do not. A majority of SLA and cognitive psychologists have evidenced that Schachter's claim is not valid (Perruchet & Pacteau, 1990; Perruchet & Pacton, 2006; Shanks, 2005; J. N. Williams, 2009).

Gass, Svetics, and Lemelin (2003) discredited the premise that noticing is a *necessary* condition for learning. Their research findings demonstrated that L2 acquisition can take place without focused attention. They assigned Italian learners at different proficiency levels to one of two conditions ([+focused attention] and [−focused attention]) for each of the following linguistic features: syntactical, morphosyntactical, and lexical. Their hypothesis that focused attention would have the least effect on syntactical acquisition and the largest effect on the lexical acquisition was rejected. In fact, their findings revealed the opposite effect. Contrary to

¹⁸ A third objection came from Carroll (Carroll, 2001, 2006a, 2006b, 2006c). Carroll criticized the Noticing Hypothesis in that the characteristics of input available for noticing are unspecified, and that linguistic representations exist in the mind to begin with and are not mediated by environmental stimuli. Carroll's (2006a) and (2006b) publications, which deal with phonetic and phonological perception (prosody, specifically), are best summarized by Godfroid (2010). "[According to Carroll,] language-specific coalitions of [prosodic] properties are acquired well-below the threshold of awareness. On this account, they constitute a case of unaware phonetic and phonological learning and may thus count as evidence against the Noticing Hypothesis" (p. 80, insertions are mine).

their predictions, learners in the non-focused-attention condition showed greater focus on lexis than syntax. With regard to proficiency, focused attention was most beneficial for beginning learners and least beneficial for intermediate learners.

The second main objection to the *Noticing Hypothesis* targeted the claim that noticing involves awareness; that is, while there is subliminal, unconscious perception, there is no subliminal, unconscious learning (Schmidt, 1995, 2001). As explained in Chapter 1, Tomlin and Villa (1994) proposed that alertness, orientation, and detection (i.e., the closest equivalent to Schmidt's noticing) do not require awareness. Similarly, Williams (2005) found that language learning might happen without awareness. This claim was addressed and largely refuted by a series of studies and reviews by Leow and co-authors (Hama & Leow, 2010; Leow et al., 2008; Leow, 1997b, 1998, 2000, 2002; Rosa & Leow, 2004). In addition, from a theoretical perspective, Robinson (1995, 2003, 2009) and Schmidt (1990, 1994a, 1995, 2001, 2010) maintained that the representation of novel linguistic forms involves focal attention *and* awareness.

Schmidt has acknowledged that the objections to the *Noticing Hypothesis* (mentioned above) have merit and cannot be dismissed easily. In 2001, Schmidt formulated a weaker version of the *Noticing Hypothesis*, arguing that noticing may not be necessary for, but certainly is *facilitative* of L2 learning in that “more noticing leads to more learning” (2001, p. 8).

Last, the *Noticing Hypothesis* has been criticized because it does not relate to unconscious linguistic representations (competence); rather it relates to the acquisition of metalinguistic knowledge (Truscott, 1998). Truscott's opposition stems from the conception that attention is a “very confused” and “poorly defined commonsensical” concept (p. 105). As noticing, in Truscott's view, merely relates to metalinguistic knowledge, and not to unconscious

language competence, it is “not only unnecessary but also unhelpful” (p. 126). In his 2000 essay “Unconscious second language acquisition: Alive and well” Truscott maintained that SLA research on attention and awareness “rejects unconscious acquisition” (Truscott, 2000, p. 115). Clearly, no SLA researcher has rejected the possibility of implicit acquisition; in fact, there is a wealth of literature on the intricate relationships among implicit and explicit processes in adult SLA (e.g., DeKeyser, 2003; Dörnyei, 2009; N. Ellis, 1994a, 1994b, 2002, 2005, 2007, 2010; R. Ellis, 2002; Hulstijn, 1989, 2003, 2005; Perruchet & Pacton, 1990; 2006; Reber et al., 1999; Robinson, 1996, 1997a, 1997b, 2005; Schmidt, 1994b; Segalowitz, 2003; Williams, 2004, 2009) but, at the same time, there is *substantial* L2 research data supporting the idea of noticing as a facilitating factor in L2 learning (e.g., Adams, 2003; Carr & Curran, 1994; Carrell, 1992; DeKeyser, 1997; R. Ellis, 1994; Fotos, 1993, 1994; Gass & Alvarez-Torres, 2005; Hama & Leow, 2010; Izumi, 2002; Izumi & Bigelow, 2000; Izumi, Bigelow, Fujiwara, & Fearnow, 1999; Izumi & Izumi, 2004; Leow et al., 2008; Leow & Morgan-Short, 2004; Leow, 1997b, 1998, 2000; Long, 1996; Mackey, 2006; Mackey, Gass, & McDonough, 2000; Norris & Ortega, 2000; Roberts, 1995; Rosa & Leow, 2004; Rosa & O'Neill, 1999; Sachs & Suh, 2007; Swain, 1995, 1998, 2005; Swain & Lapkin, 1995; Van Hell & De Groot, 1999). Thus, the prevalent notion to date is that noticing, which involves attention and some level of awareness, greatly facilitates language learning, but may not be absolutely necessary for learning to occur; at least some types of learning can also occur implicitly, without awareness.

Swain (1998) suggested that there are several levels of noticing: Learners may simply notice a structure in the target language due to the frequency or salience of that form (see also Bardovi-Harlig, 1987; Jourdenais, Ota, Stauffer, Boyson, & Doughty, 1995) or they may notice that the target language is different from their own interlanguage (i.e., *noticing the gap*, see also

Doughty & Williams, 1998; Gass, Mackey, & Pica, 1998; Philp, 2003). Godfroid (2010) did not agree with Swain's point that noticing and noticing the gap refers to different *levels* of noticing, but noticing and noticing the gap are different *kinds* or *types* of noticing. According to Godfroid, awareness may have different *levels*, one of which corresponds to noticing as proposed by Schmidt. For SLA researchers it is more useful to distinguish between two types of processes: (1) noticing *per se* during input processing and (2) noticing *the gap* during output processing. The mechanisms underlying these two processes may not be the same; noticing the gap may function as a *priming* device for noticing (Godfroid, 2010, see pp. 70-72 in particular).

As mentioned in the previous chapter, Schmidt (2001) and Leow (1997) proposed a distinction between noticing and metalinguistic awareness. They suggested two levels of noticing: simple noticing (i.e., registration with awareness) and noticing with metalinguistic awareness (understanding). Noticing might be a necessary and sufficient condition for learning, but understanding is not. In other words, learners can learn a new language structure without necessarily understanding the underlying rules of that specific structure. Recall from Chapter 1 that Schmidt (1990) claimed that unconscious learning can refer to either (1) incidental (i.e. non-intentional) learning, (2) learning without understanding (i.e. implicit learning) or (3) learning without awareness (i.e. subliminal learning). It follows from the above that Schmidt acknowledges the possibility of (2), but rejects the notion of (3).

Noticing and Learning

Usually, reports of experimental research concerned with language acquisition leave the reader to wonder whether the improvements on posttests reflect learning. Even more important, the experiment results do not provide information on which *type* of learning takes place inside the learners' mind. Just based on the results reported in the research articles (Alanen, 1995;

Leow, 1997, 2000; Long, 1991; Rosa & O'Neill, 1999; Sachs & Polio, 2007; Sharwood Smith, 1993; VanPatten, 1990, 1996, 2002), we do not know whether learners were able to generalize the target structure across items and apply their acquired knowledge to new situations. Leow (1997, 2000), for example, found that aware learners significantly improved their production of correct Spanish verb-morphology. According to Leow (2000), “[t]hese findings provide further empirical evidence for the association between awareness and subsequent processing of L2 data” (p. 568). Yet, did these learners “learn” the stem-changing Spanish verb forms as *items* (similar to the way that they would learn a vocabulary item) or did they learn the underlying *system* (i.e., the morphologic rule) that governs Spanish verb morphology in the past tense? In J. N. William’s (2005) words, “we must be careful in defining learning” (p. 271). In his research (1997, 2000), Leow assumed that his participants were equally aware of *all* the items (i.e., verb forms they filled out in the cross-word puzzle task). This operationalization of one level of awareness (i.e., either aware or unaware) for all stimuli has implications for how *learning* can be traced back to the processing of the input. That is, based on this operationalization, we do not know (a) whether the learners’ awareness of the structure *increased* with the completion of the cross-word puzzle task, and consequently, (b) whether the learners’ processing shifted from noticing (attending with a low level of awareness) to understanding (attending with a high level of awareness) over the course of the task. Thus, it remains unknown whether the learners established a foundation or data base of consciously stored exemplars, based on which they derived the understanding of the morphological rule. This is what usage-based accounts of language learning (e.g., N. Ellis, 2002a) would predict. When Leow’s (aware) learners demonstrated their learning, they only showed evidence of *item-based* learning. As such, the notion of *item-based learning* and *system-based learning* is quite relevant (Shanks, 2005),

especially when it comes to studying the fine details of the learning process. System learning involves the understanding of underlying rules so that rule generalizations to new items are possible and likely. Item learning, in contrast, refers to the rote memory of items.

Shanks and St. John (1994) investigated the issue of item and rule learning and, thereby, the possible existence of independent explicit and implicit learning systems. Their research also had ramifications regarding the involvement of *awareness* and the depth and *type* of learning. Shanks and St. John inquired whether learning mainly involve the encoding of language fragments (which we may think of as ‘items’) or, whether it goes deeper in that it involves the induction of abstract rules? Based on the premise that implicit learning equals unconscious rule learning, Shanks and St. John found, among other facts, that unconscious learning has not been satisfactorily established in artificial grammar learning (see also DeKeyser, 2003). “When subjects cannot report the “implicitly learned” rules [...], this is often because their knowledge consists of *instances or fragments of the training stimuli* rather than rules” (p. 367, highlights are mine). “Human learning is almost invariably accompanied by conscious awareness, and in tasks such as artificial grammar learning, where learning is frequently thought to involve rule abstraction, performance is most often based on the acquisition of *instances or fragments from the training stage*” (p. 394, highlights are mine). Shanks and St. John therefore suggested that the distinction between (conscious) instance learning and (conscious) rule learning is more theoretically sound and useful than the distinction between conscious and unconscious learning.

In contrast, according to Robinson (1995), once specific linguistic input is encoded,¹⁹ generalizations across instances *can* happen without the learners being aware of those

¹⁹ Recall that according to Robinson (1995, 2000), *noticing* means that specific linguistic input is encoded in working memory

generalizations. If learners are aware of their generalizations, according to Schmidt (1990, 2008), this corresponds to “awareness at the level of understanding” (1990, p. 145). Leow (1997, 2000) concluded that awareness at the level of understanding produces notably more learning than awareness at the level of noticing (see also Qi & Lapkin, 2001; Rosa & Leow, 2004; Rosa & O'Neill, 1999). However, as established by Leow, Johnson, and Zárte-Sández (2011), it is both conceptually and practically very difficult to determine different *levels* of awareness and/or a complete lack thereof.

It is important to note that none of the noticing studies cited in this dissertation claims that there is a *causal* relationship between noticing and learning. It is unlikely that any learning is caused by one single factor (e.g., noticing) or one instance of the occurrence of this very factor. We can, however, safely assume that there is a positive relationship between noticing and learning, and that noticing is a complex concept composed of and interacting with a range of other cognitive factors (e.g., attention, awareness, aptitude, developmental level, the specific language form, social and linguistic context, etc.). Although there is a general consensus among prominent SLA researchers that noticing is at least facilitative of learning, the operationalisation of noticing and the research methods used to tap into the assumed relationship between noticing and learning have been subject to much debate (e.g., Bowles & Leow, 2005; Bowles, 2008, 2010; Yoshida, 2008). The following section addresses the issue of operationalization.

The Measurement of Noticing

The scrutiny concerning the operationalization of noticing is largely due to the difficulty in defining *noticing* so that it can be measured. Operationalizing noticing is difficult because it requires introspection—the assessment of learner-internal cognitive activities. (For a detailed overview of L2 introspection studies, see Bowles, 2010; Gass & Mackey, 2000.) However, not

all measures of learner-internal cognitive activities are introspective: the registration of eye-movements (i.e., by means of eye-tracking technology) is not, for instance.

Schmidt (1995) proposed to operationalize noticing in terms of learners' self-reports either during or immediately after exposure to input. However, a lack of self-reporting should not be interpreted as lack of awareness, as some thought processes during exposure to input are difficult to verbalize (e.g., Jourdenais, 2001; Schmidt, 1995, 2001). In addition, noticing incidences might involve a low level of awareness that is not 'reportable'. The challenge facing the measurement of noticing is to accurately link observable and/or self-reported behaviors by language learners to the construct of noticing.

Methods used to qualitatively and quantitatively account for learners' noticing of specific target language features, words, or phrases fall into two categories: online, which measures language learners' noticing during performance of a certain language task, and offline, which employs a post-treatment assessment of noticing. However, neither online nor offline methodologies enable an absolute account of the learners' attentional processes (e.g., Egi, 2004; Robinson, 1997a, 1997b; Williams, 2005). Nevertheless, each has some value: Offline measures do not interfere with the language task completion by the learners, while online measures offer instantaneous assessment of cognitive processes.

Think-aloud protocols, which require the participants to monitor (in most cases) and orally self-report their mental processes, are the most common online method (see Bowles, 2010, for a thorough review and meta-analysis of studies that have investigated the reactivity of the think-aloud method). Some research has attempted to measure language learners' noticing through written online tasks such as note taking (Hanaoka, 2007; Izumi & Bigelow, 2000; Izumi, Bigelow, Fujiwara, & Fearnow, 1999), circling of words (Leow, Hsieh, & Moreno, 2008) and

underlining (Izumi & Bigelow, 2000; Schierloh, in press). The advantage of online measures, as opposed to post-exposure measures, is their instantaneous access to L2 processing, thus minimizing the risk of possible memory decay by the L2 learner (Gass & Mackey, 2000, 2007). Stimulated recall has evolved as a frequently used post-exposure method of obtaining data of the language learners' thought processes (e.g., Mackey, 2006; Mackey et al., 2000; Roberts, 1995; Swain & Lapkin, 2002). During stimulated recall, learners are prompted with a stimulus (e.g., the learner's written product or a video displaying the learner's actions while engaging in a language task), and they are asked to report on their thought processes during the performance of the language task.

The problem with protocol analyses such as think-aloud protocols and stimulated recall is the *reactivity* (Bowles, 2010). Reactivity refers to the extent to which the observed participant performance (e.g., their mental engagement with a task or their posttest performance) is a result of or caused by the actual research treatment (i.e., the verbalization).

Think-aloud protocols, for example, might be reactive; that is, the verbalization may change the nature of actual cognitive processes (e.g., R. Ellis, 2001; Jourdenais, 2001). In addition to interfering with cognitive processes, thinking aloud may add an additional processing load on the learners. Therefore, verbal reports might not reflect real thought processes. The think aloud data method itself acts as an additional task which must be considered carefully when examining learner performance" (Jourdenais, 2001, p. 373).

Godfroid (2010) identified eight studies on the reactivity issue associated with think-aloud protocols in SLA: Bowles & Leow, 2005; Bowles, 2008; Leow & Morgan-Short, 2004; Rossomondo, 2007; Sachs & Polio, 2007; Sachs & Suh, 2007; Sanz, 2009; and Yoshida, 2008. As described by Godfroid, the findings of these studies are mixed. Leow and Morgan-Short

(2004) provided empirical data showing no reactivity of thinking aloud on text comprehension, intake and production with beginners. Bowles and Leow (2005) explored the reactivity issue by investigating the differential effects of type of verbalization (non-metalinguistic vs. metalinguistic) with advanced language learners, showing that neither type of think-aloud protocol caused reactivity. Yet, the learners' overall text comprehension was negatively affected by metalinguistic verbalization. Similarly, Bowles (2008) found that metalinguistic verbalization interferes with item learning and extends the time needed for completion of the task. Sachs and Polio (2007) investigated the reactivity of think-alouds as a research tool as a secondary question in their study on learners' uses of different types of written feedback on an L2 writing revision task. The think-aloud protocols appeared to be reactive, so Sachs and Polio emphasized that think-aloud protocols need to be implemented and interpreted with caution. Rossomondo (2007) had her participants engage in *either* silent reading *or* think-alouds to investigate the role of lexical cues to infer meaning in the incidental acquisition of the Spanish future tense. The think-aloud protocol appeared to have an effect on the recognition and production of the Spanish future tense. The findings were interpreted in that thinking aloud, in contrast to reading silently, encouraged the learners to pay attention to the target structure. Last, Sachs and Suh (2008) evidenced a reactivity effect of thinking-aloud on a text completion task, but not on subsequent learning.

Most recently, Goo (2010) investigated the issue of reactivity in an examination of the relationship between working memory capacity and think-alouds, focusing on the issue of reactivity. His data evidenced that working memory capacity is related to reactivity because it interacts with the cognitive load consumed by verbalization. The think-aloud protocols were therefore associated with either more or less reactive effects on learner comprehension.

In conclusion, the problem of reactivity is inevitable when employing think-aloud or stimulated recall methods to measure ‘noticing’ (or the level of awareness involved in focal attention). Reactivity is also likely to play a role, albeit minor, in other research methodologies that investigate ‘noticing’ by learners, such as underlining, circling, or note taking. Thus, any research employing these methods to investigate ‘noticing’ needs to be interpreted with caution. The following section discusses a new method, eye-tracking, that is not belabored by reactivity problems.

Eye-Tracking as a Measurement of Noticing

The technique of eye-tracking—the registration and recording of eye-positions and eye-movement—is a suitable way to tap into learners cognitive and, particularly, attentional processes. While eye-tracking has been used since the mid-seventies in cognitive psychology and the cognitive sciences to investigate cognitive processes during scene perception, reading, and visual search (Rayner, 1998, 2007, 2009), it has only been recently that SLA researchers have utilized eye-movement recordings to study language learners’ ongoing cognitive processes during reading (for a review of syntactic processing studies, see Frenck-Mestre, 2005). “One great virtue of eye-movement data is that they give a good moment-to-moment indication of cognitive processes during reading” (Rayner, 2009, p. 1461). Studies investigating language (lexical, morphological, syntactic, and discourse) processing during reading collectively rely on the premise that there is a close link between overt attention (i.e., eye-location) and covert attention, the *mental* focus and processing of what is attended overtly (for a detailed discussion of overt and covert attention, see Wright & Ward, 2008). Godfroid (2010, p. 99) coined the term “eye-mind assumption” based on Just and Carpenter (1980) and Morrison (1984) who convincingly theorized that there is a tight link between the eyes and the mind: Our cognition

controls our eye-movement during reading. According to Godfroid, the ‘eye-mind assumption’ essentially corresponds to “what Reichle, Pollatsek and Rayner (2006) coined the ‘eye-mind link’ (p. 4) and what Engbert, Nuthman, Richter and Kliegl (2005) call the ‘immediacy-of-processing-assumption’ (p. 796)” (2010, p. 99). In conclusion, where an individuals’ eyes fixate most likely coincides with his or her focused attention.

Eye fixation times during reading can be influenced by any or all of the following variables: (1) frequency, (2) familiarity, (3), plausibility/predictability, (4) length, (5) number of meanings, (6) morphological complexity, (7) contextual constraint, and (8) age-of-acquisition (for overviews and reviews, see Liversedge & Findlay, 2000; Rayner, 1998, 2007, 2009; Starr & Rayner, 2001).²⁰ The most influential factors, the ‘big three’ predictors, are frequency, predictability, and word length (Kliegl, Grabner, Rolfs, & Engbert, 2004; Kliegl, Nuthmann, & Engbert, 2006), which are described below.

Word Frequency

Numerous studies have shown frequency effects on the different fixation time measures: low-frequency words (as determined based on corpus data) are usually fixated on longer as they are perceived as ‘difficult, and high-frequency words are usually fixated on less as they are perceived as ‘easy’ (see Rayner, 1998; Rayner, Sereno, & Raney, 1996; Reichle, Rayner & Pollatsek, 2003 for summaries). Interestingly, the frequency effect is attenuated when words are encountered several times in a reading passage. By the third encounter of a high or low frequency word, there is no difference between the two in terms of fixation times. That is, the

²⁰ For citations of all studies examining each of the variables, see Rayner, 2009, p. 1472.

durations of fixations on low frequency words decreases considerably with repetition²¹ (Rayner, Raney, & Pollatsek, 1995).

Disappearing text studies have corroborated the frequency effect: Even if a word disappears after 60 milliseconds, the eyes remain fixating on the empty space (K. Rayner, 2009; K. Rayner, Liversedge, et al., 2003; K. Rayner et al., 2006; K. Rayner, Liversedge, S.P., White, S. J., & Vergilino-Perez, D, 2003). How long the eyes keep fixating on the empty space is determined by the frequency of the word that disappeared. This phenomenon also supports the eye-mind assumption. Although the (low frequency) words that disappeared were most likely encoded within 60 milliseconds, they were probably not yet processed in the reader's mind, so the eyes remain on the empty space.²² According to Rayner (2009), this is “very compelling evidence that the cognitive processing associated with a fixated word is the engine driving the eyes through the text” (p. 1473).

Word Predictability

In addition to the frequency of a word, the predictability of a word in context influences the time needed to process that word. That is, the more predictable or plausible a word is given the textual context, the easier or faster the processing of the word. Thus, predictability is often defined in terms of the ‘contextual constraint’ (Frisson, Rayner, & Pickering, 2005, p. 862). To explain the issue of contextual constraint and predictability, I use an example from Frisson, Rayner and Pickering (2005): “For example, stamp is highly predictable in the (high-

²¹ Durations of fixations on high frequency words actually also decreases, but not as dramatically as for low frequency words.

²² This is yet another example that underscores the major flaws in design and in the operationalisation of ‘noticing’ in Smith’s (2010) study. Smith claimed that noticing happens when the eyes fixate on the target item 500 ms or longer, and no noticing occurs that if readers fixate for less than 500 ms on the target item. This is pure non-sense given the overwhelming eye-tracking literature that speaks of lexical processing within as little as 60msec.

constraining) context, *He mailed the letter without a _____*, but low predictable in the (low-constraining) context *He saw a beautiful _____*.” (p. 862).

A large amount of eye movement research has shown that high-predictable words are read faster than low-predictable words. This effect has been evidenced with different types of eye-fixations, including *first fixation duration*, *gaze duration*, and *total time*, the variables analyzed in this dissertation study. Most important for this dissertation study, word predictability has an effect on the *first fixation duration* (Altarriba, Kroll, Sholl, & Rayner, 1996; Rayner, Ashby, Pollatsek, & Reichle, 2004) as well as on *gaze duration*, the sum of all fixations on a word before fixating on another word (Rayner & Well, 1996) and on total time spent reading the target word (Calvo & Meseguer, 2002; Rayner & Well, 1996). More predictable words are also skipped more often compared to less predictable words (Altarriba et al., 1996; Brysbaert, Drieghe, & Vitu, 2005; Drieghe, Brysbaert, Desmet, & De Baecke, 2004) and individuals rarely fixate on them more than once (Calvo & Meseguer, 2002). The issue of predictability is important and I will return to it in the Discussion Chapter.

Word Length

The last influential factor as part of the ‘big three’ predictors is word length. Rayner (2009) simply describes this phenomenon: “As word length increases, the probability of fixating a word increases” (p. 1461). Just as the *predictability* of a word increases the probability of it being skipped, so does the *length* of the word (Brysbaert et al., 2005). Words composed of eight or more letters are rarely skipped, whereas shorter words are fixated on less frequently, only 25% of the time (Rayner, 2009, p. 1461, citing McConkie, Kerr, Reddix, Zola, & Jacobs, 1989; Vergilino & Beauvillain, 2000, 2001; Vergilino-Perez, Collins, & Dore-Mazars, 2004).

In light of this dissertation study, the appeal of eye-tracking is based on the relationship that has been found between eye movements and ongoing language processing during reading (Rayner, Liversedge, et al., 2003; Rayner et al., 2006). By means of eye-tracking, focal attention can be measured through an increase in fixation times, that is, longer gazes and more regressions (e.g., Godfroid, et al., 2010). Noticed, or more specifically, attended features (here verb stem-vowel changes) are likely to be marked by greater viewing activity (longer fixations and more regressions),²³ relative to unnoticed features. Thus, eye-tracking data yield information on the readers' (language learners') attentional foci and their general processing of written input (Rayner, 1998, 2007). Clahsen, Felser, Neubauer and Sato (2010) similarly emphasize the great benefit of eye-movement recordings. This method "can provide a window on how linguistic representations are constructed in real time during language comprehension and production and reduce the possibility of participants relying on their explicit or metalinguistic knowledge, compared to commonly used offline tasks" (p. 22). Thus, the great advantage of studying eye-movements of language learners is that compared to other online measurements of noticing (e.g., underlining, circling, or thinking aloud) there is no potential interference that may increase learners' cognitive load, distract them, or decrease what is being noticed. Furthermore, reading sentences or passages on the screen during an eye-tracking experiment is relatively natural (e.g., Duyck, Van Assche, Drieghe, & Hartsuiker, 2007). Thus, eye movement recordings have become very appealing for SLA researchers trying to tap into the learners' ongoing cognitive processes as they read the L2.

²³ Some caution is necessary for the interpretation of fixations on the irregular verb as evidence for attention. Words can be recognized and processed when the fixation lands few characters before the critical word, which provide a parafoveal preview of the target word (Rayner, 1998; Clifton, Staub and Rayner, 2007).

The next section briefly describes models of ocular movements as they relate to cognition: the E-Z reader model, followed by a brief description of the Saccade-generation with Inhibition by Foveal Targets (SWIFT) model. The E-Z Reader model has more relevance to this dissertation study than the SWIFT model, therefore the shortened explanation of the SWIFT model. Future elaboration of the *other* eye movement control models that have been proposed (i.e., several oculomotor and cognitive control models) would go well beyond the scope of this study and does not contribute to the theoretical foundation and/or the discussion appropriate for this study.

The E-Z Reader Model

The intuitive notion that readers move their eyes through a text to acquire information about its content inspired the establishment of the E-Z Reader Model. The E-Z Reader is a computational model that simulates eye-movement during reading. It works on the premise that cognitive processes associated with processing a fixated word serve as the engine driving the eyes through the text and that words are fixated in a serial fashion, word by word (a different view is presented by the SWIFT model). A wealth of literature has described and tested or used the E-Z reader model (most notably Pollatsek & Rayner, 2009; Rayner, Ashby, Pollatsek, & Reichle, 2004; Rayner, Pollatsek, & Reichle, 2003; Reichle et al., 2006).

According to the E-Z Reader Model, the identification of a word is completed in three stages: an early visual processing stage (V) that is pre-attentive, and later *lexical* processing which has two sub-stages. These two sub stages require different types of attention. There are two ‘events’ that occur during these later, attention-demanding stages: (a) a familiarity check

(level 1) that signals the oculomotor system (eye-muscles) to program a saccade²⁴ to the next word, and (b) the lexical access (level 2) that signals the reader's mental 'attention system' to move attention to the next word. It has been assumed that brain fires signals to move the eye (when needed) and also fires signals to attend to the input at the same time. Pollatsek, Reichle and Rayner (2006) claimed that the brain mediating a familiarity check triggers the programming of the saccade. Likewise, the completion of the lexical access might subsume the latency of an attention shift. "[T]he completion of lexical access is a point when processing of a word is complete enough so that the processing system can shift attention to the next word so that lexical processing of that word can begin with no "crosstalk" from the processing of the prior word" (Reichle et al., p. 11). Based on this account, Godfroid (2010) convincingly argued that "the dependence of the saccadic program from the successful completion of the lexical familiarity check initiates the *eye-mind link* in E-Z Reader" (p. 105, italics are mine). In summary, the E-Z Reader Model stipulates that an eye-movement is programmed or primed to occur when only part of the processing (i.e., familiarity check) has occurred. As such, the "time" (a few milliseconds) that elapses between the completion of the processing of the current word and movement of the eyes to the next word, is influenced by the completion of the familiarity check. If the next word is highly predictable by the sentence context, it might be skipped. Researchers concerned with eye-movements during reading usually view the E-Z Reader Model as a useful and appropriate way to account for and predict eye-movements by readers (e.g., Reichle et al., 2003).

²⁴ A saccade is the rapid movement of the pupil as it jumps from fixation on one point to another, for example in reading (Rayner, 1998).

While the E-Z Reader Model suggests that words are fixated on one at a time, the SWIFT model suggests that words are fixated on in a parallel fashion. For example, the characteristics of the following word influence the fixation time and processing of the currently fixated-on word; they are processed nearly simultaneously, in an analogous fashion. Studies evaluating both, the E-Z Reader Model as well as the SWIFT model (e.g., Kennedy & Pynte, 2008; Kennedy, Pynte, & Ducrot, 2002; White, 2008) revealed that both models are valid to some extent,²⁵ yet the E-Z Reader Model provides “a more transparent theoretical framework than SWIFT and has, importantly, been proven highly successful in accounting for well-established phenomena of eye-movement control” (Godfroid, 2010, p. 109-110).

The majority of SLA studies have used eye-tracking to investigate the effects of language transfer, either from (linguistic knowledge of) the L1 to the L2 or vice versa (N. Ellis and Sagarra, 2010; Roberts, Gullberg & Indefrey, 2008). The eye-tracking technique has also been adopted to study aspects of vocabulary acquisition, more specifically, idiom processing (Siyanova-Chanturia, Conklin & Schmitt, 2011) or the noticing of novel lexical items (Godfroid, 2010; Godfroid, Housen, & Boers, 2010b). Smith (2010) also used eye-tracking to explore the relationship between recasts, noticing, and performance during computer-mediated communication (CMC). However, there are substantial problems with the design of Smith’s study, particularly with the operationalization of noticing (i.e., fixations of longer than 500 millisecond = noticing; fixations shorter than 500 milliseconds = no noticing). Choosing a cut-

²⁵ The E-Z Reader Model has been criticized (1) as it does not take into account for sentence-level or discourse processing, which are also important in reading, (2) for the assumption of serial processing, and (3) for its reliance on word-frequency as a predictor for fixation length, as predictability might be even more important than frequency (see Rayner, 2009 for a review).

off point (500 milliseconds) is completely random and an inappropriate way of operationalizing the complex construct of noticing.²⁶

Only two studies (that did not employ eye-tracking) have investigated learners' noticing of (Spanish) verbs with a stem-vowel change (Leow, 1997, 2000). Thus, the investigation of attention to stem-vowel changes utilizing eye-tracking technology is a novelty in the field of SLA. Since a vast number of research studies has supported that noticing plays an important role in the acquisition process, this study explored (a) whether learning can occur based on a single exposure to a novel morphological structure,²⁷ and (b) how important the amount of attention is for possible learning gains. Godfroid (2010) inspired this research as she pursued an almost identical question (see also Godfroid, Housen & Boers, 2010). However, her eye-tracking research focused on vocabulary acquisition. She examined the processing of novel words in texts by advanced learners of English.²⁸ They read English paragraphs, half of which contained unknown (pseudo) words in different conditions. The words were either preceded or followed by a synonym, or they occurred without any contextual clues. The learners allocated more attention

²⁶ As stated by Godfroid (2010), "[t]he definition and operationalisation of noticing [...] requires a good understanding of its component processes, i.e. attention and awareness, the theoretical details which belong to the fields of cognitive psychology and neuroscience rather than to traditional language acquisition research" (pp. 177-178). The study of eye-movement in reading is also challenging and complex, as eye-fixations and movements are influenced by a large variety of factors. Smith (2010) appears to be inconsiderate of both, a) how attention and awareness have been operationalized and researched in cognitive psychology, and b) the wealth of eye-tracking research that informs future eye-tracking studies in the field of SLA.

²⁷ Note that the participants in this study were exposed to many exemplars illustrating the same underlying phenomenon (i.e., the stem-vowel change) but they only saw each exemplar once.

²⁸ This is an important difference to the current study. Godfroid's participants were English language majors at a university who had learned English in secondary school in Brussels or Flanders since the age of 13. In this study, the learners had been exposed to the target language, German, for only for 4 to 5 weeks, so they can be characterized as "true beginning learners."

to the novel words (as measured by eye fixation times), and this effect was enhanced when the context clarified the possible meaning of the word. The learners' total fixation time on a new word predicted vocabulary learning, as measured by an immediate recognition post-test. That is, the longer a participant looked at a word X during the reading task, the more likely he/she was to recognize word X in its original context on a subsequent test. The study showed that incidental noticing during reading might be *an important step* in advanced learners' integration of new words into their L2 mental lexicon.

To date, no eye-tracking research has investigated the noticing of verb morphology by German learners. Research into the attention to, and the processing and acquisition of verb morphology, such as vowel stem changes, will add much to the understanding of how adult learners attend to and become aware of irregularities that may or may not be a feature in their native language. As such, this dissertation project has theoretical and practical significance: It will inform SLA theory by using a new and empirically-based operationalisation of noticing, and it will inform foreign language pedagogy as this study will add to the knowledge of how learners of German process morphological irregularities when they first encounter them.

CHAPTER 3 GERMAN STEM-CHANGING VERBS

L2 Morphology and Its Noticeability

As explained in Chapter 2, the majority of noticing studies have addressed learners' noticing of specific syntactic, morphological, or morphosyntactical features (Alanen, 1995; Fotos, 1993; Hama & Leow, 2010; Hanaoka, 2007; Izumi & Bigelow, 2000; Izumi & Izumi, 2004; Jourdenais et al., 1995; Leow, 1997a, 1997b, 2000; A Mackey, 2006; Philp, 2003; Robinson, 1997a, 1997b; Rosa & Leow, 2004; Song & Suh, 2008). Two researchers (J. Williams, 2001; Hanaoka, 2007) looked through a broader lens and examined what learners naturally notice. J. Williams investigated whether learners initiate attention to form, and if so, what types of forms they notice. The results suggest that—though relatively infrequently—learners do attend to form. Learners' proficiency, the nature of the language activity and learners' associated objectives played an important role in learner-generated attention to form. More important, however, Williams found evidence that learner-generated attention to form is usually related to questions on lexis. "More than anything else, learners want to know about words" (J. Williams, 2001, p. 304). Although conducted in a different context, Hanaoka's (2007) study similarly found that Japanese learners of L2 English by and large attended to lexical features. The proficiency level of the learners also played a role: Proficient learners noticed more features than less proficient learners. The role of the proficiency level, or the developmental level of the learners, will resurface in the discussion chapter of this dissertation.

These two studies report that L2 learners primarily notice lexis as opposed to morphological features. Thus, this dissertation study builds upon these findings and investigates the extent to which beginning learners attend to content words, such as verbs, that also carry

morphological information. In this study, stem-vowel changes of irregular verbs in German constituted the morphological structure under investigation.

German Stem Changing Verbs

In German, several verbs change their stem-vowels (i.e., irregular verbs, also called “strong verbs”) in the second and third person singular, for example:

ich sehe → du **sieh**st (*I see, you see*)
ich sehe → er **sie**ht (*I see, he sees*)
ich esse → du **is**st (*I eat, you eat*)
ich fahre → er **fä**hrt (*I drive, he drives*)
ich laufe → er **lä**uft (*I run, he runs*)

Research investigating the acquisition of this irregularity has shown that adult learners, regardless of their proficiency level, have difficulty learning the inflectional morphology of verbs, particularly those with stem-vowel changes (e.g., Leow, 1997, 2000, for Spanish), which might be related to the fact that most stem-change patterns are lexically arbitrary and therefore unpredictable (Bybee & Newman, 1995). The acquisition of verbal morphology poses a challenge even to advanced learners of German (e.g., Clahsen, 1999; Neubauer & Clahsen, 2009; Prévost & White, 2000) learners of Spanish (VanPatten, Keating, & Leeser, in press), and learners of English (Haznedar, 2001; Lardiere, 2007). Just recently, in her 2010 review article on ‘L2 Morpheme Processing and Acquisition,’ Larsen-Freeman addressed the difficulty that (verb) morphology poses to language learners:

Learning inflectional morphology is a vexing problem for second language (L2) learners. Children acquiring their native language also experience some difficulty, which results in their committing overgeneralization errors. However, relatively quickly, children sort out the regulars from the irregulars and one allomorph from another. This is not the case for learners of L2s, at least not for older learners. Long after individuals have achieved a high level of proficiency in the L2, they are still plagued by uncertainty when it comes to grammatical inflections (Todeva, 2010), and their production in the L2 is still characterized by morphological omissions, commissions, and substitutions of one allomorph for another. (p. 221)

Larsen-Freeman then addressed the fundamental question that Pinker (1999) brought up over a decade ago: “Is language acquisition rule based or item based?” (p. 222); that is, in regard to this dissertation study, are irregular verbs learned as *items* or as *rules*? When learning irregular verbs, do learners use their declarative or their procedural memory?²⁹ Obviously, there is no ready answer to such a general question because the type of learning is likely to depend on the particular grammar structure in focus, the learners’ L1, learning contexts and situations, etc. In concluding her review article, Larsen-Freeman cautions that “we are susceptible of falling into the same trap again by using the acquisition of morphology to decide between a rule-based and associative learning account on SLA” (p. 228). Results from single experimental studies are not in the position of evidencing whether learners generally engage in item or rule-based learning. In this respect, with this dissertation study I do not attempt to probe into whether irregular verbs are learned item or rule based, but the study findings might allow me to consider and speculate on the design and possible results from future longitudinal research studies that investigate this issue.

Although descriptive linguistics has developed a system of rules for verb inflection and stem-vowel changes based on morphological and phonological regularities, it is not clear to what extent and when these regularities start to correspond to the L2 learners’ mental storage, processing, and production (Neubauer & Clahsen, 2009) and to what extent the learners’ attentional foci and levels of awareness relate to the acquisition of verb morphology.

Considering the relative importance of the reliable inflection of regular and irregular verbs

²⁹ Declarative knowledge corresponds to the storage of facts and events. In contrast, procedural knowledge corresponds knowing how to perform a certain task or completing an activity (Segalowitz, 2003).

coupled with the difficulty of the learning task, the question arises as to how L2 learners acquire correct verb inflection. In the context of *noticing* novel words or morphological features, no study has investigated whether learners of German who have not been formally introduced to these irregularities attend to these irregular features during reading. A recent review article by Clahsen, Felser, Neubauer and Sato (2010) synthesized the findings of different experiments that investigated native and non-native morphological processing. The overarching synthesis was that “L2 learners are less sensitive to morphological structure than native speakers and rely more on *lexical storage* than morphological parsing during processing” (p. 21, highlights are mine). If L2 learners are less sensitive to morphological structures, do they notice morphological changes, for example, in verbs? What exact roles do attention and awareness play in the acquisition of novel L2 forms, and do these cognitive mechanisms depend on the nature of the L2 structure? The following section addresses the latter question.

The Difficulty and Noticeability of L2 Forms

DeKeyser, in his 2005 review article, addressed the following important question: “What makes learning a second language grammar difficult?” He identified three major factors: (1) complexity of form, (2) complexity of meaning, and (3) the complexity of the form-meaning relationship. Regarding the third factor, DeKeyser argued “it is the transparency of form-meaning relationships to a learner who is processing a language for meaning that determines the difficulty of acquisition, at least for learners who are left to their own resources instead of presented with a reasonably complete set of rules about form-meaning relationships” (p. 3).

SLA researchers have debated connections between the difficulty level of the target language structure and the learner’s attentional resources (e.g., DeKeyser, 2005; Ellis, 1994; Gass et al., 2003; Long, 1996; Philp, 2003; Robinson, 2003; VanPatten, 1996).⁴ The attentional

resources consumed by the learner depend on the nature of the L2 task, because language tasks requiring more complex structures require greater processing (and thus greater processing capacities) than tasks requiring simple structures (e.g., McLaughlin, Rossman & McLeod, 1983; McLaughlin & Heredia, 1996). Complex structures may also capture more attentiveness by the learner and may therefore be more noticeable (White, 1998).

In the SLA literature there are two dominant views on how the difficulty level of the L2 task or the complexity of the target structure may influence the learners' attentional demands and noticing. One possible view is premised on the *Noticing Hypothesis* by Schmidt: the view is that the difficulty level of the structure may draw the learner's attention; that is, a more complex structure may have more saliency and, therefore, might more likely be noticed and learned (DeKeyser, 2005). Conversely, the higher complexity of a structure may increase its perceptual saliency or noticability and, therefore, stimulate processing for form, which may result in language learning gains (Skehan, 1998). Beyond that, the complexity of the L2 structure or task may not only increase its saliency, but may also promote the noticing of one's linguistic limitations (i.e., noticing the gap by the learner) (Gass & Varonis, 1994; Schmidt, 1990). From an information processing/cognitive theory of attention (e.g., McLaughlin et al., 1983; McLaughlin & Heredia, 1996), however, L2 learners may experience attentional constraints in specific processing conditions. In other words, if learners are limited-capacity processors (where their capacity for complex or difficult processing is limited), then attention to complex, relatively new structures may result in a cognitive overload that negatively impacts L2 gains. That is, if an L2 structure or L2 task is too complex, the attentional demands may result in a cognitive overload and impede learning processes. Izumi (2003) further elaborated on the interplay among cognitive overload, existing knowledge, and noticing during reading in that "the restricted L2

knowledge of the learners may make them rely on certain strategies (e.g., use of semantic information or contextual cues) more than others (e.g., syntactical cues)” (p. 184). Thus, such learners’ attention to form might be limited.

As described in Chapter 1, VanPatten (1996) also addressed the issue of attentional constraints within the notion of learners as limited-capacity processors and proposed that learners’ processing of the input depends on the attentional resources they currently have. Notwithstanding the criticism VanPatten’s IP Principles have received, they still made a significant contribution to the field. According to VanPatten (2004), processing is equalized with “making form-meaning/function connections during real time comprehension” and “an online phenomenon that takes place in working memory” (p. 4). As such, if learners are limited-capacity processors (as suggested by McLaughlin et al. 1983), their simultaneous attention to meaning and form may result in a cognitive overload. Thus, they may not attend to and notice a particular structure if they are processing the L2 input primarily for meaning.

This dissertation study has implications many of the above raised issues. It will provide a fine-grained examination of the role of focused attention in the learning of verb morphology. If learners pay attention to these features, do they use their selective attention? Do they rehearse a novel feature in working memory? To what extent is awareness involved in their focused attention? How well does eye-tracking methodology gauge learners’ L2 processing? Are beginning learners able to process sentences from the top down, or do they mainly rely on bottom-up processes? Does their learning of irregular verbs suggest item or rule-based learning? And last, does the particular grammar structure in focus determine whether learners pay attention to form or not? The dissertation study lends toward the answers to some pieces of information to

several, but possibly not all, of these questions, and the results from this study are likely to raise new questions and pave ways for new eye-tracking research agendas.

The strength and uniqueness of this dissertation study lies in the fact that it is the first eye-tracking study that has been conducted with true *beginning* adult learners. It is also the first eye-tracking study to investigate attention to verb morphology in the L2.

CHAPTER 4 METHODOLOGY

Research Questions and Hypotheses

This study was motivated by the following research questions and hypotheses.

1. Do adult beginning learners of German who are unfamiliar with stem-changing verbs pay attention to those irregularities during reading?
2. Is increased attention to irregular verbs associated with subsequent learning of them?

I addressed the first question via eye-tracking methodology (measurement of fixation times) and mixed factorial ANOVAs. I hypothesized that during the reading of simple sentences, learners would fixate on vowel-stem-changing verbs significantly longer than regular verbs in which the stem does not change. In other words, I expected to find a significant effect of verb-type (regularity) on reading. This prediction was based on previous research demonstrating that more visual attention is deployed on unknown words or structures (e.g., Godfroid, 2010) and on theories suggesting that eye-movements are governed by cognitive processing (Engbert et al., 2005; Just & Carpenter, 1980; Morrison, 1984; Rayner, 1998, 2007; Reichle et al., 2006).

I addressed the second question using pre- and post-production test gain scores and correlating gain scores and with total time gains. For this second question, I hypothesized that the eye fixation durations would not be predictive of the uptake of stem-changes in that there would be no statistically significant improvement of accuracy in producing (writing) stem-changing verbs after learners had been exposed to them through written input. The understanding of a partially rule-governed target structure (i.e., a vague underlying system) such as stem changes may not develop immediately but may develop gradually over time with exposure to several exemplars of the same underlying rule (e.g., N. Ellis, 2005). I predicted,

however, that learners would improve on the inflections (second person) and overall spelling of verbs due to the exposure to the verbs during the reading activity.

Participants.

Forty-seven adult learners of German enrolled in German 101 courses at Michigan State University (MSU) participated in the study on a volunteer basis.³⁰ Six German 101 courses were offered in the Fall semester of 2010. Every semester, students are enrolled in German 101 courses if they have no previous experience in German or receive a low score³¹ on the German Placement Test (administered at MSU). The German placement test is administered online by MSU's Center for Language Teaching Advancement. The test consists of 24 listening, 24 reading, 12 vocabulary, and 16 grammar items. The test items were developed by German department professors who based the design of the items on the ACTFL Proficiency Guidelines (www.actfl.org) and department course materials and instructional objectives. The test's reliability is only moderate, but the professors in the department consider the placement tests to be adequate for placing students into the courses. One limitation is that students may

³⁰ The data of 43 participants was included in the analysis of this study.

³¹ Based on personal communication (February 25, 2011) with Dr. Dan Reed, head of the language testing office at Michigan State University, the specific cut scores (for German 101, 102, 201, and 202) are regularly reviewed by the department and programmed into a system, so they are not easily accessible. The language test consists of 39 multiple-choice items that each student responds to (and these are selected from a pool, so not the same for each student). The 39-item test can be expected to have an average standard error of about 3. In 2007, Rasch-based item reliabilities averaged in the mid 70s, but the test has had some improvements since then, so the numbers are likely better to date. It is important to note that the multiple choice test is only part of the placement system, which begins with a background questionnaire and ends with the advisor making decisions based on all available information. So asking the reliability of the MC part is like asking about the reliability of a subset of the test and underestimates the reliability of the whole placement system. Nonetheless, the system as a whole works relatively well due to all the checks and balances (background information collected, MC results, essay results for higher-level placement, and advisor recommendations).

intentionally perform below the level of their true abilities in order to place into an “easier” course; it remains unknown how typical that is (Personal Communication with Tom Lovik, Department Chair, August 26, 2010).

Based on this information, I assumed that the 43 participants had comparable German proficiency. Based on consultations with the German course instructors and the German program coordinator, and based on my analysis of the Fall semester’s syllabus, I confirmed that the stem changing verbs would not have been introduced to the students, so none of the students had explicitly learned that stem-changes existed or how to use stem-changes accurately. The Teaching Assistants of the German 101 courses as well as the coordinator of the German undergraduate program supported my class visits for recruitment purposes.

The 43 participants consisted of 24 females and 19 males. Their age ranged from 17 to 36, with an average of 20 years. Thirty-eight participants reported that English was their native language, and the remaining five participants reported that their native language was Mandarin (one participant), Uzbek and Kazah (one participant), and two participants stated they were bilingual, one in French and English and another in Malay and English. Twenty-five participants stated they were of German heritage (but not German heritage-speakers); two reported having family members with whom they occasionally exchange words and phrases in German. I followed up by email and learned that these two learners’ German interactions were quite insignificant; for example, one stated, “my Grandma taught me to count in German.” Only one student reported having had prior German instruction in middle school, but he did not attend regularly, nor did he remember any German from his middle school time. Nine students reported that they had “some German” in high school. After a follow-up via phone and email, I learned that they had either dropped out of class early in the school year or “not remembered anything.”

None of the participants had received any private instruction or tutoring. Thirty-four participants indicated that they have studied other foreign languages; 24 had studied Spanish or a combination of Spanish and French, one student had also studied Thai and one student reported having studied Latin. Ten participants reported that they had been to Germany with the length of stay ranging from one to 14 days. The questionnaire inquired whether the participants had studied or interacted in German during their stay, which was unanimously negated. The students varied notably in terms their indication of weekly hours spent on homework assignments or just studying German: It ranged from zero to 10 hours per week with an average of 3.09 hours. The participants assessed their German proficiency in the four language skills (listening, speaking, reading, and writing). The descriptive statistics are summarized in Table 1.

Table 1: Descriptive Statistics

	Beginner	Low Int.	Intermediate	High Int.	Advanced
Listening	27*	11	5	0	0
Speaking	31	9	3	0	0
Reading	27	10	6	0	0
Writing	28	11	4	0	0

*Number of students who indicated the proficiency level on the background questionnaire.

Materials

The materials used in this study consisted of the background questionnaire (Appendix A), the pre- and the post-test, and the four counterbalanced versions of the reading task (i.e., sentences and pictures) on the eye-tracker screen and a vocabulary familiarity test.

The Pre- and Post-test

The pre- and the post-test were identical with the only difference being the order in which the pictures were presented to the students. Both tests required the participant to write down a

sentence using the action verb that was depicted on the picture. The exact procedures are described under the ‘procedures’ section.

The Reading Task

The target words (i.e., the regular and irregular verbs in the sentences) were elements of short, simple sentences that the participants were instructed to read for meaning. For a comprehensive list of all verbs, their inflection, and translation please see Appendix B. The learnability of the regular and irregular verbs used in this experiment was assumed to be equal, as the verbs have similar frequency (see Appendix D) and number of syllables, that is, exactly two. All verbs (regular and irregular) occurred as the second element of the sentence, which is the word order used in main clauses. The vocabulary familiarity test established that students had at least “seen the word” (i.e., the relevant verb) before.³² The sentences were also matched by number of words, with an exact average of 4.5 words per sentence. All stimuli sentences were designed with the learners’ proficiency level in mind, and they have also been shared with another experienced German foreign language teacher, who confirmed that the learners would not experience difficulty comprehending the gist or the overall meaning of the sentences. As the words used in the sentences were covered by the textbook by the time the data collection started, it was safe to assume that the vocabulary items used in the sentence were familiar to the participants. Every sentence was preceded by a fixation cross in front of the sentence.

I informed students that the study’s purpose was to investigate how beginning learners read from a computer screen, and this information may help language learning software developers such as *Rosetta Stone*. The reading task included two sentences, one baseline

³² Three verbs were not covered by the textbook, so the Teaching Assistants introduced the infinitive and the meaning of the verbs in their instruction.

sentence (starting with *ich* (*I*) = the first sentence, termed the ‘baseline sentence’) and a critical sentence (starting with ‘*du*’(*you*) or ‘*er*’ (*he*), ‘*sie*’ (*she*) = the second sentence termed the ‘critical sentence’) which used the same verb but with different inflections due to the first and second/third person difference. In the experiment condition the difference was not only established by inflection, but also by the stem-vowel change; thus, the term *critical* condition.

The participants worked through 24 verbs and saw three screens per verb (trial). On the first screen of each trail, participants saw a picture of an action verb. When they clicked to the next screen, they saw two sentences, which they read. On he third screen, participants saw three pictures, one of which, again, depicted the action verb of that trail. Then, the participants looked at a fixation cross (drift correct) before they proceeded to the next trial, which again contained three screens. On the whole, the participants saw 12 experimental trials (irregular verbs) and 12 control trials (regular verbs).

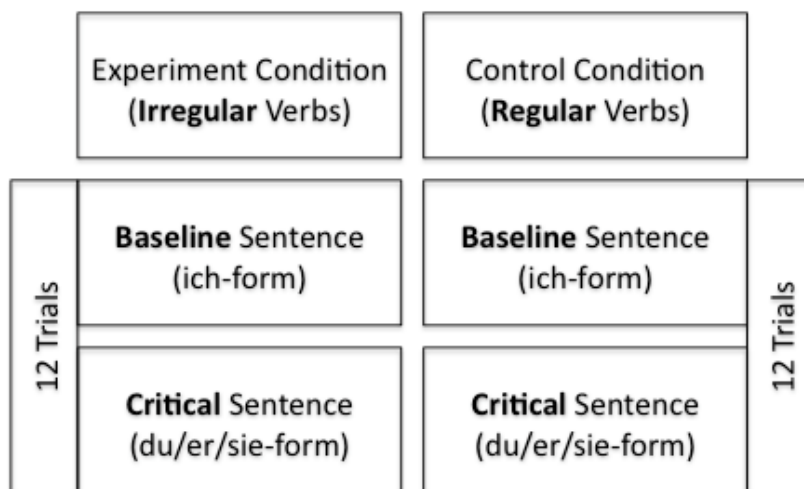


Figure 7. Simple Illustration of the Experimental Design

The purpose of the fixation cross between the trials was to perform a drift correct, which provides additional security to avoid a drift in the gaze position. Although the calibration procedure at the beginning of the experiment notices any drifting problems, additional drift corrections via fixation-crosses add an additional safety measure.

Another type of fixation-cross appeared right at the beginning of the two sentences. Before the participants read the sentence, they briefly fixated the cross. This ensured an additional recalibration of their eye fixation coordinates on the screen. The sentences appeared in Tahoma 16-point font on a Power Point slide in presentation mode with a subtle yellow fading to blue background (for eye comfort). The font was kept at a regular size to make the reading exercise appear as natural as possible. A picture showing the action (e.g., a woman running) preceded the screen that showed the two sentences (i.e., baseline and critical sentence) using the action verb. The purpose of showing a picture depicting the action was to ease the overall reading task for the beginning learners. Following the reading of the two sentences, the third screen showed three pictures. One of those pictures, again, depicted the action verb described by the sentences shown on the previous screen (but this picture was different from the initial picture depicting the action verb). The other two pictures depicted other actions. The participants were instructed to look at the correct picture, that is, the picture that corresponded to the preceding sentences. Figure 8 illustrates what a trial looked like on the eye-tracker screen.

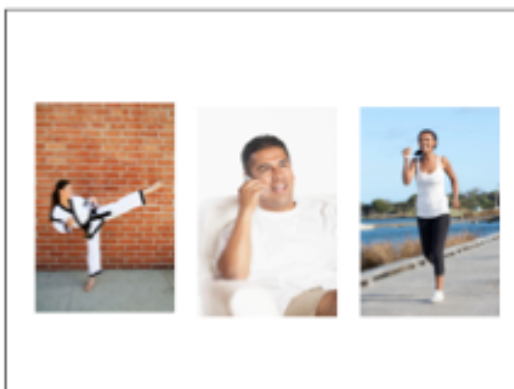
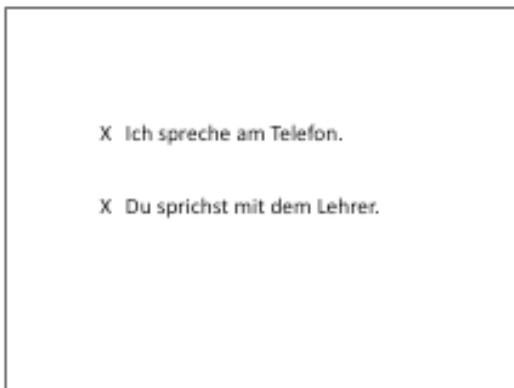


Figure 8. Example of Eye-tracker Screens

For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation.

The stimuli sentences that occurred in the experiment were all counterbalanced by sentence endings and by the person ('du' or 'er'/'sie') used for the critical condition. That is, all sentences occurred with either stimuli endings A or with stimuli ending B and they contained either the 2nd person or the 3rd person singular. Table 2 below illustrates the counterbalancing of sentence endings and persons:

Table 2: Counterbalancing of Sentence Endings and Persons

Infinitive	Sehen	Sprechen	Essen
Group I:	Ich sehe das Meer. <i>I see the ocean.</i> Du siehst die Natur. <i>You see nature.</i>	Ich spreche mit dem Lehrer. <i>I speak with the teacher.</i> Du sprichst am Telefon. <i>You speak on the telephone.</i>	Ich esse gern Brot. <i>I like to eat bread.</i> Du isst oft Müsli. <i>You often eat cereal.</i>
Group II:	Ich sehe die Natur. Du siehst das Meer.	Ich spreche am Telefon. Du sprichst mit dem Lehrer.	Ich esse oft Müsli. Du isst gern Brot.
Group III:	Ich sehe das Meer. Er sieht die Natur.	Ich spreche mit dem Lehrer. Er spricht am Telefon.	Ich esse gern Brot. Sie isst oft Müsli.
Group IV:	Ich sehe die Natur. Sie sieht das Meer.	Ich spreche am Telefon. Sie spricht mit dem Lehrer.	Ich esse oft Müsli. Er isst gern Brot.
Second person stimuli ending A			
Second person stimuli ending B			
Third person stimuli ending A			
Third person stimuli ending B			

Recall from Chapter 2 that fixation times are likely to be influenced by the words after the target word (Kliegl, Nuthmann and Engbert, 2006). Therefore, I controlled for that by counterbalancing the sentence endings across the groups to which the participants were assigned. Note that these groups did not constitute experimental groups and control groups; rather, they dictated what counterbalancing condition the participant would be assigned to.

Due to the counterbalancing of stimuli endings across groups, I assumed that the experimental variance was not influenced by the words that followed the target word. In other words, any sentence ending's effect on processing time was controlled. During the actual experiment, the stimuli were presented in a fully randomized order, that is, a random mix of trails containing regular or irregular verbs. However, each experiment included all 24 trials (i.e., all 24 possible verbs).

The Vocabulary Familiarity Test

The vocabulary familiarity test included all words that appeared in the reading task. The participants indicated on a scale from one (*don't know the word*) to seven (*know the word and its meaning*).

Procedures

To ensure that the participants had not been formally introduced to the phenomenon of stem-changing verbs in German, all data were collected early in the fall semester, specifically, during the fourth and fifth week of the semester. According to the German 101 syllabus and the coordinator of the German undergraduate program, the stem-changing verbs were not formally introduced to the students until the end of the 6th week of the fall semester.

Participation in the study was voluntary. I recruited the participants by briefly visiting the six 101 German classes on the first days of classes (September 3rd and 4th) to inform students about the opportunity to participate in the study. Without revealing the actual research goals and purpose of the study, I informed the students about the procedures of the study and the risks and benefits involved. Students who were interested in participating were asked to document their

names and email addresses on a sign-up sheet so I could subsequently email them to confirm their appointment and provide guidance to the location of the research room.

The German 101 volunteers participated individually in a quiet room (the eye-tracking research laboratory) for about one hour, depending on the participant's pace of completing the research tasks. Before the research procedure started, the participants carefully read and signed the consent form. Each participant was informed of the sequence of the activities prior to completing the pre-test, the reading activity, and the post-test. As the first step in the research procedures, the participants completed a background questionnaire that elicited information about the participants' demography, language learning background, prior German exposure, and a self-assessment of their German proficiency.

Following the completion of the background questionnaire, the participants took the pre-test. As the entire experiment followed a pre- and posttest design, the purpose of the pre-test was to establish a baseline against which the post-test scores can be assessed. To avoid test anxiety, the participants were told that the pre-test was a warm-up activity, but that they should do their very best. A typical testing session proceeded as follows: I sat down with the participant and I had a stack of 14 pictures illustrating 14 verbs, 12 irregular verbs that are used in the experiment as well as two regular verbs (distractors). The participant had a sheet of paper with 14 lines and a pencil in front of him or her. Then I held up a picture and asked the participant in English what the man or the woman in the picture was doing. If the picture showed a woman talking on the phone, the participants usually responded: "She is talking on the phone." After positive acknowledgement I encouraged the student to say the sentence in German. Typically, the student then said "Sie spricht am/in/auf Telefon" (sic.). I asked the student to write down the sentence on the paper. I also informed them that I wouldn't provide them with any feedback on

whether their sentences were correct or incorrect, but I encouraged them to write down the sentence just as they said it. The whole procedure was repeated until all fourteen pictures had been shown. The advantage of this form of pre- and post-testing lies in the vagueness of the testing purpose. That is, the participants did not know whether the test was targeting grammar, vocabulary items, pronunciation, or their ability to form simple sentences. Thus, the participants did not get the idea that I was looking for their use of stem-changes in inflected verbs. There was no time pressure to complete the pre- or posttest; the time spent on the pre- or posttest ranged from ten minutes to eighteen minutes.

Upon students' completion of the pre-tests, the participants sat in front of a desk with a computer screen approximately 80 cm in front of them. The participants placed their chin on a forehead rest (head supported as opposed to remote/head free). A camera attached to the computer screen captured the participants' pupils' movements as they read the sentences and looked at pictures presented on the screen. The participants had a game controller in their hands that allowed them to click through the slides at their own pace. The experiment session began with a short built-in calibration exercise on the computer screen. To explain the calibration exercise, the participants looked at a moving dot presented on different parts of the computer screen to calibrate the Eye-Link camera. Godfroid (2010) explains the calibration exercise precisely: The calibration "adjust(s) the gaze position measurement for data collection. During calibration of the EyeLink II, a participant is required to fixate on [...] points that appear at various locations on the screen. Thus, the system knows a participant's exact gaze position on the screen and can adjust the pupil position as measured by the cameras to this benchmark. Hence, calibration is the process of adjusting the eye gaze measurement with a view to subsequent data collection" (p. 133).

I explained some practical aspects of the experiment, for example how to use the game controller to move from one screen to the next, how to successfully complete the calibration exercise, and that they had to focus their eyes on the fixation cross (drift correct) every three slides and wait for me to confirm the accurate fixation (by pressing a button on a separate keyboard).

After every trial (i.e., three slides: one with one picture, one with two sentences, and one slide with three pictures) the participants looked at a fixation cross, also called a *drift correct*, that ensured that the camera still accurately captured the participants' eye-movements. The entire reading task took about fifteen minutes.

As a last step in the experiment, the participants completed the post-test, which was the same as the pre-test, with the only difference being the order in which the pictures were presented. The purpose of the pre- and posttests was to determine whether students had improved their writing of stem-changing verbs and to see whether there is a possible correlation between the time spent looking at the target words and scores tests. After completion of all research tasks, the participants were paid \$10.00 cash.

The German Teaching Assistants administered the vocabulary familiarity test in the German classes on the last day of the data collection period. All students in their classes took the test, even if they hadn't participated in the research. The reason behind administering the test in the German classes as opposed to integrating it into the experimental procedures was due to logistic reasons, mainly time limits. Completion of the vocabulary familiarity test took up to twenty minutes. I excluded the students from the test pool who did not volunteer in the experiment.

Analyses

Overall, the data consisted of the gain scores from pre- to post-test, the participants' eye-movements and fixations as recorded during the reading task, and the vocabulary familiarity test. As a reminder, *noticing* was measured through the participants' eye-movements and fixations, whereas *language uptake* was measured by means of the pre- and post-tests.

Analysis of the Eye-tracking Data

The stem-vowel changes of the German verbs constituted the independent variable (i.e., target words at two conditions: regular (baseline) and irregular (critical) in this experiment. The time spent by the participants looking at the target words (various types of fixations) was defined as the dependent variable. Please refer to the different types of fixation measures listed in Table 3.

Table 3: Variables in the Interest Area Report

Dependent Variables	Description
First Fixation Duration	Duration of the <u>first</u> fixation event that was within the interest area.
Gaze Duration (First Run Dwell Time)	Dwell time (i.e., summation of the durations across <u>all</u> fixations in the current interest areas)
Total Time (Dwell Time)	The summation of the duration across <u>all</u> fixations on the interest area

Based on the insight provided by the literature on eye-movements during reading (e.g., Rayner, 1998), I mainly considered the First Fixation duration, Gaze Duration, and the Total Time, which have been mainly used in the cognitive science of reading research (Staub & Rayner, 2007) and are therefore likely to provide an accurate reflection of how much attention and cognitive processes occurred upon processing the target word. First fixation duration is the duration of the first fixation on a word, regardless of whether this word has been fixated a

number of times. Gaze duration is the sum of all fixations on a word *during the first pass*, and total time is the summation of the duration across all fixations, regardless of whether it was the first or the second pass. To clarify the distinction of these measures, I am citing French-Mestre (2005), who explains:

We can distinguish between ‘first pass’ reading time, ‘second pass’ reading time and ‘total’ reading time (the sum of the former two measures), for any given region of interest (ROI). First pass reading can furthermore be broken into ‘first fixation’, i.e., the first time the eyes land in a given region of interest (ROI), and ‘gaze duration’, i.e., all fixations in the ROI from the first fixation until the eyes exit to the right or left” (p. 176).

To answer research question 1 (Do adult beginning learners of German who are unfamiliar with stem-changing verbs pay attention to those irregularities during reading?) I first computed the differences in the various types of fixation times between a target word (i.e., irregular verb) and a control word (i.e., regular verb). The calculated differences—the fixation gain—established the dependent variables. I then conducted a mixed design factorial ANOVA as I had two independent variables (fixed factors). According to Field (2009), a Factorial ANOVA is very powerful because it allows us to assess the effects of each independent variable, plus the effects of the interaction. The first independent variable (i.e., fixed factor) was established by the regularity of the verb at three levels, or two levels, depending on the target of the analysis (i.e., the type of stem-change as opposed to irregularity in general).

1. The vowel change from e to ie/i, which received the value 1:
essen (du isst; er/sie isst)
geben (du gibst; er/sie gibt)
lesen (du liest; er/sie gibt)
nehmen (du nimmst; er/sie nimmt)
sehen (du siehst; er/sie sieht)
2. The vowel change from a to ä, which received the value 2 (or also 1, if I was looking at irregularity in general):

fahren (du fährst; er/sie fährt)
fangen (du fängst; er/sie fängt)
schlafen (du schläfst; er/sie schläft)
tragen (du trägst; er/sie schläft)
waschen (du wäschst; er/sie wäscht)

3. No vowel change (simply regular), which received the value 3:

kochen (du kochst; er/sie kocht)
segeln (du segelst; er/sie segelt)
studieren (du studierst; er/sie studiert)
trinken (du trinkst; er/sie trinkt)
wandern (du wanderst; er/sie wandert)
zeigen (du zeigst; er/sie zeigt)

4. No vowel change (simply regular), which also received the value 3:

begrüßen (du begrüßt; er/sie begrüßt)
gehen (du gehst; er/sie geht)
kaufen (du kaufst; er/sie kauft)
lachen (du lachst; er/sie lacht)
schreiben (du schreibst; er/sie schreibt)
singen (du singst; er/sie singt)

The second independent variable (i.e., fixed factor) was the group (i.e., one of the four counterbalancing conditions) to which each subject was assigned. The factor subject itself, although designated as a random factor, was nested within the group. According to Sall, Creighton and Lehman (2006), “nested effects occur when a term is only meaningful in the context of another term, and thus is kind of a combined main effect and interaction with term within which it is nested” (p. 336). All statistical tests were calculated using the IBM Statistical Package for the Social Sciences (SPSS), version 19.

Analysis of the Test Scores

To answer research question two (Is increased attention to irregular verbs associated with subsequent uptake of them?) I conducted correlation analyses were I investigated whether increased first fixation times (i.e., total time gains) correlate with gain scores on the corresponding test items. Only the target words (regular and irregular verbs) were considered for

scoring, all other words written by the participants in the sentences were ignored. The tests were scored twice with two different scoring procedures. With the first scoring procedure I took a holistic approach and scored the overall improvement by the participants. With the second scoring procedure, I only gave points for improvements on the vowel-stem change feature. The two scoring procedures are described below, starting with the first, the holistic approach.

A participant received two points if he or she provided the correct form of the verb, including correct inflection and the correct stem-vowel change. A verb was conjugated correctly if all letters were present (including umlauts: ä, ö, ü) and in the correct order, I did not count as incorrect any typographical/spelling errors, e.g., *du seihst* instead of *du siehst*. I gave one point if a verb was conjugated correctly but the stem-vowel change was incorrect or simply disregarded (e.g., *er 'sprecht' (he speaks)* instead of *er 'spricht'*). To give an example, if a learner wrote '*er sprecht*' on the pre-test and '*er spricht*' on the posttest, he or she received one point on the pre-test and two points on the post-test. I gave zero points if a learner failed to conjugate a verb correctly (e.g., '*er sprechen*' or '*er ist sprechen*'). If the improvement on the posttest laid in the correct verb inflection (i.e., a change from '*er sprechen*' to '*er sprecht*'), I gave half a point, regardless of the missing stem-vowel change. I tabulated all verbs and scores and calculated the difference in scores (i.e., gain score) for each item (verb) between the pre-and the posttest.

The second scoring procedure disregarded any improvements on inflection or spelling; I paid attention to vowel changes only. If a participant did not change the stem-vowel correctly, I gave zero points. Only if the stem-vowel was changed correctly (with the exception of typographical/spelling errors, e.g., *du seihst* instead of *du siehst*) I gave two points.

The next chapter shows the findings of the statistical analyses employed to address the two research questions.

CHAPTER 5 RESULTS

The results presented in this chapter are organized by the dependent variables, starting with total time, gaze duration, and first fixation duration. Therefore, I first look at the sum of all fixation durations on a word (i.e., total time) before looking at smaller fixation time measures (gaze duration = the summation of the durations across all fixations in the current interest areas during the first pass reading, and first fixation duration = only the first fixation on a word). The main ANOVA design was nested with two crossed-fixed factors (verb type and group) and participants nested as a random factor within group. Recall from Chapter 4 that a nested ANOVA can be thought of as an extension of a one-way ANOVA in which each group is divided into subgroups. Specifically, the data stemmed from 43 participants, and I assigned those participants to groups, each of which read the sentences with a particular sentence ending and subject person (2nd or 3rd person singular) combination. Figure 9 below illustrates the nested design.

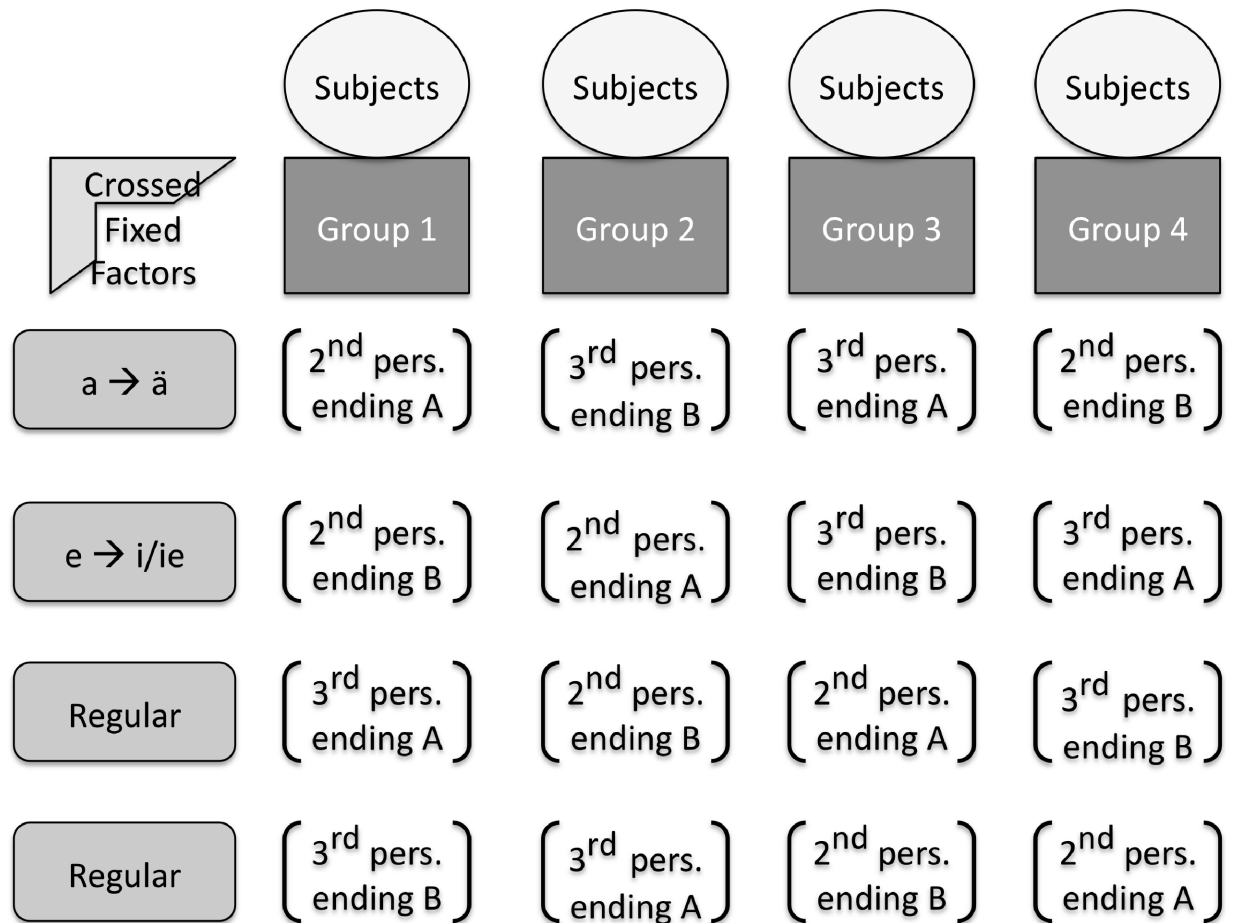


Figure 9. Nested Design

After presenting results pertaining to the three types of fixation times, the observation data collected for this study, I present the experimental results for learning that occurred as a result of the participants' attention to the stimuli. These data consist of correlations between pretest-posttest gains and total time gains (total time of the critical condition minus the baseline total time). This correlation analysis addresses my second research question—whether students' attention to irregular verbs is associated with their subsequent uptake of them.

Results Addressing RQ1

Total Time

My preliminary assessment of the descriptive statistics in the nested ANOVA revealed the existence of outliers that may bias parameter estimation and, by implication, statistical inference (Field, 2009). To avoid any bias in estimation, I followed several steps. The first step was to remove influential outliers according to Cook's distance test (Cook, 1977). Cook's distance is a measure of the overall influence of an observation on the model (Maindonald, 2010). However, the removal of outliers based on Cook's distance test did not normalize the residuals associated with the model. As a second approach, I removed all observations whose corresponding standardized residuals were larger than 3.5 or smaller than -3.5.³³ Both methods (Cook's distance and the removal of large standardized residuals) still failed to ensure normality of the model's residuals. In the fields of second language studies and bilingualism, researchers often omit observations that fall outside of two standard deviations from the item mean (Hoover & Dwivedi, 1998; Irena O'Brien, 2007; Lotto & De Groot, 1998; Sherman, 2007; van Heuven, Dijkstra, & Grainger, 1998), which make up about 5% of the data points in a normally distributed data set. In other words, researchers usually keep those values that lie within two standard deviations from the mean. Such values make up a confidence interval, which is a range of values that are assumed to contain, with a certain probability, usually 95% of the true population mean (Field, 2009). In this study, I worked with a 93.4 % confidence interval, which ensured the normality of residuals, and thereby a fundamental assumption to underlying ANOVA models. This means, with 93.4% confidence, the general targeted population is likely to share the characteristics of the study participants (e.g., Bird, 2004).

³³ According to Meyers, Gamst and Guarino (2006) "A traditional remedy for statistical assumption violation is to eliminate cases with large residual values (those with absolute standardized residual value greater than 3). This elimination can produce a better fit by the regression equation" (p. 203).

The influential outliers that were removed were cases of long fixation times (usually around 2000 milliseconds or more). Given the large number of unique data points in this study (n=1,032), the removal of 84 records still left me with 948 records for analysis. It is important to note that outliers are defined as observations that appear to be inconsistent with other observations in the data set: outliers have low probability to originate from the same statistical distribution as the other observations in the data set (Field, 2009; Sall, Creighton, & Lehman, 2005). Since 84 cases in this study revealed extremely long fixation times, they might be fundamentally different in terms of the type and amount of processing they engage in when reading stem changing verbs. Because I operationalized noticing as an increase in fixation time (relative to a baseline condition), I decided to investigate the outliers in a separate analysis (see section x). Below I present data with the 84 outliers removed.

Total time ANOVA

Since there was no main effect of group, I conducted a nested ANOVA with the fixed factor *verb type* crossed with the random factor subject/group (subject nested within group) with the trimmed data set to model the main effects of verb type (the independent variable) on total time (the dependent variable). Recall that *total time* here refers to ‘total time gain’, which is the baseline-critical-item differential; in other words, for each individual, it is his or her average fixation time on the baseline condition minus his or her average fixation time on the critical condition. If total time gain has a *negative* value, irregularity (stem-vowel change) did not attract longer fixations in comparison to the baseline *ich*-form (regular, no vowel change). Analogous, if the total time gain has a positive value, then the irregularity attracted longer eye-fixations or processing time compared to the baseline *ich*-form.

As expected, there was a significant effect of verb type, $F(3, 902) = 4.278, p = 0.005$,

$\eta^2_p = 0.014$, and a significant effect of subject $F(42, 902) = 2.719, p = 0.032$. The latter accounted for a great deal of the variance in total time gain, as indicated by the effect size

(η^2_p) of 0.112. Thus, 11.2% of between-subject variance in total time gain was accounted for by individual differences in reading speed among the participants.

Using a Multivariate ANOVA for these data was appropriate, as judged by the F test of

goodness of fit ($p > .05$): $F(126, 776) = 1.102, p = 0.226$ and $\eta^2_p = 0.152$. The latter indicates that 15.2% of the variation in total time gain was accounted for by this model, which is considered to be large (Field, 2009).³⁴

Post Hoc Tests

The Tukey HSD *post hoc* test revealed that the participants' average total fixation time on verbs *changing their stem-vowel from a → ä* was significantly longer than their average total fixation time on regular verbs *with no vowel change* ($p = .035$). There was no significant

³⁴ The result of the model indicated a non-significant intercept. That is, the value at which the line crosses the y-axis does not differ significantly from zero; more precisely, the total fixation time predicted by the model when the values of all independent variables are set to zero does not differ significantly from zero. I kept the intercept in the model as (1) no-intercept models do not adjust for the average of the independent variables, which leads to interpretation difficulties, and (2) no-intercept models should be fit only when theoretical justification exists and the data appear to fit a no-intercept framework better (Field, 2000 pp. 144 -145). The latter does not apply to this eye-movement study, as the models preserve the same magnitude of significance for the independent variables.

difference between participants' average total fixation time on verbs changing their stem-vowel from $e \rightarrow ie$ and regular verbs ($p = .117$). Figure 10 below illustrates such differentials.

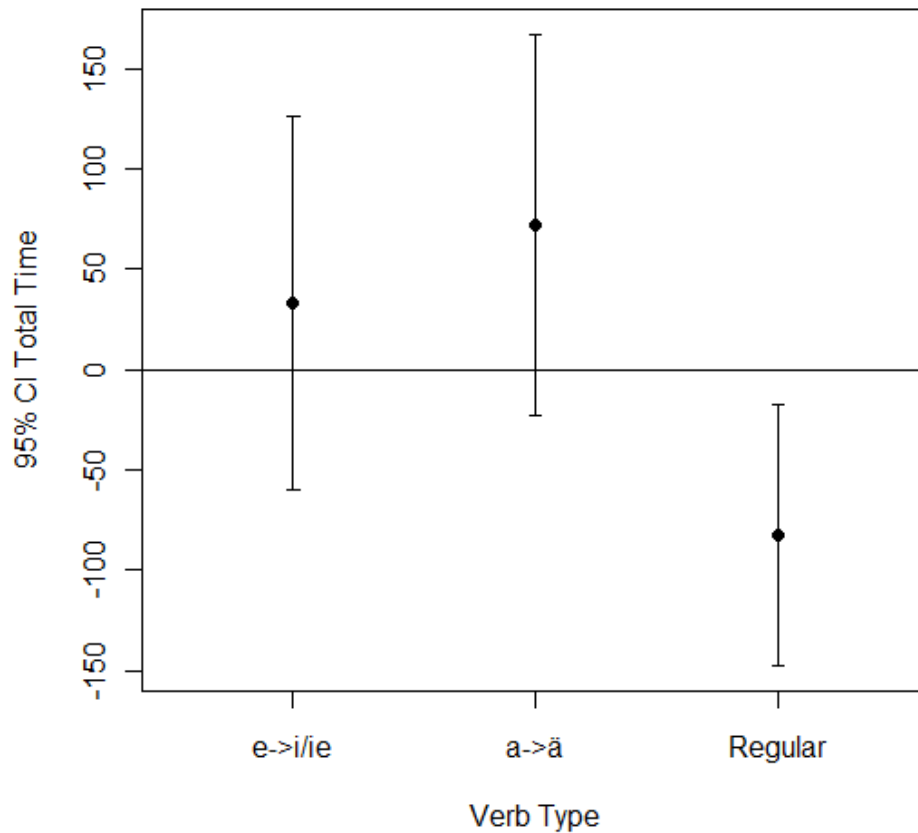


Figure 10. Post-Hoc Results

Pair-wise Comparisons

Because there was a significant difference between the stem-vowel-change $a \rightarrow \ddot{a}$, I was interested to see whether there was a significant difference in total time gain between regular verbs on the one hand and irregular verbs on the other. The pair-wise comparison of irregular versus regular verbs (using *verbcoding2*) showed that the participants' average total fixation time on *irregular verbs* was significantly higher than their average total fixation time on *regular*

verbs. Using a Multivariate ANOVA for these data was appropriate, as judged by the test of

goodness of fit ($p = .005$): $F(42, 862) = 1.226$, $p = 0.157$ and $\eta^2_p = 0.056$. The latter indicates that only 5.6% of the variation in total time gain was accounted for by this model, which is considered to be moderate (Field, 2009).

Nested ANOVA Assumptions

The assumptions for Mixed models include: (1) linearity of model, (2) independence of observations, (3) residual normality of the independent variables' groups, (4) homogeneity of residuals variances of the independent variables' groups, (5) lack of multicollinearity, and (6) lack of influential outliers to produce unbiased parameter estimation, and be able to generalize the inference from the results to the entire population of interest (Leon-Guerrero & Frankfort-Nachmias, 2010; Rutherford, 2001; Smith, Gratz, & Bousquet, 2008). I satisfied the first and second assumption by my modeling and the research design. The fifth assumption was not applicable, as I did not have continuous predictor variables. The lack of influential outliers was ensured due to the trimmed data set (see explanation above). Therefore, I turned my attention to examine assumptions 4 (homogeneity of residuals variances), and 3 (normality of the residuals).

Homogeneity of Residuals Variances

Homogeneity of error variance is one assumption of the fitted mixed models which implies that the standard deviations of the error terms are constant across the levels of the independent variables and do not depend on the x-value (Field, 2009). For the ANOVA model for total time, the probability associated with Levene's test for equality of variances ($F(171, 776) = 1.779$, $p = 0.01$) was less than the .05 significance level, so the assumption of equal variances was not satisfied. It is important to note, however, that factorial analyses are not sensitive to

violations of the equal variance assumption when samples are moderate to large and samples are approximately of equal size (Field, 2009), which is the case in this study.

Normality of the Residuals

Visual inspection of the histograms and QQ-Plots of the standardized residuals suggested that the residuals for the different levels of verb types (verb coding2) were normally distributed. Furthermore, the formal tests of normality (Kolmogorov-Smirnov and Shapiro-Wilk) indicated that the total time data was normally distributed among the categories of the independent variables ($\alpha = 0.001$). The data appeared to be normally distributed since the histograms followed approximately bell-shaped curves and the QQ-Plots had diagonal lines; the data did not appear to have any nonlinear patterns. Based on these observations I concluded that residual normality was established (Field, 2009 p. 148).

Analysis of Omitted Outliers

To analyze the characteristics of the omitted observations from the study, I conducted a Chi-square test of independence. Specifically, by means of the Chi-square I analyzed whether observations of extreme long fixation times occurred more often with irregular verbs than with regular verbs than might be expected due to chance. An alpha level of $\alpha = .05$ was set for all Chi-squares. A Chi-square test measures “the difference between a statistically generated expected result and an actual result to see if there is a statistically significant difference between them” (Cohen, Manion, & Morrison, 2007, p. 562). I conducted the first Chi-square test with irregular versus regular verbs and positive values versus negative values. It is important to understand that positive and negative values were computed with reference to the baseline sentence (1st person singular = *ich*, where the verb-stem never changed). Recall that I subtracted

the fixation time on the verb in the critical sentence from the fixation time on the verb in the baseline sentence. Thus, negative values indicated longer fixations on the baseline condition compared to the critical condition, which can tentatively be interpreted as “no noticing.” Positive values indicated longer fixations on the critical condition compared to the baseline condition, which can tentatively be interpreted as “noticing.” The hypotheses in consideration of the Chi-squares were as follows:

H₀: The verb types (regular versus irregular) and fixation the times on the critical verb are independent.

This means that the outlier values for a given verb type (regular versus irregular) are evenly distributed among positive values (noticing) and negative values (no noticing).

H₁: The verb types (regular versus irregular) and fixation the times on the critical verb are dependent.

This means that the outlier values for a given verb type (regular versus irregular) are not evenly distributed among positive values (noticing”) and negative values (no noticing).

The results of the Chi-square test indicated that the total fixation time (*posneg*) was significantly associated with verb-type (irregular versus regular), $\chi^2 = 5.098, df = 1, p = 0.015, OR = 0.309$ with 95% C.I. (0.117, 0.812). As shown in Figure 11, for the majority of the outlier observations, fixation times were significantly longer on irregular verbs as opposed to regular verbs (both compared to the baseline condition).

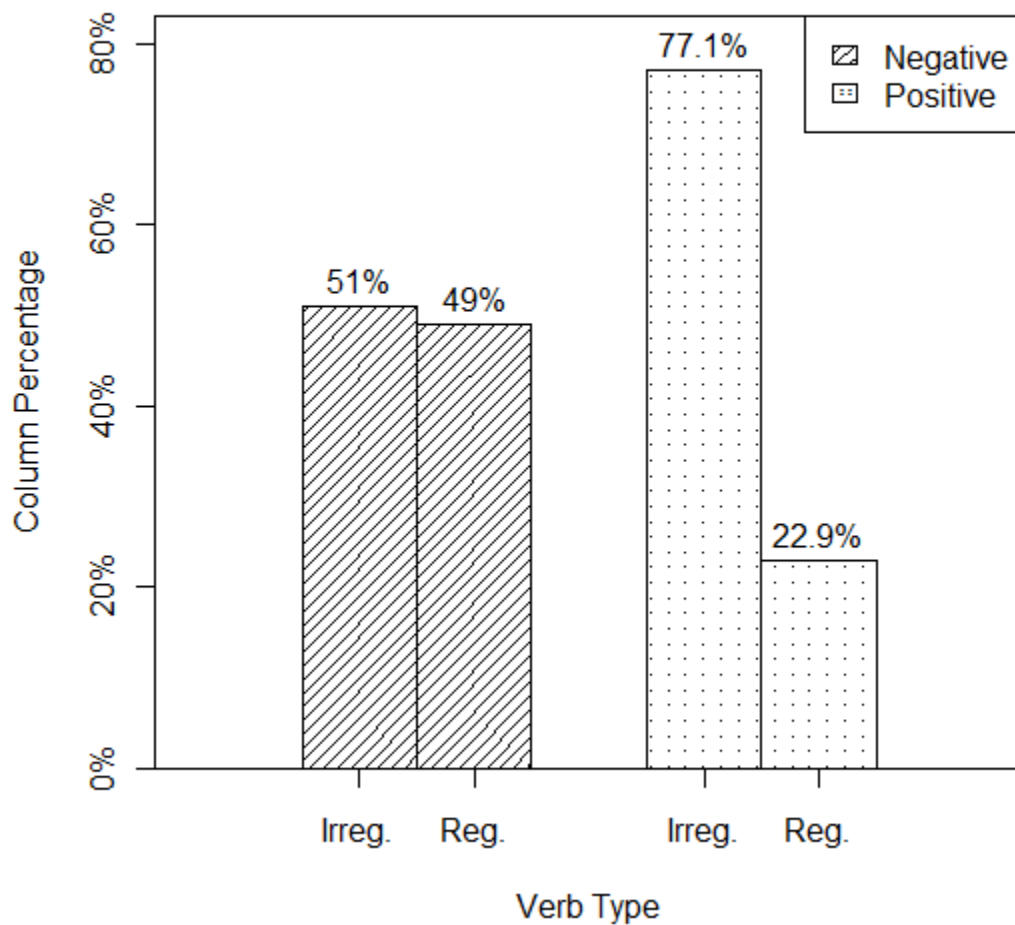


Figure 11. Verb Type and Total Time Gain (Regular versus Irregular)

Figure 12 shows that 77% of the outlier cases revealed extremely long fixation times on vowel-changing verbs (compared to the baseline condition), whereas only 22.9% of the outlier cases revealed longer fixation times (compared to the baseline condition) on regular verbs. This notable difference suggests that the outlier observations demonstrate noticing (i.e., longer fixations) of the irregularities as opposed to regular verbs.

To quantify the association between verb type and total time, I used the Mantel-Hänszel estimator as a common odds ratio for the categories of sign of total time. Hosmer and Lemeshow (2000) explained “The odds ratio [...] approximates how much more likely or unlikely it is for the outcome to be present among those with $x = 1$ than among those with $x = 0$ ” (p. 379). For example, in the case of this study, if the data have an odds ratio of 2, then the odds ratio estimates that long fixation times are twice as likely to occur with irregular verbs than regular verbs in the dataset. The results of the Mantel-Hänszel test indicated that extremely long fixation times are 3.24 time more likely to occur with irregular verbs than regular verbs, with a 95% confidence interval of (1.231, 8.526), $\chi^2 = 4.794$, $df = 1$, $p = 0.029$. This finding is very important in relation to RQ1, yielding that learners whose *observations* fell into the outlier pool fixated significantly longer on irregular verbs with vowel changes than regular verbs with no vowel change.

As a next step, I investigated whether I would see a similar pattern when further separating the irregular verbs by the type of stem-vowel change. I conducted the same Chi-square analysis with *verbcoding3* (1=irregular e → i/ie, 2=irregular a → ä, and 3=regular) and *posneg* (1=positive total time, and 0=Negative total time). The results of the Chi-square test

revealed that Total Time (*posneg*) was significantly associated with verb-coding3, $\chi^2 = 6.267$, $df = 2$, $p = 0.044$, Cramer’s $V = 0.273$. The results are illustrated in Figure 12.

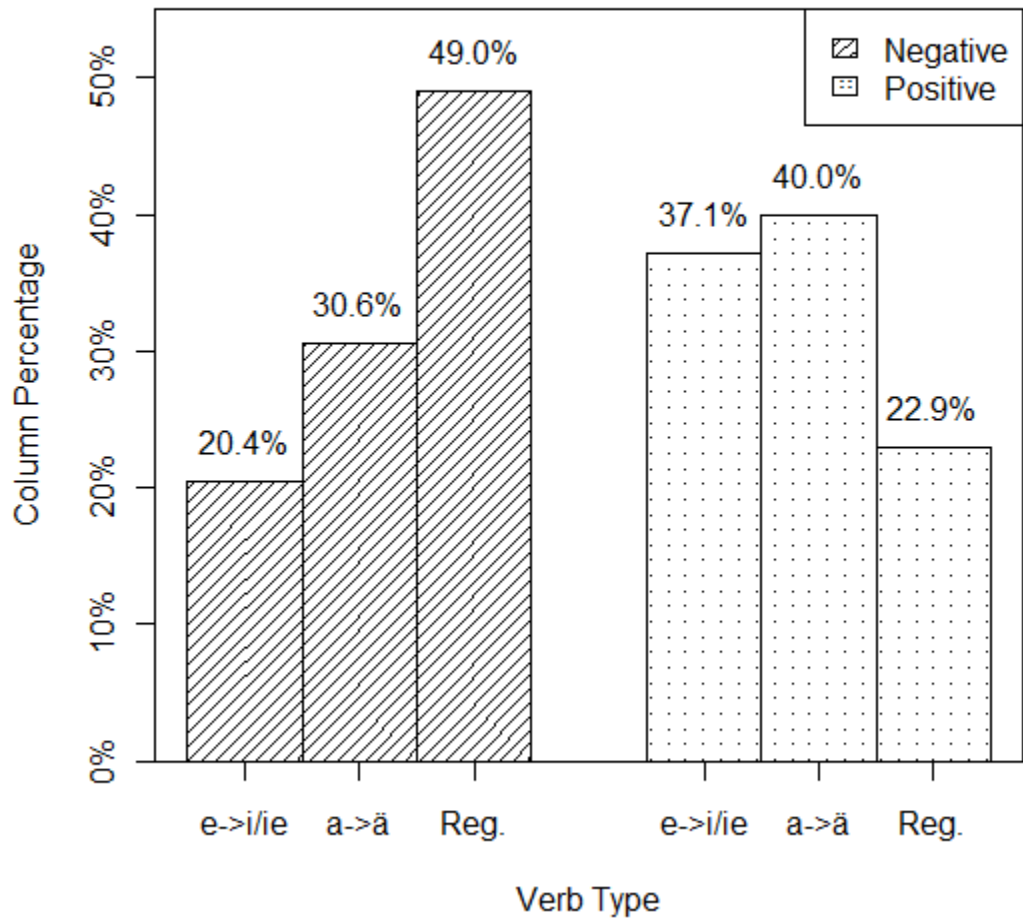


Figure 12. Verb Type and Total Time Gain (by Stem-change)

Figure 13 illustrates that a negative fixation time (which can be interpreted as no noticing) is most common with regular verbs (in nearly 50% of the cases), less common with a vowel change $a \rightarrow \ddot{a}$, and least likely to occur with verbs that change the vowel from $e \rightarrow i/ie$. Positive fixation time (i.e., longer fixations on the critical sentences than baseline sentences, which can be interpreted as *noticing*) occurs more commonly with $a \rightarrow \ddot{a}$ changes (40%), similar to $e \rightarrow i/ie$ changes (37%), and least common with regular verbs (only 22.9%). In summary, the

chart suggests that observations of *no noticing* occur with regular verbs almost half of the time, while *noticing* is considerably more common for stem-changes $a \rightarrow \ddot{a}$ and $e \rightarrow i/ie$.

I investigated this phenomenon further by looking at the odds ratio using logistic regression (for detailed information on logistic regression, please refer to Hosmer and Lemeshow (2000)). A logistic regression is a version of multiple regressions using dichotomous data. In my analysis the outlier data were coded as either positive total time gain values (coded with 1) or negative total time gain values (coded with 0). My predictor variables, verb types, were categorical as well. The logistic regression yielded which verb types (i.e., $e \rightarrow i/ie$; $a \rightarrow \ddot{a}$; or regular) predicted whether a learner fixated longer on the critical sentence (where the stem-change occurred) or not. The results of the logistic regression are summarized in Table 4.

Although the difference between the stem-changes are presented in the first row of the chart (for the purpose of completeness), I will only elaborate on the predictor variables of the $a \rightarrow \ddot{a}$ change and regular verbs (second row), and the regular verbs and the $e \rightarrow i/ie$ change (third row), Table 4 is best interpreted when looking at the odd ratios. As explained above, the odds ratio is a measure of effect size, describing the strength of association (or non-independence if smaller than 1) between binary data values (Field, 2009).

Looking at the second row, extreme fixation points on verbs with an $a \rightarrow \ddot{a}$ change are 2.8 times more likely to have a positive value (i.e., reveal noticing) compared to fixation times on

regular verbs with 95% C.I. (0.949, 8.262), $\chi^2 = 3.4785, df = 1, p = 0.0622$. This finding approached statistical significance. Looking at the third row of the chart, extreme fixation times on verbs with an $e \rightarrow i/ie$ change are 3.9 times more likely to have a positive value (i.e., reveal

noticing) compared to fixation times on regular verb with 95% C.I. (1.236, 12.303), $\chi^2 = 5.3909$, $df = 1$, $p = 0.0202$. This finding was statistically significant.

Table 4: Logistic Regression with Total Time

Fixation Time	Predictor Variable	Odds ratios	Confidence Interval		p-value
			Lower	Upper	
Positive Value* (=longer fixation on critical item than on baseline)	$e \rightarrow ilie$	1.393	0.464	4.185	0.5549
	$a \rightarrow \ddot{a}$	1 (Reference)	1 (Reference)	1 (Reference)	
	$a \rightarrow \ddot{a}$	2.800	0.949	8.262	0.0622
	<i>Regular</i> *	1 (Reference)	1 (Reference)	1 (Reference)	
	$e \rightarrow ilie$	3.900	1.236	12.303	0.0202
	<i>Regular</i> *	1 (Reference)	1 (Reference)	1 (Reference)	

*To obtain odds ratios for Negative value, take the reciprocal of the odds ratios of the Positive value

Assumptions of the Chi-Square and Logistic Regression

The use of a Chi-square test involves four assumptions: (1) the sample is random; (2) the sample size is reasonable; (3) adequate expected cell count (larger than 5); and (4) the observations are independent (Field, 2009). Concerning the first assumption, one could argue that the outlier sample was not truly random as these observations shared the outlier behavior of long fixation times. This is a minor limitation that needs to be acknowledged. My inspection frequencies of outliers revealed that observations from 74.5% of all participants fell into the outlier pool, and these participants contributed fairly evenly with the exception of participant 18 (11.9%) and participant 24 (10.7%). All verbs were present in the outlier pool except for the verb ‘studieren’ (= to study). Based on these inspections it was safe to assume that the first assumption was met. The second assumption was met by outlier sample-size ($n = 84$).

The third assumption was also met; the analysis indicated no cells with expected frequencies less than 5. The fourth assumption, the assumption of independence, was violated as is often the case in research social science research. According to Howell (2009), “[i]t is not uncommon to find studies in which the assumption of independence of observations is violated, usually by having the same participant respond more than once” (p. 152). Yet, violating the assumption can cause Type I errors and an overestimation of significant differences in the data (Field, 2009).³⁵ Thus, the assumption of independence of data in this particular study warrants further consideration. The unit of analysis here (i.e., fixation times) does not only originate from the individual, but it can be thought of as co-constructed by the particular trial (out of 24), the sentence ending (ending A or ending B) and the person used (2nd or 3rd) in the sentence.

Therefore, even though certain individuals contributed to some of the extreme fixation times that went into the outlier pool, they can be viewed as independent as they co-occurred with several related factors. In fact, Saito (1999) stated: “there is no suitable statistical procedure for analyzing multiple frequency data. It is therefore best to admit this difficulty and accept it as a limitation” (p. 648).

As with Chi-square tests, one assumption of the logistic regression is that the data are independent, which was assumed here (see explanation above). Logistic regression also works on the assumption that the dependent variable is binary or dichotomous. This was the case in this analysis as I only looked at either positive or negative values. I used the forced entry mode, which means that “all covariates are placed into the regression model on one block” (Field, 2009,

³⁵ It is important to note that Chi-squares do not test for interactions in the data, but it simply compares the frequencies observed in different categories (Field, 2009). The logistic regression, which I ran as a follow up, provided information on interactions in the data.

p. 226), which is an appropriate method for theory testing as opposed to finding a model that fits my data.

First Fixation Duration

First Fixation Duration ANOVA

In this section I present the results for the ANOVA with first-fixation duration presented in the split-plot design: two crossed fixed factors (verb-type and group) and a random subject factor (nested within group). The effect of group was non-significant and therefore this factor was removed from the model, yielding a two-way ANOVA with verb-type as a fixed factor and subject (nested within group) as a random factor. Similar to my analysis with total time, the preliminary analysis indicated lack of normality of the residuals by levels of the independent variables *verb type* and *group*. The lack of normality of the residuals could not be fixed by (a) removing influential outliers based on Cook's distance test, or by (b) removing cases whose corresponding standardized residuals were larger than 3.5 or smaller than -3.5. Therefore, analogous to my analysis with total time, I applied a truncation method on the dependent variable by including observation only within the range of -270 to 270 milliseconds that make up about 93.3% of the original data. Truncation insured normality of residuals across the levels of the independent variables and left me with 942 valid cases for the analysis. The 90 outlier-cases were analyzed separately with a Chi-square test and a logistic regression in a manner similar to the analyses presented in the previous section.

Because there was no main effect of group, I conducted a customized, nested ANOVA with the trimmed data set to model the main effects of *verb type* on *first fixation duration* with the random factor group and subject (nested within group). Similar to the modeling with total time gain, there was a significant effect of group and the nested subject, $F(39, 896) = 1.742, p =$

0.004, $\eta^2_p = 0.070$, which accounted for moderate variation in first fixation duration, as indicated by the effect size of 0.070. This means that 7.0% of between-subject variance regarding the first fixation duration was accounted for by this interaction. Contra to my expectations, verb type was not significantly associated with first fixation duration, $F(3, 896) =$

0.730, $p = 0.534$, $\eta^2_p = 0.002$.

Nested ANOVA Assumptions

The assumptions for the nested ANOVA with first fixation duration as the dependent variable were the same as the ones I described for total time. I closely examined the assumption of homogeneity of residuals variances and the assumption of normality of the residuals. Regarding the latter, the formal test of normality (Shapiro-Wilk) indicated that first fixation duration was normally distributed among the categories of the independent variables ($\alpha = 0.05$).

Similar to my analysis with total time, the Levene's test of homogeneity of variances indicated lack of constant variance across the levels of independent variables ($F(171, 770) = 1.569$ with $p < 0.001$). However, this finding should not influence the results as explained in section x. Using a Multivariate ANOVA for these data was appropriate, as judged by the F test

of goodness of fit: $F(126, 770) = 1.071$, $p = 0.294$ and $\eta^2_p = 0.149$. The latter indicates that 14.9% of the variation in first fixation duration was accounted for by this model which is considered to be large (Field, 2009).

Analysis of Outliers

I analyzed the omitted observations via a Chi-square test of independence (see analysis with total time outliers for detailed information regarding the Chi-square test as well as the relevant assumptions for its conduction). I conducted the first Chi-square test of independence ($\alpha = .05$) with irregular versus regular verbs (*verbcoding2*) and positive first fixation durations (noticing) versus negative first fixation durations (no noticing). Negative values indicated much longer fixations on the baseline condition compared to the critical condition, which can tentatively be interpreted as no noticing), and the positive values indicated longer fixations on the critical condition compared to the baseline condition, which can tentatively be interpreted as noticing). The hypotheses in consideration for the Chi-square are as follows:

H₀: The verb types (regular versus irregular) and first fixation time on the critical verb are independent.

This means that the outlier values for a given verb type (regular versus irregular) are evenly distributed among positive values (noticing) and negative values (no noticing).

H₁: The verb types (regular versus irregular) and first fixation time on the critical verb are dependent.

This means that the outlier values for a given verb type (regular versus irregular) are not evenly distributed among positive values (noticing) and negative values (no noticing).

The results of the Chi-square test indicated that first fixation duration (*Fposneg*) was not significantly associated with verb-type (irregular versus regular), $\chi^2 = 2.195, df = 1, p = 0.138$ OR = 0.531 with 95% C.I. (0.229, 1.231). Although the finding was not *statistically* significant, Figure 13 shows that the occurrence of outliers (either positive or negative) does vary as a

function of verb-type. Figure 14 shows that 57.1% of the outlier cases revealed extremely long first fixation times on irregular verbs, whereas long first fixation times on regular verbs were less common (42.9%). Recall that the same pattern transpired with total time gain, but the results of total time gain were statistically significant. Yet, although the Chi-square test results of the first fixation duration were not statistically significant, I infer that the outlier observations demonstrate some degree of noticing (i.e., longer fixations) of irregular verbs as opposed to regular verbs. The trend transparent with the negative values for first fixation time (i.e., no noticing) supports my inference. No noticing occurred in 58.5% with regular verbs, and 41.5% with irregular verbs. The Chi-square test with total time gain did not reveal a difference between regular and irregular verbs when it came to no noticing (i.e., negative total time gain).

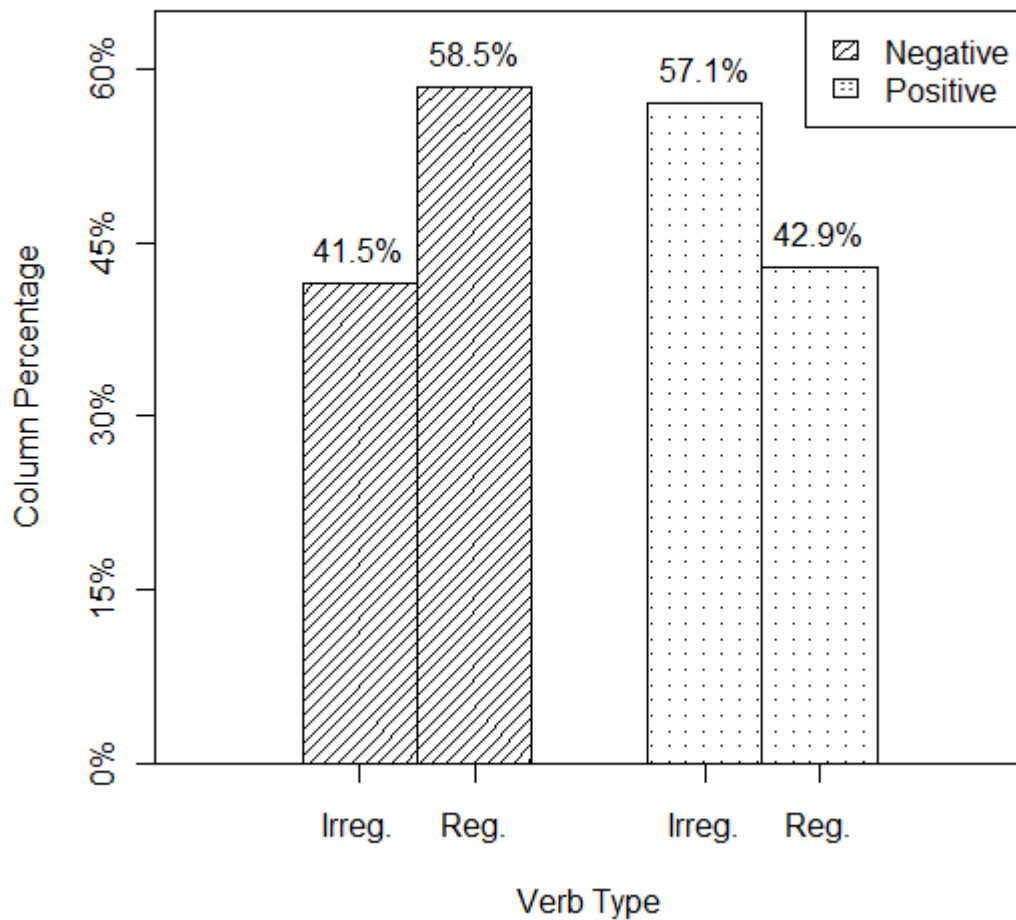


Figure 13. Verb Type and First Fixation (Regular versus Irregular)

As a next step, I investigated whether there was a significant pattern when further separating the irregular verbs by the stem-vowel change. I conducted the same analysis with *verbcoding3* (1=irregular e → i/ie, 2=irregular a → ä, and 3=regular) and positive first fixation duration versus negative first fixation duration). The results of the Chi-square test revealed that First Fixation Duration Time (Fposneg) was not significantly associated with verb-coding3,

$\chi^2 = 2.2230$, $df = 2$, $p = 0.328$, Cramer's $V = 0.157$. The results are visualized in Figure

14.³⁶

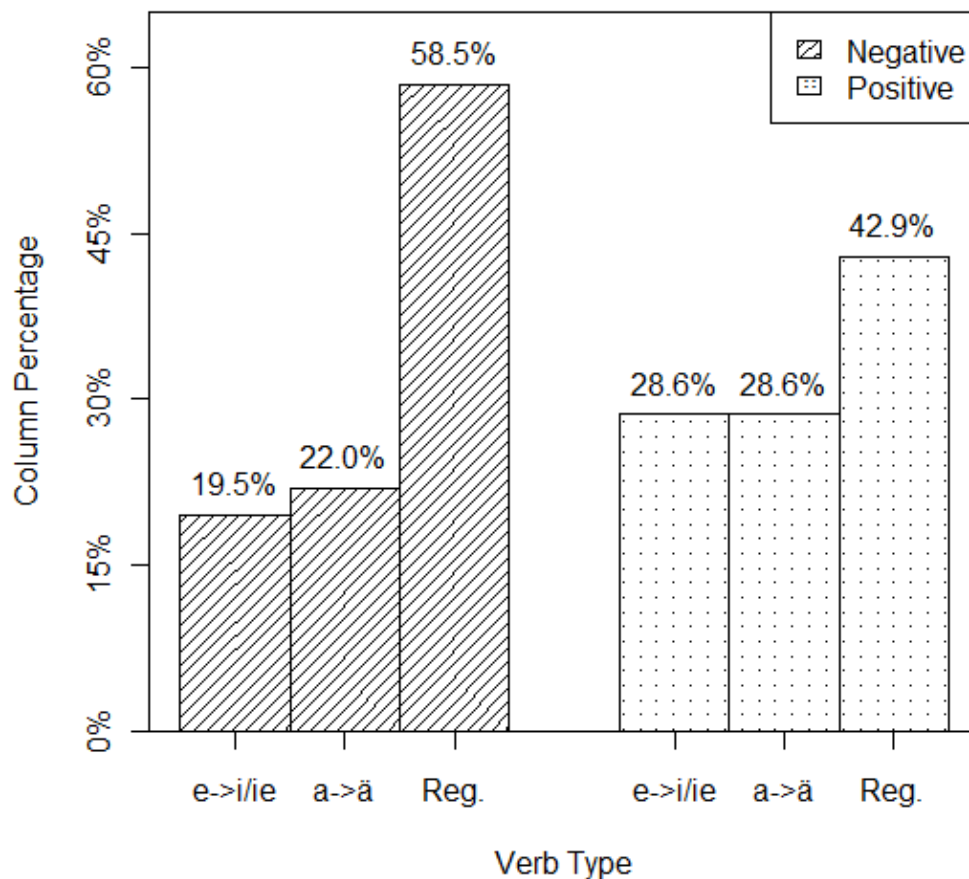


Figure 14. Verb Type and First Fixation (by Stem-change)

³⁶ The fact that the bars of regular verbs are higher is expected. There are both regular verb groups combined together, so it is 57.2%, which is a greater count than the irregular verbs. As such, this graph might be misleading unless the bar of the regular verb was reduced by half.

Figure 15 shows that negative first fixation duration (which can be interpreted as no noticing) is most common with regular verbs, and about equally common with the two types of vowel changing verbs. I investigated this phenomenon further by looking at the odds ratio using a logistic regression analysis. The logistic regression yielded which verb types (i.e., $e \rightarrow i/ie$; $a \rightarrow \ddot{a}$; or regular) predicted whether a learner fixated longer on the critical sentence or not.

Table 5 is best interpreted when looking at the odds ratios. The logistic regression revealed that extreme fixations on verbs with an $a \rightarrow \ddot{a}$ change are 1.778 times more likely to have a positive value (reveal noticing) compared to fixation times on regular verbs with 95% C.I.

(0.640, 4.939), $\chi^2 = 1.2179$, $df = 1$, $p = 0.2698$. Extreme fixations on verbs with $e \rightarrow i/ie$ are 2.0 times more likely to have a positive value compared to regular verb with 95% C.I. (0.702,

5.702), $\chi^2 = 1.6814$, $df = 1$, $p = 0.1947$. Although the findings were not statistically significant, they exhibit trends that are notably similar to the findings for total time gain.

Table 5: Logistic Regression with First Fixation Duration

Fixation Time	Predictor Variable	Odds ratios	Confidence Interval		p-value
			Lower	Upper	
Positive Value* (= longer fixation on baseline item than on critical)	$e \rightarrow i/ie$	1.125	0.337	3.760	0.8483
	$a \rightarrow \ddot{a}$	1 (Reference)	1 (Reference)	1 (Reference)	
	$a \rightarrow \ddot{a}$	1.778	0.640	4.939	0.2698
	<i>Regular</i>	1 (Reference)	1 (Reference)	1 (Reference)	
	$e \rightarrow i/ie$	2.000	0.702	5.702	0.1947
	<i>Regular</i>	1 (Reference)	1 (Reference)	1 (Reference)	

*To obtain odds ratios for negative value, take the reciprocal of the odds ratios of the positive value.

Gaze Duration ANOVA

As a last analysis to address RQ1, I conducted a nested ANOVA with the fixed factor *verb type* crossed with the random factor subject/group (subject nested within group) to model the main effects of verb type on gaze duration (the dependent variable). Recall that gaze duration is the sum of all fixations made in an interest area before the eyes leave that area. As with total time and first fixation duration, the preliminary analysis indicated lack of normality of the residuals by levels of the independent variables *verb type* and *group*. The lack of normality of the residuals could not be fixed by a) removing influential outliers based on Cook's distance test, or by b) removing cases whose corresponding standardized residuals were larger than 3.5 or smaller than -3.5. Therefore, analogous to the previous ANOVAs, I applied the truncation method on the dependent variable by including observation only within the range of -650 milliseconds and 600 milliseconds, which make up about 93.5% of the original data. Truncation insured normality of residuals across the levels of the independent variables. After the removal of influential outliers (78 cases) I ran the nested ANOVA with 954 observations.

Similar to the models for total time and first fixation duration, there was a significant effect of group and the nested subject, $F(39, 908) = 2.395, p < 0.001, \eta_p^2 = 0.093$, which accounted for moderate variation in gaze duration, as indicated by the effect size η_p^2 of 0.093. Thus, 9.3% of between-subject variance in gaze duration was accounted for by individual differences in reading speed among the readers. More important, however, verb type had no significant effect on the gaze duration' $F(3, 908) = 0.176, p = 0.913, \eta_p^2 = 0.001$. This means

that the extra processing time readers allocated to the processing of the critical sentences did not differ as a function of verb type.

Assumptions

The formal tests of normality (Kolmogorov-Smirnov and Shapiro-Wilk) indicated that the gaze duration was normally distributed among the categories of the independent variables ($\alpha = 0.001$).

The Levene's Test of homogeneity of variances indicated a lack of constant variance across the levels of independent variables ($F(171, 782) = 1.275$ with $p = 0.017$). This finding should not influence the results, as explained above. Using a Multivariate ANOVA for these data was appropriate as judged by the F test of goodness of fit: $F(126, 782) = 0.988$, $p = 0.522$

and $\eta_p^2 = 0.137$. The effect size indicates that 13.7% of the variation in gaze duration was accounted for by this model.

Analysis of Outliers

Analogous to my outlier-analyses with total time and first fixation duration, I conducted a Chi-square test of independence with irregular versus regular verbs and positive values of gaze duration (noticing) versus negative values of gaze duration (no noticing). The hypotheses tested for the Chi-square are as follows:

H_0 : The verb types (regular versus irregular) and gaze duration on the critical verb are independent.

This means that the outlier values for a given verb type (regular versus irregular) are evenly distributed among positive values (noticing) and negative values (no noticing).

H1: The verb types (regular versus irregular) and gaze duration on the critical verb are dependent.

This means that the outlier values for a given verb type (regular versus irregular) are not evenly distributed among positive values (noticing) and negative values (no noticing).

The results of the Chi-square test indicated that Gaze duration was not statistically significantly associated with verb type (irregular versus regular), $\chi^2 = 1.363$, $df = 1$, $p = 0.243$, OR=0.563 with 95% C.I. (0.213, 1.485). In other words, the regularity of verbs was not statistically significantly associated with the distributions of extreme gaze durations between negative values (no noticing) and positive values (noticing). Figure 15 illustrates this. These findings are different from those with total time duration where I found a significant difference. The bar graph, however, shows the same trend found with total time and first fixation duration with regard to negative values. No noticing occurred with regular verbs. Looking at the positive values, I expected more positive value counts for irregular verbs based on the trend discovered with total time and first fixation. In contrast, the bar charts of positive fixation times are approximately the same for both, regular verbs and irregular verbs. This discontinuity of the “noticing trend” found with total time and first fixation duration might be due to the fact that the outlier pool wasn’t exactly the same for each dependent variable. That is, for gaze duration truncation left me with 79 outliers, with total time I identified 84 outliers, and with first fixation I identified 90 outliers. Beyond that, it is important to keep in mind that the graphic representations include 5% chance for error (confidence interval of 95%).

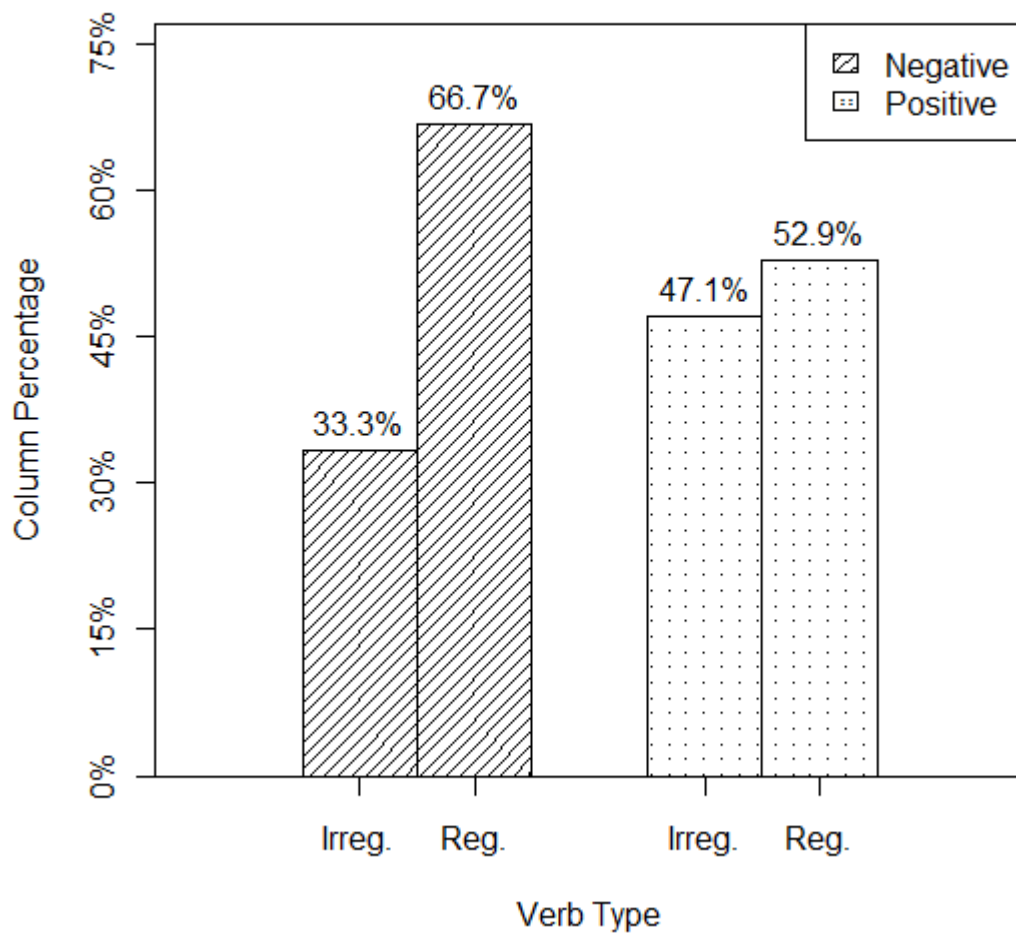


Figure 15. Verb Type and Gaze Duration (Regular versus Irregular)

As a next step for the gaze duration outlier analysis, I investigated whether there was a pattern when further separating the irregular verbs by the type of stem-vowel change. I conducted the same analysis with *verbcoding3* (1=irregular e → i/ie, 2=irregular a → ä, and 3=regular) and the values for positive gaze duration, and negative gaze duration. The results of the Chi-square test revealed that gaze duration was not significantly associated with the type of

stem-change, $\chi^2 = 4.351$, $df = 2$, $p = 0.114$, Cramer's $V = 0.236$. The results are presented in Figure 16.

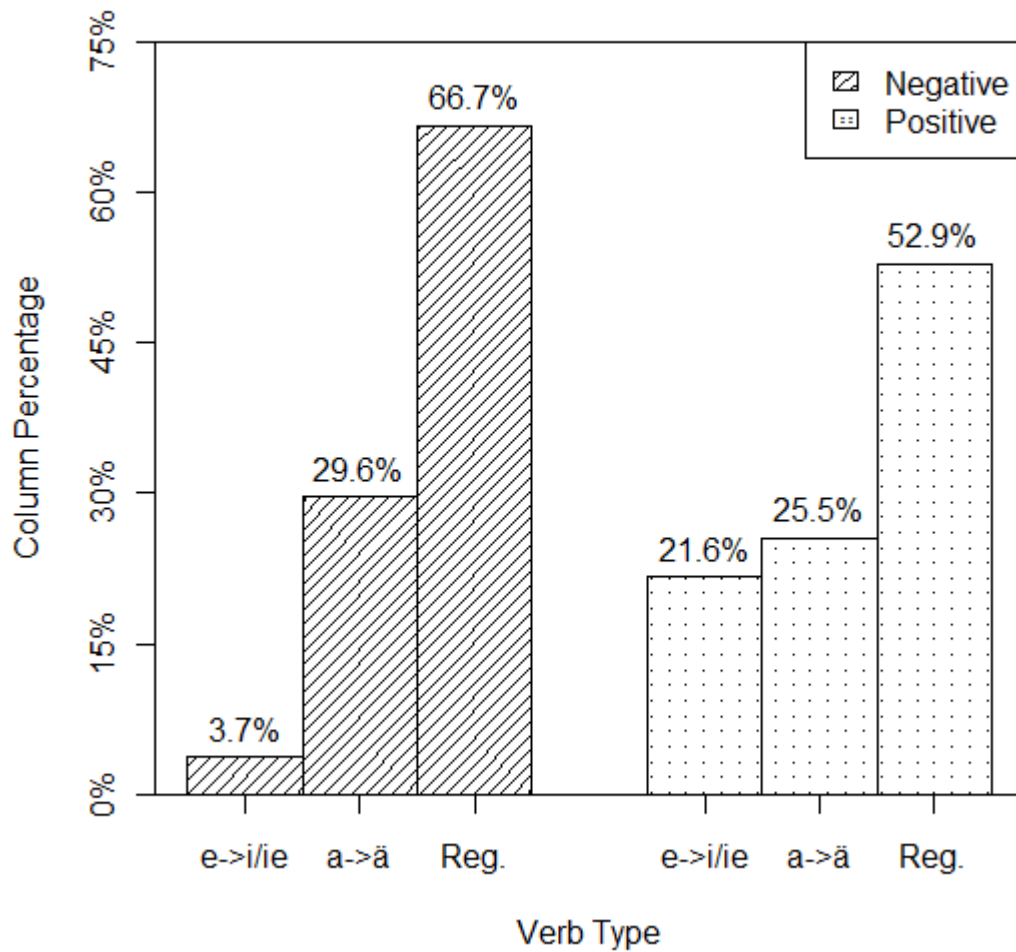


Figure 16. Verb Type and Gaze Duration (by Stem-change)

Figure 16 shows that extreme gaze duration values on both ends—negative and positive—are most common with regular verbs, less common with a vowel change from a → ä

and least likely to occur with verbs than change the vowel from e → i/ie. I investigated this phenomenon further by looking at the odds ratios using a logistic regression.

Table 6 is best interpreted when looking at the odds ratio. Gaze durations on verbs with a → ä vowel changes are 1.083 times more likely to have a positive value compared to gaze

durations on regular verbs with 95% C.I. (0.374, 3.138), $\chi^2 = 0.0218$, $df = 1$, $p = 0.8827$. Gaze

durations on verbs with e → i/ie vowel changes are 7.333 times more likely to have a positive value compared to verbs with an e → i/ie change with 95% confidence interval (0.870, 61.840),

$\chi^2 = 3.3542$, $df = 1$, $p = 0.0670$. This finding is approaching significance. The findings are in line with the outlier findings of total time and first fixation, although a stronger effect of a → ä changes was expected. The results of this analysis are summarized in Table 6.

Table 6: Logistic Regression with Gaze Duration

			Confidence Interval		
Fixation Time	Predictor Variable	Odds ratios	Lower	Upper	p-value
Positive Value* (=longer gaze duration on critical item than on baseline)	<i>e → i/ie</i>	6.769	0.729	62.859	0.0926
	<i>a → ä</i>	1 (Reference)	1 (Reference)	1 (Reference)	
	<i>a → ä</i>	1.083	0.374	3.138	0.8827
	<i>Regular</i>	1 (Reference)	1 (Reference)	1 (Reference)	
	<i>e → i/ie</i>	7.333	0.870	61.845	0.0670
	<i>Regular</i>	1 (Reference)	1 (Reference)	1 (Reference)	

*To obtain odds ratios for Negative value, take the reciprocal of the odds ratios of the Positive value.

Summary of Findings for Research Question 1.

Looking at the big picture of the results, the irregularity of the verbs appears to have an effect on fixation times. Even though this finding was not statistically significant in every

analysis, there was a trend that learners fixated longer on verbs where the stem-vowel changed. The nested-ANOVA with total time revealed a significant effect of verb type, and the post-hoc test revealed that there was a significant difference of total time between a → ä vowel changes and regular verbs (longer fixations were made on verbs with vowel changes). There also was a notable difference of total time between e → i/ie vowel changes and regular verbs, even though this difference was not statistically significant. My analysis of the outlier observations (cases with very long fixation times) revealed a similar trend: Total time was significantly associated with verb type (regular versus irregular verbs). When separating the irregular verbs by type of vowel-change, I found that verbs with an e → i/ie change were fixated significantly longer than regular verbs. Learners also fixated longer on a → ä changes relative to regular verbs, although this difference was not statistically significant.

Although my analyses (regular data and outliers) with first fixation duration did not yield statistically significant results, the findings exhibited a trend that was notably similar to my findings for total time. Similarly, my findings with gaze durations (regular data and outliers) were not significant, but my analysis of outliers showed that verbs with an e → i/ie change were fixated longer than regular verbs, and this finding approached statistical significance with moderate and occasionally large effect sizes.

Results Addressing RQ2

While the first research question asked whether beginning learners pay attention to verb irregularities during reading, the second research question inquired whether increased attention to irregular verbs is associated with subsequent uptake (i.e., a learning gain on post-tests) of those irregularities (i.e., stem-vowel changes). It is important to note at the outset that I addressed RQ2 by means of a correlation analysis. As is well known, correlation does not

indicate causation, and thus I will interpret any significant results as indicating a statistically significant *association* rather than a *causal relationship* between attention and learning. Thus, I use the Spearman (two-tailed) correlation analysis to address RQ2. It is also important to mention at the outset, that all analyses consider the experimental trials (i.e., trials with irregular verbs) only, as the participants were not tested on their improvements of regular verbs, which do not contain stem changes. Figure 17 below illustrates the design of this study for clarity purposes. The shaded part of the design is relevant for this correlation analysis.

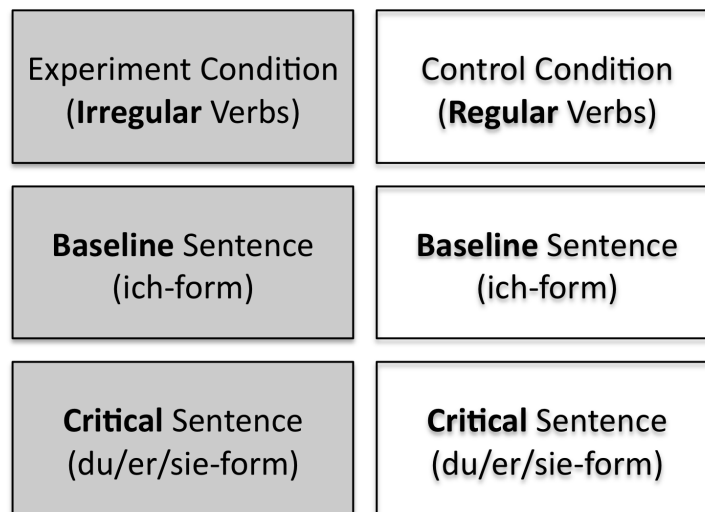


Figure 17. Design of the Study

Based on the ‘eye-mind assumption’ (Chapters 2 and 3), I expected to find a positive correlation between eye-fixation duration (total time) and the learning gains as measured by the difference from the pretest too posttest. Since the ANOVA models generally revealed that the subject variable accounted for a significant variation in fixation time (i.e., participants differed significantly in reading speed), I first ensured that the total time gain values that I used in the correlation analysis were adjusted for the individual’s overall reading speed by all participants. In particular, I aimed to control for the possibility of an interaction between total time gain and

overall reading speed (whereby slower readers would generally show larger total time gains than faster readers) by adjusting for participants' overall reading speed.

To this aim, I used two approaches: The first one was a manual calculation of a total time gain value that was corrected, or better yet, *adjusted* for the differences on reading speed. To accomplish this adjustment, I first computed the average reading time per participant. That is, for each individual, I calculated the summed total time of all reading trials (only the sentence slides) and I divided it by 24 (each session contained 24 trials or sentences-slides). This computation left me with a mean reading time per trial for each participant. Next, I computed a grand mean, which was the sum of all participants' mean reading time per trial divided by the number of all included participants included ($n = 43$). For each participant, I then divided his or her average reading speed by the grand sample mean sample value. This outcome provided me with the value by which to correct each participant's total time gain value (for each verb). The *adjusted total time gains* for each verb (= 1,032 observations/cases in total) were the values I used to run the Spearman rank correlation with the learning gains.

Another way of controlling for reading speed is by running a semi-partial correlation. According to Field (2009), when running a partial correlation between two variables, the effect of a third variable on *both* variables is controlled for. When running a *semi* partial correlation (which can only be done by means of a regression analysis), the effect that a third variable (e.g., *overall reading speed*) has on only *one* of the variables in the correlation (e.g., *total time gain*) is controlled. My preliminary analysis showed a significant relation between reading speed and

holistic gain³⁷ ($r = .162, p = .000$), but no significant relationship between reading speed and gains on stem changes ($r = .035; p = .125$). Regarding the latter, in a strict sense, a semi partial correlation (regression analysis) would be more appropriate. However, I conducted partial correlation analyses with *both* types of learning for a) ease of interpretation and b) because a regression analysis would bring several other issues (satisfying assumptions, etc.) which would not bring a greater value to the overall results I aim to obtain. Figures 19 and 20 below present the concepts of a partial and semi partial correlation. I opted to conduct the partial correlation (controlling for reading speed) to investigate the relationship between total time spent on the irregular verb only (*not* total time gain) and holistic learning and the relationship between stem change learning and total time gain. I present the findings after presenting the findings for the analyses with *total time gain adjusted* (i.e., the other option of controlling for reading speed).

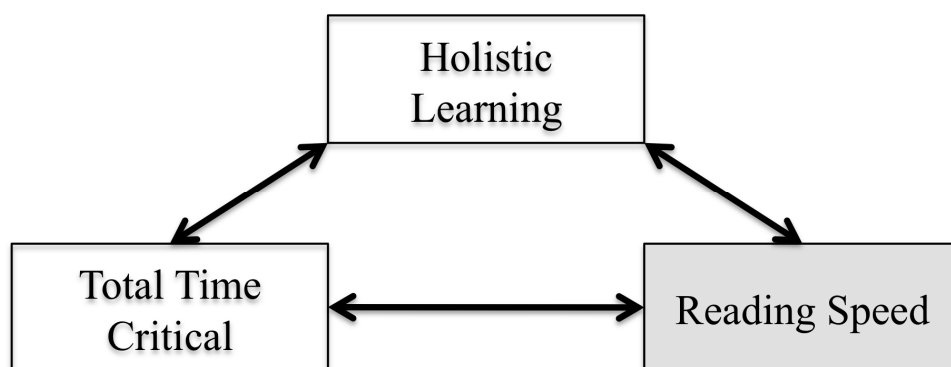


Figure 18. Illustration of Partial Correlation (controlling for reading speed)

³⁷ Recall from Chapter 4 that I scored the pre-and posttest in two different ways: I scored them (1) holistically, giving points for improvement on spelling and verb inflection, and (2) with regard to the vowel-stem change only.

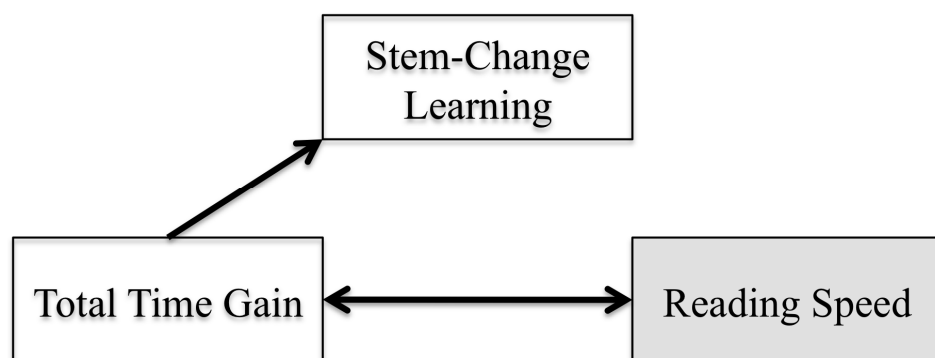


Figure 19. Illustration of a Semi Partial Correlation (Controlling for Reading Speed)

I used the Spearman's correlation coefficient, r , which is a non-parametric statistic and therefore the recommended test of correlation when the data have violated assumptions such as non-normally distributed data.³⁸ The histogram of fixation time gain showed a positive kurtosis (a high peak in the middle) because the data had many values around zero which are indicative of similar total times for the baseline item and the critical item. An alpha level of .05 was set for all correlations.

*Correlation of Fixation Times with Learning Gains*³⁹

Overall, *total time gain adjusted* was not significantly correlated with stem-change learning gains $r = .039, p = .566$. There was also no relationship between verb-type (a → ä vs. e → i/ie) and learning gains, which means that the learners' improvement was not related to the type of stem-change on which they were tested. Because adjusted total time gain corresponds to

³⁸ I also ran the Pearson's correlation (parametric) and Kendall's tau (non-parametric) and received nearly identical results. For accuracy purposes, however, I am reporting the results of Spearman rank correlation.

³⁹ Recall from Chapter 4 that I am only including adjusted total time gains for the irregular (stem-changing) verbs, not for the regular verbs. The pre-and posttest did not include the entire regular verb set, but only a subset, which functioned as distracters.

the *difference* between the baseline and the critical verb and is reflective of a participant's increased attention to irregularities, this difference might not be relevant in regard to *holistic* learning. Therefore, in order to investigate the relationship between fixation time and holistic learning, I ran a correlation analysis with the total time for the *critical* item (du/er/sie form) and *holistic* learning. The total time spent on the verb with the vowel change was significantly related to holistic learning gains, $r = .087, p = .048$. This relationship was relatively weak, which means that we need to be cautious when generalizing this finding to the overall population.

Partial Correlation of Fixation Times with Learning Gains

The partial correlation, where the effect of reading speed was controlled, revealed that there was no relationship between the total time spent looking at the irregular verb and holistic learning: $r = -.056, p = .104$.⁴⁰ There also was no relationship between the learning gains regarding stem-changes and total time, $r = -.002, p = .484$.

I was then interested to see whether *positive* fixation times (longer fixation times on the critical verb, which had a stem-change, than on the baseline verb, which had no stem-change) were correlated with learning gains. I re-ran the Spearman correlation twice with negative values excluded list-wise. I correlated positive fixation times (using the adjusted values) with the learning of stem-changes, and I correlated positive fixation times with overall (holistic) learning gains. The positive fixation times (i.e., noticing) were not correlated with the learning of the stem-changes, $r = .092, p = .131$. However, the positive fixation times were positively correlated with *holistic* learning gains, $r = .187, p = .002$. This means that there was a relationship between longer fixation times on stem-changing verbs (relative to the baseline condition) and *overall*

⁴⁰ There was a significant relationship between the two variables when I did not control for reading time, which corroborates the effect of overall reading speed on fixation times on the verbs.

gains on written production of all verbs, but not on the learning of the stem-changes alone. The relationship was not very strong, which means that 18% of students learning in all probability is related in art (i.e., 18%) to focusing on and attending to them.

Due to the significant finding with positive fixation times, I also investigated whether there might be a relationship between *negative* fixation values (i.e., longer fixation on the baseline condition) and the test scores. In other words, I correlated *negative* fixation times with learning gains on the stem changes, and I also correlated *negative* fixation times with overall (holistic) learning gains. There was no significant relationship between *negative* fixation values and learning gains on *stem changes*, $r = -.055$, $p = .395$. There was, however, a *significant negative correlation* between negative fixation values and holistic learning gains, $r = -.185$, $p = .013$, demonstrating that longer fixations on baseline verbs (as opposed to the critical verbs) was related to better holistic learning.⁴¹ This means that there was a relationship between longer fixation times on *baseline* verbs (relative to the stem-changing verb condition) and *overall gains* on the written production of irregular verbs, but not on the learning of the stem-changing verbs alone.

This finding logically suggests the possibility of a *quadratic* association between adjusted total time gain and holistic learning gains.⁴² For example, a participant who has a total time gain of, for example, -10 milliseconds (i.e., fixating both, the baseline and the critical verb, about equally long) is less likely to demonstrate learning gains on the posttest than a participant of has a total time gain of, for example, -500 milliseconds (which is indicative for longer focus on the baseline verb with no vowel change). In other words, there might be a *parabolic* relationship

⁴¹ A *negative* correlation coefficient indicates that as one variable increases, the other decreases, and vice-versa.

⁴² A quadratic relation is represented graphically by means of a parabola.

between the adjusted total time gain and holistic learning gains. To check whether this assumption is correct I conducted a correlation analysis between the squared value of adjusted total time gain (x^2) and *holistic* learning gains. There was a significant correlation between the squared adjusted total time gain and holistic learning gains, $r = .125$, $p = .004$. Although the relationship is not very strong, it indicates a strong likelihood of a linear relationship between total time gain squared and holistic learning gains. In other words, it indicates a quadratic association between total time gain and holistic learning gains. Even if the participants processed the *ich*-form longer than the *du/er/sie* form (vowel change), the participants improved in the posttest overall. Consequently, longer fixations on *either* form, the non-changing *ich*-form or the vowel-changing *du/er/sie* form, was associated with holistic learning gains. Figure 20 below represents (in a simplistic manner) the quadric relationship between the adjusted total time gain and the test scores.

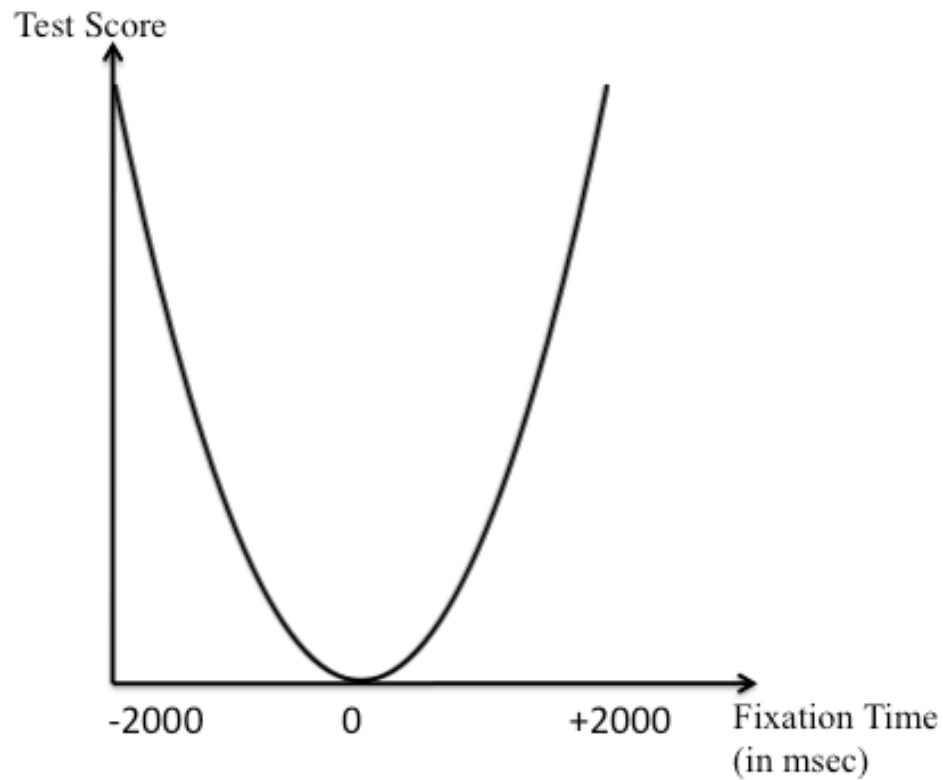


Figure 20. Simplistic Depiction of the Parabolic Relationship

My visual inspection of the testing data showed that the pre-and posttest scores for each verb varied to some extent. That is, some irregular verbs were “learned” to a greater extent than others, and the initial knowledge of the correct spelling of the verbs differed (see Figures 21 and 22).

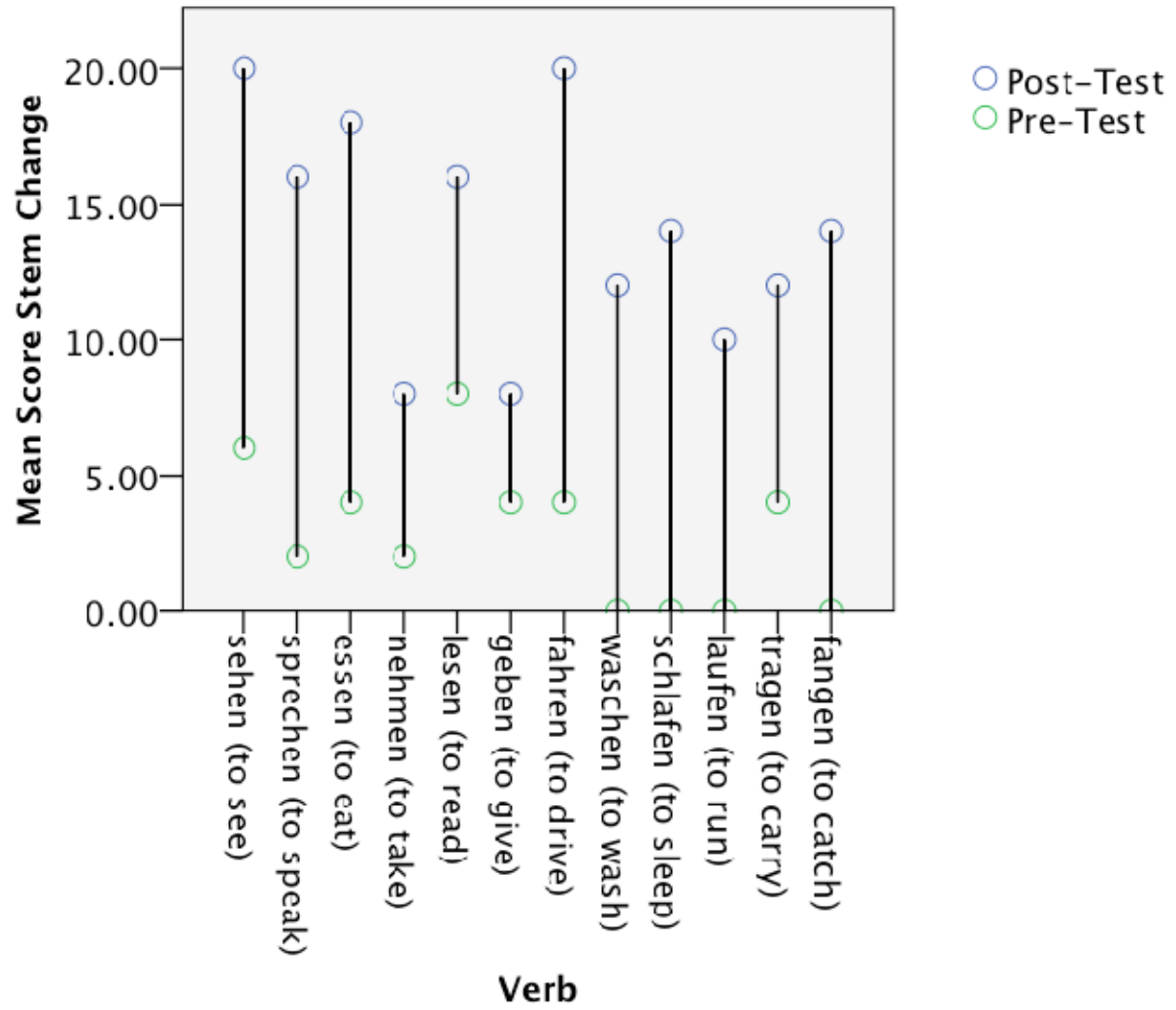


Figure 21. Verb and Mean Scores Stem-change

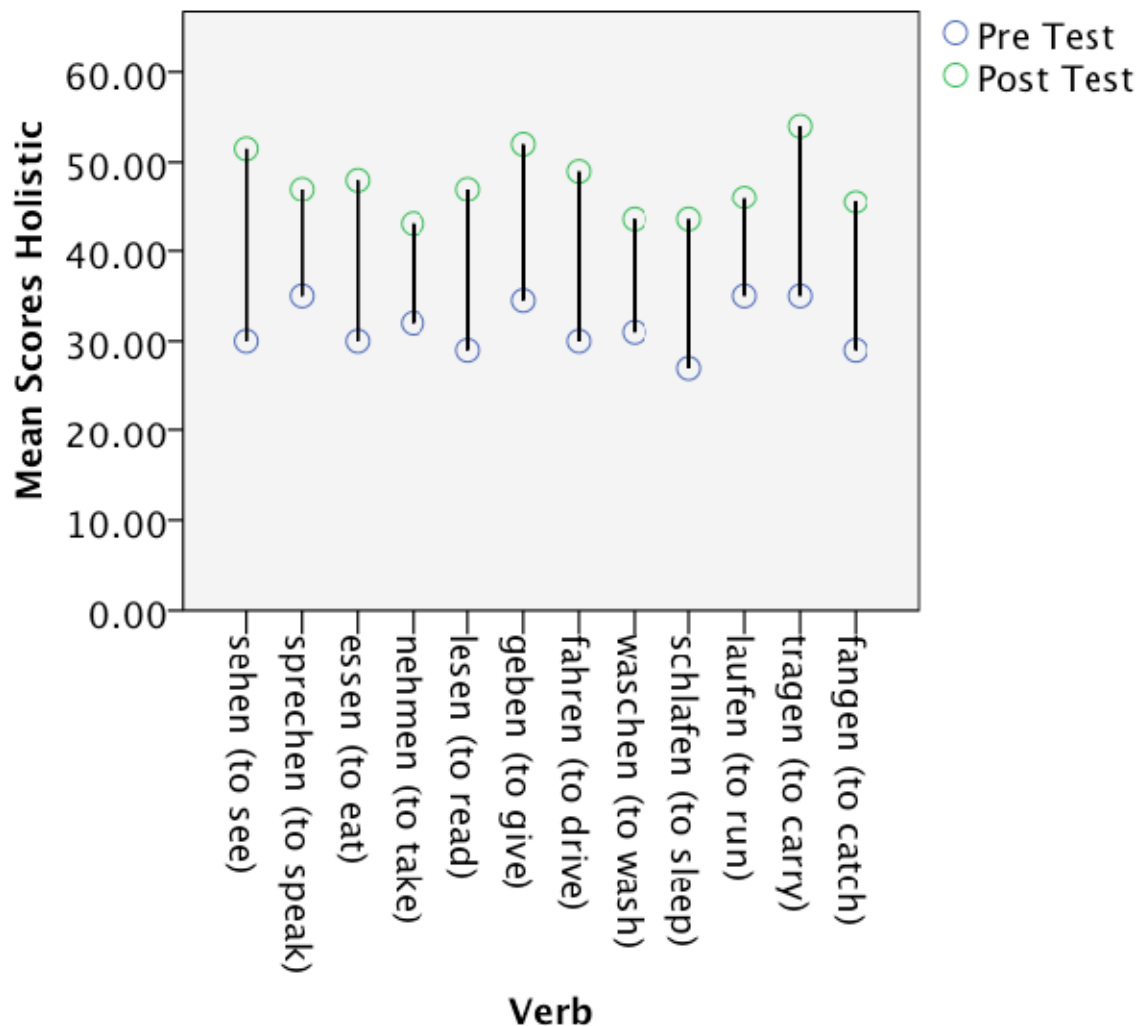


Figure 22. Verb and Mean Scores Holistic

I expected that the participants' initial knowledge of stem-changes equaled zero, but very few participants used the correct stem-vowel on the pre-test. It remains unknown whether this was due to chance or whether the students had read ahead in their textbook. Based on these visual inspections, I was interested whether there were significant correlations between fixation times and learning when looking at the different verbs individually. Two verbs showed statistically significant positive correlations between fixation times and learning gains: (1) Fixations on 'sehen' (*to see*) were positively correlated with *holistic* learning gains, $r = .322$, $p =$

.036, and (2) fixations on ‘fahren’ (*to drive*) were positively correlated with stem-change gains, $r = .472, p = .001$. The remaining 10 irregular verbs did not reveal significant correlations between their fixation times and learning gains. This finding will be discussed in the next chapter.

Summary of Findings for RQ 2

The correlations between fixation times and “learning gains” (measured by improvements from the pre- to the posttest) revealed a significant relationship between *positive* fixation times (i.e. longer fixation times on stem-changing verb forms than on the corresponding baseline forms) and *overall* gains on the written production of those verbs. Similarly, there was a significant negative correlation between *negative* fixation values and holistic learning gains. These results gave rise to the assumption of a possible parabolic relationship, which was confirmed. This means that longer fixation times on the *ich*-form verb (compared to the *du/er/sie*-form verb), or longer fixation times on the *du/er/sie*-form verb (compared to the *ich*-form verb) were associated with overall learning gains. This means that it was not relevant *which* of the two verbs was fixated longer, but that one verb was processed for a longer time.

When looking at the verbs separately (selecting them list-wise from the data pool) and when considering both negative and positive fixation times (total time gain adjusted) in a single analysis, the verbs “sehen” and “fahren” stood out in that fixation times on those verbs correlated with learning gains.

CHAPTER 6 DISCUSSION

This dissertation was guided by two RQs:

- (1) Do adult beginning learners of German who are unfamiliar with stem-changing verbs pay attention to those irregularities during reading?
- (2) Is increased attention to irregular verbs associated with subsequent learning of those verbs?

In discussing the findings pertaining to RQ1, I consider the following two outcomes that emerged from the data in this study: (a) why, based on analyses conducted with data from the majority of the learners,⁴³ I found evidence for attention to irregular verbs in regards to *total time*, but *not* with *gaze duration* or *first fixation duration*; and (b) why there were “outlier cases” that exhibited extreme fixation times, and what differentiates an “outlier case” from an average case. I then turn to RQ2 and discuss the following: (a) the lack of a linear association between fixation time (i.e., total time gain) and learning; (b) the issue of non-saliency; (c) item learning versus system learning, (d) the role of the developmental level of the learner; (e) the parabolic relationship; and (e) why the learners in this study reacted differently when reading and when being tested on the two verbs 'sehen' and 'fahren', compared to the other verbs.

Research Question One

As established in Chapters 2 and 3, based on the “eye-mind assumption” (e.g., Godfroid, 2010; Just & Carpenter, 1980; Rayner, 2009), I assumed that there is a close link between eye position and attention during reading (Rayner, 2007). Recall that for RQ1, I predicted participants would fixate longer on and make more regressions to stem-changing verbs, whose

⁴³ When using the term “*majority* of learners” I am referring to the trimmed data set (i.e., the overall data minus the outliers).

irregular stem-vowels would constitute a novelty to them, than to regular verbs, which have no stem-vowel changes.

Attention to Irregular Verbs

When considering the summation of all the fixations a learner makes on a word (i.e., *total time*), the irregularity of the verb appeared to have effect. To give an example, when a learner fixated on a verb for a longer time (possibly due to multiple fixations on the verb, some of them being regressions), the verb was significantly more likely to be an irregular verb than a regular verb.⁴⁴ Moreover, when examining which type of irregular verb learners fixated on longer (either the e → i/ie type or the a → ä type), I found that the latter verb type garnered longer fixation durations compared to non-vowel changing, regular verbs. Learners also tended to fixate longer on verbs with an e → i/ie change, as opposed to regular verbs, but this difference did not reach statistical significance. These joint findings lead me to assume that the learners attended to the vowel differences in the irregular verbs. At the minimum, it might be safe to state that, overall, the learners' eyes dwelled⁴⁵ significantly longer on verbs that had a vowel change, which was a new form for them. In L1 contexts, "[h]ow long readers fixate on a word is related to the ease or difficulty with processing that word" (Rayner, 1998, p. 186). It might also be the case that some word incongruity, for example, a change in the syllable nucleus, affects the processing of that word. The longer fixations and re-fixations on the irregular verbs are likely to be accompanied by more processing activity, given the assumption that there is a close link between the eyes and

⁴⁴ Since both conditions, regular and irregular/vowel changing consisted of a *baseline* verb (*ich*-form) and a *critical* verb (*duler/sie*-form) in both conditions, the irregular condition and the regular condition), it is safe to use the term "longer" fixations. We can say that learners fixated *longer* on the irregular, critical verbs than regular verb forms, as both instances were compared to a corresponding baseline condition.

⁴⁵ Recall that 'dwell time' is the actual term for 'total time'.

the mind, or that the learner's mind was the impetus guiding the eyes (Clifton, Staub, & Rayner, 2007; Just & Carpenter, 1980; Pollatsek & Rayner, 2009; Rayner, 2009; Rayner, Pollatsek, et al., 2003). Yet, do these longer fixations indicate whether the learners actually *noticed* the vowel-stem change? My conceptualization of noticing (see Chapter 1, pp. 33-34) is that during reading, noticing involves focal attention to surface structures and lexical items that might create a novelty for the learner.⁴⁶ I assumed that awareness is an *involved* process, even if only peripherally. Based on this premise, I maintain that the irregularity attracted learners' focal attention (i.e., noticing) even if peripherally and with—possibly—some level of awareness involved. I will address the question of whether the learners “noticed” again as I take into consideration the learning gains as well.

Based on the findings with total time, I expected to find a similar pattern with first fixation and gaze duration, as the majority of eye-movement studies evidenced the same pattern with all three types of fixation time measures (e.g., Calvo & Meseguer, 2002; Clifton et al., 2007; Frazier & Rayner, 1982; Godfroid, 2010; McConkie, Kerr, & Dyre, 1994; Reichle, Rayner, & Pollatsek, 1999; Williams & Morris, 2004). Contrary to my expectations, however, these fixation time measures did not indicate that irregular verbs were fixated on longer than regular verbs. This means that when the learners' eyes fixated on the verb during *first pass* (the first time the eyes went over the verb), there was no difference depending on whether the verb included a stem-vowel change or not. Irregularity did not have an effect on either first fixation or gaze duration. Since the measure of *total time* also includes the instances when the learners *re-fixate on* (i.e., go back or make a regression to) a verb, these regressions and subsequent

⁴⁶ Recall also that I did not operationalize noticing at the level of understanding.

readings appear to be triggered by the irregularities. Another possibility is that learners re-read the verb because of contextual cues the remaining part of the sentence may have contained. Based on these results alone I can only make tentative claims as to why learners made regressions and therefore viewed the verb for a higher total amount of time. One possibility is that they took time to process the vowel-change, a feature the learners had not encountered before. Future analyses need to also look at second pass times and/or regressions separately, an issue I will address again in the section on Future Research.

“Outlier Cases” with Extreme Fixation Times

Recall that the preliminary analysis of my data revealed outliers, which were cases of extremely long fixation times (see Chapter 5). That is, considering the entire pool of observations, there were cases⁴⁷ in which learners fixated much longer on certain verbs than normally observed (about 2,000 milliseconds in total compared to an average of 200-300 milliseconds in total). These cases were, accordingly, *outlier* cases in that they differed markedly from the average observations in this study; they created “noise” in the quantitative statistical methods of data analysis. Dörnyei (2009) argued that “‘Noise’ is important” (p. 107) because it reflects the Dynamic Systems Theory described by Larsen-Freeman and Cameron (2008), which attempts to account for the complexity of how humans use, learn, and teach languages. “[W]hat has been considered as ‘noise’ in quantitative studies *does* matter and should not be eliminated through the quantitative focus on the central tendency at the group level” (Dörnyei, 2009, p. 107, highlights are mine).

⁴⁷ Specifically, 75.5% of all learners displayed some outlier behavior, plus the long fixations occurred with every verb except for ‘studieren’ (*to study*).

In fact, the “noise” in my data (i.e., outlier cases) showed a very interesting pattern: I found that the extremely long fixations (as measured by total time) occurred with irregular verbs (rather than regular verbs) more than 77% of the time (37.1% with e → i/ie vowel changes, and 40.0% of the time with a → vowel changes). The follow-up logistic regression analysis confirmed this trend. I take this to mean that these learners *noticed* the irregularity of the verbs, especially given the length of processing time and the *eye-mind assumption*. This supposition could be interpreted as supporting the traditional *Noticing Hypothesis*: Learners attended to the novel linguistic form in the input, the items were registered in working memory, and then became candidates for further processing and, eventually, long-term storage (i.e., learning). Note that the data in this study only lend support to the *Noticing Hypothesis* (the data in this study appear to fit, if you will, the hypothesis), but the data do not (and cannot) unequivocally prove the hypothesis is true. Also, it needs to be acknowledged that outlier (noticing) cases were rare. In fact, this “noticing phenomenon” occurred in only in 84 out of 1,032 overall cases (i.e., in 8% of all observations), which means that such processing behavior constitutes an exception rather than the rule. At the same time, it is remarkable that 75% of all participants produced a long fixation on some verb that fell into the outlier pool. This means that these “outlier” data do not stem from a small subset of participants only. In fact, some noticing of some verb occurred with more than three quarters of all participants.

One possible explanation for this finding is that if learners attend to some structure, they keep it in their working-memory, which is the place where a myriad of cognitive processes take place. According to N. Ellis (2005), working memory is “home of explicit deduction, hypothesis formation, analogical reasoning, prioritization, control, and decision making. It is where we

develop, apply, and hone our metalinguistic insights into an L2. Working memory is the system that concentrates across time, controlling attention in the face of distraction, [...]” (p. 337).

The controlling of attentional resources in working memory is the process that Gass (1997) coined ‘apperception.’ According to this notion, apperceived words or structures are attended to, and possibly recognized, due to one or both of the following two things: (a) the learner’s prior knowledge or experience, and/or (b) something salient in the word (i.e., a peculiarity that does not match the current representation in the learner’s mind). In this study, I collected data from extreme beginners of German who had not yet been introduced to stem-vowel-changing verbs. Thus, if apperception is occurring with these learners, it must be because they notice that the stem-vowel-changing verbs are different—they are unexpected. However, what learners have already established in their minds and what they expect is, of course, variable. Yet these data do appear to demonstrate that apperception is occurring in these learners.

Alternatively, the outlier cases (as well as the total time results of all observations) might reflect classic examples of learners engaging in *attention as selection* (Robinson, 2003), which influences the perception and the *selection* of stimuli that enter working memory (e.g., Carroll, 1999; Gass, 1997; Leow, 1993; Tomlin & Villa, 1994). Thus, selection as attention is a very important function of attention in general. According to Robinson (2003), attention as selection is likely to involve awareness,⁴⁸ but it is also quite possible that selective attention occurs without the learners being aware that they are selecting information from the stimulus array or that they are further processing of one piece of information over the other. “Selection of detected input happens [...], and attention is required for it to happen, but it need not (but very

⁴⁸ Yet, awareness is not *required* for attention as selection to take place.

often does) implicate awareness” (Robinson, 2003, p. 636). As such, there is no clear implication as to whether the viewing behavior of the outliers, which I tend to interpret as “selective attention,” involves awareness. This research did not employ any awareness measures, so future research studies may benefit greatly from employing a multi-method procedure that yields both data pertaining to the learners’ (focal) attention as well as data on their awareness, or lack thereof (Godfroid, 2010).

Research Question Two

Regarding RQ2, I hypothesized that the eye-fixation durations would not be predictive of learning gains related to stem changes because the learning of verb morphology such as stem-vowel changes is known to be difficult for learners with a language background that has limited morphology, such as English. I predicted, however, that learners would improve their production of verbs overall due to the exposure they had had during the treatment phase. It is important to recall at the outset that the analyses that were conducted to address RQ2 included the irregular verbs only.

No Association between Fixation Time and Learning of Irregular Verb Morphology

There was no association between the learners’ total time spent processing the *ich*-form and the *du/er/sie*-form and their learning gains. Given the fact that learners did fixate longer on the irregular verbs compared to the regular verbs (the main finding for RQ1), it’s possible to assume that these increased fixation times would be reflected in the learners’ performance on the posttest. Yet, no relationship was found. I speculate that there are three, possibly complimentary, ways to interpret this finding.

First, the limited saliency and redundancy of the target structure might not have lent itself to being noticed. Or taking this assumption a step further, the target structure may not have been

conducive for further processing following noticing—on average only noticing at the level of apperception occurred.

Second, the low developmental level of the learners (i.e., beginning learners) may have restricted their capacities to attend to the morphologic structure of stem-vowel changes. Below I discuss the findings in regard to these three premises in more detail.

And third, the irregular vowel change is a *rule-based* structure. That is, the stem vowel of a verb will change in the 2nd and 3rd person if it is a *strong* verb. On connectionist accounts, knowledge of a rule-based structure emerges gradually (e.g., N. Ellis, 2002; Dörnyei, 2009), so it is likely that the participants did not become aware of (albeit they displayed focused attention to) the possible existence of an underlying “rule” for vowel changes.

Non-saliency of Irregular Verb Morphology

As N. Ellis and Sagarra (2010) suggest, “early adult learning is characterized by a general tendency to focus on *lexical* cues because of their *psychological salience* resulting from their simplicity of form-function mapping and from learners’ prior first language knowledge” (p. 85, highlights are mine). At the early stages of acquisition, the (lack of) saliency of morphological cues affects the learners’ attentional processes in that they attend to lexical cues rather than morphological cues (Bell, 2006; Kintsch, 2005; Tsui & Fullilove, 1998; J. Williams, 2001; 2005). Beginning learners usually have processing difficulties in their L2, so they depend more on *contextual* cues found elsewhere in the sentence or derived from other sources, such as pictures, the situational context, or their background knowledge. As explained in Chapter 4, the experiment in this dissertation provided pictures that contextualized the sentences presented on the eye-tracker screens; that is, the pictures demonstrated the action verb that the participants

encountered in the sentences that followed. It is possible that some learners in this study might have relied on the pictures as a contextual cue—more so than linguistic cues only—when processing the meaning of the sentences, which may, in turn, have diminished their top-down processing of verb morphology and associated attentional foci.

Furthermore, while advanced learners and native speakers are able to perceive differences as in ‘lauft vs. läuft’, ‘fährt vs. fährt’, or ‘seht vs. sieht’) both visually and aurally, beginning learners of L2 German may not be. Put more carefully, beginning learner of L2 German (L1 English) are likely to be able to perceive phonemic contrasts between ‘lauft vs. läuft,’ but they may not do so under incidental processing conditions. That is, they may not attend to the relevant cues in the input. The learners’ decreased attention or failure to attend to the relevant morphological feature (i.e., a changed-stem vowel) might be exacerbated by the fact that these cues are redundant as the information in number and person is provided by an overt sentence subject plus the inflectional ending of the verb. Contrary to pro-drop Japanese and Spanish, for example, German is not a pro-drop language. In other words, the inflectional ending of verbs in German competes with lexical cues provided by the head/subject noun of the sentence. Beyond that, the vowel-stem change also competes with the inflectional ending of the verb, which also carries the information regarding number and gender. From the perspective of a beginning learner of L2 German, whether the sentence reads ‘Er läuft einen Marathon’ or ‘*Er lauft einen Marathon’ (*He runs a marathon*) may not create any meaning difference for the learner. As the beginning learners in this study had very limited knowledge about morphosyntactical structures in German, they likely attended to the most *salient* and *reliable* cues to comprehend the sentence. In fact, there is strong empirical evidence that (a) low salience, (b) low reliability, and (c) redundancy of morphological cues affect their learnability (DeKeyser, 2009; N. Ellis, 2006a,

2006b; 2011; N. Ellis & Sagarra, 2010; Goldschneider & DeKeyser, 2001; Jiang, 2007; Park, 2007; Tokowicz & Warren, 2010; Williams & Lovatt, 2003). Regarding the third factor, redundancy of morphological information, DeKeyser (2005) explained that it is the *lack of transparency* of the form-meaning relationship that poses the challenge to the learner, especially when learners are “left to their own resources instead of presented with a reasonably complete set of rules about form-meaning relationships” (p. 3), as it was the case in this study. What DeKeyser calls “resources” are the L1 cue mappings⁴⁹ that the learners have stored in their minds and which are changed and adapted only gradually through extensive L2 exposure (Bordag & Pechmann, 2007; N. Ellis & Sagarra, 2010; Parodi, Schwartz, & Clahsen, 2004). “Learners’ sensitivity to verbal morphology will therefore depend on the degree to which their L1 makes extensive use of it” (Ellis & Sagarra, 2010, p. 88).

Lee (1998) investigated how native speakers of English process Spanish verb morphology (subjunctive form) when reading. His research question was “Are comprehension and input processing affected by the morphological characteristics of the input?” (p. 37). Lee’s experimental groups (A and B) read texts in which for Group A verb morphology was substituted with infinitives [*le enseñar los modelos últimos (to show you the latest models)*] and for Group B verb morphology was substituted with a nonsense morpheme; that is, the appropriate subjunctive endings were replaced with -a -u. [*le ensenu los modelos ultimos (show you the latest models)*]. The control group read the same text with normal (unaltered) Spanish verb morphology [*le enseñe los modelos último (he shows you the latest models)*]. Learners

⁴⁹ In some cases, perhaps, also some L2 cue mappings from another foreign language the participants have learned in high-school, for example L2 Spanish.

comprehension of a passage was measured by a recall task,⁵⁰ and word identification task to determine how many words from the passage the learners could identify after reading it. Lee's hypothesis was that "if learners noticed, detected, or otherwise cognitively registered a form, then they would identify it in the recognition task" (p. 38). Contrary to his expectations, the results showed that the learners who were exposed to the *correct* subjunctive version recalled *significantly less* than those who read the infinitive or invented -u versions. Similarly, the comprehension scores of the passage with the *correct* subjunctive form were *significantly lower* than the comprehension scores of the passages that had infinitives and nonsense forms. Lee concluded that "[w]hen naturally attending to meaning, learners comprehend better when the forms in the input are *less complex* than when they are *more complex*" (p. 82). In addition, the learners' native language, English (which has limited morphology), might have affected their processing, or rather, non-processing of verb endings. Based on these findings, Lee argued that:

The subjunctive forms contain more linguistically encoded information than the other forms. VanPatten (1990) established that directing learners' attention to forms in the input detrimentally affected comprehension. On their own, without orienting directions, all participants detected the forms in the input. It may be that because the subjunctive forms were more varied and were linguistically richer, noticing and detecting them required more processing capacity than noticing and detecting the infinitives and invented forms. Such a draw on capacity would indeed affect comprehension (Hulstijn, 1989; VanPatten, 1990). (p. 41).

The learners in this dissertation study, mostly native speakers of English, might not have been inclined to attend to changes in verb morphology in the present tense, which was then reflected in their limited learning gains. It is likely that the learners paid attention to the verb

⁵⁰ Specifically, the learners were asked to recall in English as much as they could of what they just read. The learners were encouraged to recall not only main ideas but details, too (Lee, 1998, p. 38).

forms as a whole,⁵¹ but they failed to notice the relevant dimension of the input (i.e., the vowel change). In other words, it is likely that the learners were *aware* of the verb forms at a *global* level, but they failed to attend to the relevant dimensions of the input. The lack of attention to pertinent features may also be due to the rule-based nature of the target structure, which I will discuss in more detail in the following section.

Furthermore, the results of this study point to the theoretical existence of noticing thresholds (Sharwood-Smith & Truscott, in press). Whether focal attention leads to deeper processing that facilitates learning depends on the level of *activation* of awareness, which, in turn, appears to be dependent on the target structure (whether the feature carry important, meaningful information, whether they are based on a simple or rather abstract rule).⁵² As this study did not measure the learners' awareness, stronger conclusions cannot be drawn. Yet, given that prior research supported that learners demonstrating higher levels of awareness learn more than learners with lower levels of awareness (e.g., Bell & Collins, 2009; Broderick, 2005; Leow et al., 2008; Leow, 1997, 2000), I assume that learners who demonstrated some learning gains on the posttests were likely to have become more aware of the relevant dimensions of the input (verb forms) as their focal attention dwelled on the verb. Again, the lack of an awareness measure precludes drawing stronger, persuasive conclusions. Nevertheless, these data appear to suggest that, generally, adult learners of L2 German are unlikely to incidentally acquire correct

⁵¹ Recall that the interest area defined on the eye-tracker screen consisted of the whole verb, not just the critical vowel.

⁵² Recall that the cognitive science literature suggests that attention and awareness belong to differential cognitive circuits.

verb morphology in regards to stem-vowel changes⁵³ when reading. In fact, Ellis and Sagarra (2010) argued that adults need to *learn how to pay attention* to cues that are not part of their L1 if they are to succeed in L2 acquisition. They suggested learners' difficulty in attending to morphological cues can be overcome by "pedagogical interventions that make these cues non-redundant and more salient" (p. 102).

Based on my data, I surmise that global attention to a language form is not sufficient for learning, even if that language form is an exemplar of a particular target structure. While one could argue that the learners should have some understanding of what dimension in the input to pay attention to, it is also quite possible that the learners devoted much of their attention to the basic facts of German verb morphology (i.e., the verb-endings, which they had just learned). I will elaborate on this possibility in more detail below.

The majority of the learners demonstrated limited attention and limited learning in most cases *despite the input flood* and the fact that the learners were exposed the two *contrasting* (i.e., contrasting in their stem-vowel) verb forms via sentence pairs. There were a few exceptions, or outliers, however, which deviated from the general trend.

Another possible explanation of the finding is that the learners may have experienced the effects of their attentional constraints. Usually, learners (or people in general) have a limited capacity available for processing (e.g., Izumi, 2003; McLaughlin et al., 1983; McLaughlin & Heredia, 1996; see also references in the introduction of my dissertation), so the new L2 structures (i.e., the irregular stem-vowel changes) might have cognitively *overloaded* the learner,

⁵³ Stem-vowel changes are a fairly infrequent, non-salient and redundant type of verb morphology. Thus, it would not be justified to generalize this conclusion to other types of German verb morphology (e.g., inflectional endings), which are more frequent and more reliable in their form-meaning mapping.

who had only just been taught the regular simple present. This, in turn, may have negatively impacted L2 gain. If the developmental level of the learners had been more advanced, they might not have experienced this type of cognitive overload.

The Developmental Level of the Learner and Attention

The premise that working memory limitations restrict learners' attention to multiple features in the input has been well-established in the cognitive science literature (e.g., Baddeley, 2003; Cowan, 2005; Hasegawa, Carpenter, & Just, 2002; Just & Carpenter, 1992) and in the field of SLA (e.g., N. Ellis & Sinclair, 1996; Elman, 1993; Gass, Svetics, & Lemelin, 2003; Goo, 2010; Sagarra, 2007; van den Noort, Bosch, & Hugdahl, 2006). Is the efficacy of attention and working memory related to the developmental level of the learner? Ortega (2009) suggests that "as L2 proficiency develops the lag in working memory capacity between the L2 and L1 should become smaller." At the same time Ortega acknowledges that "less is known empirically about this widely held assumption" (p. 91). Several SLA researchers (B. McLaughlin & Heredia, 1996; Philp, 2003; Pienemann, 1984, 1989) have suggested that advanced, or more experienced learners may benefit from the increasing automaticity that comes with repeated practice, which then frees up attentional resources that can be allocated to other task dimensions. A learner's developmental level is then associated with their "readiness" to process a given aspect of the input successfully (e.g., Spada and Lightbrown, 1999).

Farrar (1990), in a study with children, found that recasts are most effective when the children's mental state is "ready" to acquire morphemes. Looking at L2 learners, there might also be a "prime time" in which learners are ready to attend to verb morphology. Mackey (1999) and Mackey and Philp (1998), for example, found that question formations are learned in a predictable sequence that follows a certain pattern of their interlanguage development. It is

possible that learners are unlikely to attend to features in the input that are beyond their current interlanguage level.⁵⁴

Rule-based versus Item-based Learning

Robinson (1996) and N. Ellis (2002) hypothesized that noticing leads to some memory representation of the noticed form and that different *types* of L2 forms may be represented differently in the brain; that is, associatively (the learning of a *system* based on associations) or symbolically (the simple memory of instances or items). Associative learning means that at the beginning of the learning process, the learner stores instances or exemplars (i.e., “items”) he or she encountered in the input. As the learner processes the items more frequently, their brains become sensitive to the properties of the input, and they “learn” from them. N. Ellis (2006) explains this more specifically: Over time, the brain engages in a tally of the overall frequency of each form and the likelihood of co-occurrence with other forms. This brain activity is assumed to be an implicit process of which the learners are unaware. In this dissertation study, when the learners read the forms ‘er wäscht’, ‘er läuft’, ‘du fährst’, or ‘du siehst’, ‘er nimmt’, and ‘du gibst’ they might have taken a mental note of—or even temporarily *stored*—the instances of stem-vowel changes. Yet, the low frequency of occurrences during the reading activity in this study most likely did not suffice for the learners’ mind to extract information about the properties and the frequencies in order to generate a rule based on these stored or noted instances. Similarly, Shanks and St. John (2005) argued that when participants demonstrate learning but do not demonstrate awareness at the level of understanding of some underlying rule, this is “because their knowledge consists of instances or fragments of the training stimuli rather

⁵⁴ It is important to note, however, that Robinson (2003, 2009), a pioneer of research on attention in SLA, rejects the entire notion of developmental readiness because it does not explain anything.

than rules” (p. 367). In regard to this study, I assumed that rule-based learning (i.e., the actual explicit *understanding* of a an underlying rule) would not occur with the learners given that the experiment-treatment phase created the single exposure for the learners. Recall from Chapter 1 and 2 that I conceptualized noticing a process where attention and awareness (two distinct processes in the mind) come together (initially in working memory), whereby the attention paid to a new linguistic structure exceeds a threshold so that this “novelty” is held in working memory for long enough to increase the (initially low) level of awareness and to gain access to short-term memory. If this noticing process (attention and awareness) occurs with sufficient *frequency*, the linguistic item will eventually build a representation/become a stored exemplar in long-term memory. I did not conceptualize or operationalized noticing at the level of *understanding*. Although I did not employ a measure of awareness, I conducted this study under the premise that the *type* of learning (i.e., item-based learning versus rule-based learning), which could be drawn for the pre-and posttest results, may provide insight into the *level* of awareness engaged during input processing. If the learners demonstrated correct production of the stem-vowel changes in irregular verbs, they may have become aware of and may have learned the *system* of stem-changing verbs (i.e., system-based learning). However, I assumed that if I find evidence of learning, this would be a reflection of *item*-based learning; that is learning without rules. Ortega (2009) explains learning without rules as follows:

“Learning without rules leads to the formation of memories of instances that can be accessed more easily allowing for faster performance, but without knowledge that can be generalized to new instances. That is, without the initial provision of rules (without an explicit learning condition), learning is bottom-up (i.e., data and memory driven), and it does not lead to knowledge of a systematic rule of some kind” (p. 102)

In this study, the learners’ fixation times on a given exemplar of the vowel change might be an indicator of *attention* and temporal storage of this exemplar, but the fixation time measure

coupled with the post-test measure yields little information on the actual learning that is, whether the learners were able to generate the vowel-change “rule” based on these stored exemplars, and whether they were able to generalize across items. Future studies will have to probe into rule-based versus item-based learning, an issue I address in the section on Limitations and Directions for Future Research. The rule-like nature of the target structure may also have some implications for the “validity of a retrospective awareness measure (such as stimulated recall), because later gained awareness may be projected on earlier (unaware) cases of processing. I also address the implementation of an “awareness-measure” in the section on Limitations and Directions for Future Research.

Parabolic Relationship

Recall that in this data a quadratic association between total time gain and holistic learning gains (i.e., learning gains that were related to the overall production of the verb, not only to the stem-changes) was observed. Consequently, longer fixations on *either* form, the non-changing *ich*-form or the vowel-changing *du/er/sie*-form, predicted learning well. This finding suggests that the learners did attend to the sentence verbs in general, just not sufficiently to the stem-vowel changes, possibly because of their lack of salience, absence from the L1 as a relevant cue, and because of the learners’ low level of automatization in general. When a learner fixated *longer* on a verb form, regardless of whether it was the *ich*-form or the *du/er/sie*-form, his or her comparatively longer fixation was associated with an overall improvement in producing these verbs. This corroborates the eye-mind assumption.

Based on the premise that there is a close link between overt attention (i.e., the eye-location) and covert attention, that is, a close connection between what one looks at and what one focuses on mentally, learner’s eye-fixations most likely coincided with their focused attention.

The question then is not necessarily *which* verb form they focused on, but *that* the learners showed extra processing activity (i.e., longer fixation times). Learners may not have attended to the stem-vowel change specifically, but to the verb in its entirety and, perhaps, to the inflectional ending of the verb. A closer look at the German 101 syllabus for the Fall semester of 2010 coupled with an examination of the learners' textbook revealed that the learners had learned the verb inflections one to one and a half weeks before the data collection started. They had practiced the correct inflectional ending of the verbs in several drills as well as in their homework activities. Thus, the learners may have paid attention to the very feature on which they have received prior explicit instruction. The holistic learning gains, in most cases, included improvements in the production of the German verb endings. For example, for the verb *sehen* (to see) the participants frequently wrote **“er sehe”* or **“er ist sehen”* on the pre-test, and **“er seht,”* so their holistic performance improved despite the missing the stem-vowel change (the correct form would be *“er sieht”*). One possible interpretation is that the learners' prior experience (via formal instruction) with German verb inflection might have induced the learners to focus their selective attention toward the *endings* of the verbs, rather than on the stem. Put differently, the participants may have failed to require the relevance of the stem vowel to their processing task, thus, in effect, failing to attend (at the level of learning) to this stimulus dimension. Given the learners' low developmental level, it is unlikely that they would have been able to divide their attention to more than one linguistic stimulus or to more than one dimension of the stimulus array (Eysenck & Keane, 2010).

It is important to note that Wicken's Proposed Structure of Processing Resources (2007) (see Chapter 1) makes different predictions: Resource competition exists within, but not between separate attention pools. That is, when language stimuli are processed via one modality only, as

in reading (i.e., visually), there supposedly is *no* divide of attentional resources. As such the findings of this dissertation study support Rosa and Neill's (1999) view: The multiple-resource model does *not* specify where and how attention differentiates between inputs from two language modalities and what happens if one modality requires more attentional resources than others.

Last, it is also important to note that the relationships observed in this study, while significant, were not very strong, so caution is required when generalizing these particular findings to larger populations. Nevertheless, although the relationships found were generally small (e.g., $r = .13$), they are still informative. If learners and language instructors realize that the beginning stages of the acquisition of verb-morphology are, in part (here, for example, a 13% part) related to attending to these structures, they might enhance their instructional techniques and teach their students to pay attention to morphemes in the language.

The Verbs Sehen and Fahren

While no linear relationship was found between the fixation times on the verbs and the learning gains for those verbs, the verbs *sehen* (to see) and *fahren* (to drive) differed from the overall picture: (1) Fixations on *sehen* were positively correlated with *holistic* learning gains for *sehen*, and (2) fixations on *fahren* were positively correlated with stem-change gains for *fahren*. There is no ready explanation to address this finding, but it is possible that the learners' prior exposure to the verbs impacted their processing. *Sehen* and *fahren* had been covered at several occasions by the German textbook up to the point of data collection, plus these verbs are also used in the subsequent textbook chapter⁵⁵ to explain the phenomenon of vowel-stem changes in irregular verbs. Since *sehen* and *fahren* are used as example verbs to explain the vowel change,

⁵⁵ The subsequent text book chapter had not been covered by the time the data collection was completed

there might be something prototypical about them. The comparatively high frequency of these verbs (*sehen* = 115,440 per 100 million tokens, and *fahren* = 35,278 per 100 million tokens, while the average frequency of all verbs was 28,722) might be related to the fact that these verbs stood out in regard to the linear relationship between fixation time and learning, which was not found with the other verbs.

Limitations and Directions for Future Research

While the results from this dissertation research make an important contribution to the field of SLA in informing SLA theory and pedagogy, there are some limitations to this research that need to be acknowledged. In the first part of this section, I address the limitations associated with (a) learner-internal factors, (b) the stimuli, (c) the interpretation of the results, and (d) the design of the study. I will conclude this section by suggesting an agenda for future research based on the questions that arise from this study.

Limitations

As explained above, noticing by the learners might have been affected by the previous classroom instruction learners have received (e.g., Doughty & Williams, 1998; Norris & Ortega, 2000; Rosa & Leow, 2004; J. White, 1998), learners' processing resources and individual differences in the amount of resources available (Dörnyei, 2005, 2009; Grigorenko, Sternberg, & Ehrman, 2000; Miyake & Friedman, 1998; Robinson, 2002; Sawyer & Ranta, 2001; Skehan, 1998), and learners' (developmental) readiness to pay attention to morphological details of the (German) language (Adams, 2003; Philp, 2003; J. Williams, 2001). While controlling for all of these learner-internal factors is perhaps beyond the scope of this dissertation research, future research should take these into consideration as much as possible.

In addition to the learner-internal factors, the eye-tracking data obtained in this study might be influenced by interfering stimuli variables; or as Godfroid et al. (2010a) phrased it: “The reader’s viewing behavior is determined by a complex interplay of factors” (p. 177). For example, the frequency and saliency of the input might have influenced participants’ attentional foci (Bardovi-Harlig, 1987; Carroll, 2006c; DeKeyser, 2005), and the participants’ familiarity with the words used in the stimuli might have influenced their eye-movements (i.e., focal attention). In addition to familiarity, as explained in Chapter 3, the participants’ eye-movement might have been influenced by word length, predictability, and/or frequency (Kliegl et al., 2004; Kliegl et al., 2006). While the counterbalancing conditions implemented in this study controlled for these influencing factors, they might not have been eliminated entirely.

Regarding the predictability of the words or sentences, some learners might be able to predict the content of the sentences. Recall that the stimuli were developed with help of the German textbook used in class. Consequently, some learners might have been familiar with some of the content of the sentences. However, at the same time, greater semantic accessibility might enable the learners to focus more on the morphological elements targeted in this study (N. Ellis & Schmidt, 1998).

A point worthy of special attention is the inconsistency of the fixation time findings for the fixation time measures: total time, first fixation, and gaze duration. In particular, the results for gaze duration did not coherently follow the patterns found with the dependent variables of total time and first fixation duration. Recall that with total time, I found an effect of verb type with both the regular, trimmed data set as well as the outliers. With first fixation duration my ANOVA did not evidence any statistical differences with the regular data set, but the analyses of outliers showed the same trend as for total time. Since gaze duration is also an ‘early’ reading

measure, I expected the results for gaze duration to mirror those found for first fixation duration. Yet, my findings for gaze duration were not significant overall. The differing results of gaze duration might be rooted in the design: the baseline *ich*-sentence was always presented first, followed by the critical *du/er/sie*-sentence. Thus, the learners might have anticipated the subject and verb of the second sentence, which is likely to have impacted their processing and eye-movements. Reichle, Rayner and Pollatsek's (1999) *E-Z Reader model of Eye-movement Control* might explain the different findings of first fixation duration and gaze duration. The model proposes that readers process a word at two levels: First, they (a) make a preliminary familiarity check where they encode orthographic and phonological information and check whether lexical access is immanent, and they (b) encode the word's meaning (lexical access). With this model in mind, consider that gaze duration differs from first fixation in that the former includes *all* fixations in an interest area during first pass reading (not only the first fixation) while first fixation, as its name suggests, reflects only the first fixation on a word (so the value in milliseconds is usually somewhat smaller). With native readers, first fixation and gaze duration usually reveal similar results, as the two levels of processing happen (almost) simultaneously. With beginning learners, however, the encoding of a word's meaning (level 2) as well as the preliminary familiarity check (level 1) might take a millisecond longer as they have just learned, to some extent, the verbs (vocabulary) one to three weeks earlier. Thus, the learners' gaze on the verb might have lasted slightly longer when they read the first sentence as opposed to the second *du/er/sie*-sentence where (according to the E-Z Reader Model) the processing on the second level (lexical access) was accelerated. Explained in a different way, gaze duration did not evidence an effect of verb regularity as learners had already retrieved the meaning of the verb (i.e. lexical access had already been completed) by the time they read the second verb. Plus, a

familiarity check (level 1) might have been speeded considerably due to learners' prediction. The accelerated access might have been minimal but detectable when considering at all 1,032 observations/cases.

On the whole, the E-Z Reader Model is only one way of attempting to explain the difference between gaze duration and first fixation duration findings, and how the difference might be rooted in the design (i.e., that the baseline *ich*-sentence was always presented first). This drawback notwithstanding, the absence of an effect of verb regularity on gaze duration, combined with the effects for first fixation duration and total time, suggests that part of the processing that distinguishes the noticing of irregular verbs from the ordinary reading of regular verbs took place during *second* pass reading. Future analyses should consider second pass time and/or regression measures. Likewise, future investigations of the noticing construct would be well advised to include some later eye-movement measures such as second pass time and the occurrence of regressions in their analyses. In addition, future research is needed because the results stemming from these data do not suggest definitive conclusions on beginning language learners processing during reading.

Another limitation related to the design of the study lies in its inability to measure learning gains *over time*. A delayed posttest could have shown how the learning process unfolds over time. However, I chose not to add a delayed posttest to my design for several reasons. Firstly, with a delayed posttest, I would not have been able to discriminate between learning caused by the treatment and learners' attentional processing of the treatment and learning associated with the immediate posttest. My sense of caution is much supported by Hulstijn (2003), who wrote:

In evaluations of learning experiments one must bear in mind two considerations: (i) with an immediate post-test, the researcher is able to measure the effect of cognitive processes during the learning session - nothing more, nothing less; [...] [Research] aimed at addressing questions concerning the effect of cognitive processing during a learning session in which words are presented for the first time, requires only an immediate post-test. Inclusion of delayed post-tests in such research would not make sense, because it would not be possible to differentiate the extent to which performance on delayed post-tests is affected by processing during the experimental learning session or by processes (if any) after that session. (p. 372, insertions are mine).

A last point to consider is the extent to which the research itself might affect participants' behavior. How learners behave in the laboratory may differ from how they behave in everyday life, which is likely to affect ecological validity.⁵⁶ Eysneck and Keane (2010), for example, explain "in the real worlds, people are constantly behaving so as to have an impact on the environment [...]. Thus, the responses that people make typically change the stimuli situation. [...] That is to say, the sequence of stimuli the experimenter presents to the participant is not influenced by participants behavior, but rather is determined by the experimenter's plan" (p. 4). That is, learners' participation in the study, especially the pre-tests, may contribute to raising their awareness to formal aspects of the German language. Consequently, the number of noticing events may have been higher because of the sheer fact that this was a laboratory study, thus introducing a potential limit to external and ecological validity of my findings.

"The participants focus of attention in most research is determined by the experimenter's instructions. As a result, relatively little is known of the factors normally influencing the focus of attention, relevance of stimuli to current goals, unexpectedness of stimuli, threateningness of stimuli, intensity of stimuli, and so on. This is an important limitation because we could not predict someone's cognitive processes and behavior in most situations without detailed knowledge of the factors determining attentional focus. (Eysenck & Keane, 2005, p. 4)

⁵⁶ Ecological validity refers to the extent to which the findings are applicable to everyday life.

This note of caution notwithstanding, my instructions to the learners were very open and non-constraining as to what participants had to pay attention to or what the goal of the experiment was. In addition, this note of caution applies to any study that is psycholinguistic or experimental in nature, this potential drawback is just usually not addressed.

The last limitation of this study concerns the use of a correlation analysis to address RQ2—whether increased attention to irregular verbs is associated with subsequent learning of those irregularities (i.e., stem-vowel changes). While the correlation analysis in this study revealed whether there was a systematic relationship between the fixation time measure (i.e., total time gain) and learning (i.e., improvement on the posttest), it was unable to reveal whether the fixation durations on the irregular verbs would *predict* improvement of the production of irregular verbs. Future analyses of this relationship should be based on more sophisticated statistical techniques, such as a GEE (Generalized Estimated Equation) analysis (Fitzmaurice, Laird & Ware, 2004; Hardin & Hibe, 2003). A GEE can be used to predict the improvement of the production of the irregular verbs based on the fixation times, after accounting for the prediction of each verb individually. Furthermore, a GEE could also reveal the probability of improvements on posttest. Godfroid (2010), for example, employed a GEE to analyze her data and found that “for every additional 1000 milliseconds that a participant looks at a word, the probability that she will recognize it on the post-test increases by 8%” (p. 172). RQ2 in this study could have been addressed with a more detailed analysis of the data. This would have provided a more comprehensive and intricate picture of the relationship between processing time and learning.

Suggestions for Future Research

This dissertation study paves the path for a challenging, yet fascinating, research agenda. While this dissertation study provides insight into the learners' attentional foci on verb morphology during reading, it does not allow for inferences about the *level of awareness* learners experienced at the time they attended to verb forms in the sentences they read. The design of this study also does not provide information of the *type* of learning learners might engage in (i.e., item-based versus system-based learning), so future experimental eye-tracking studies should look at different grammar structures and how these might be attended to, processed, and learned differently. Future studies should investigate whether learners are able to generalize the target structure to new items, and how the ability to generalize is dependent on the nature of the grammar structure. It would also be important to explore how different forms determine the learners' awareness of these forms. With a target structure that has a fairly straightforward underlying rule, the measure of awareness should happen concurrently (i.e., via think-alouds). Regarding grammar structures that are rather abstract, it might be interesting to trace the learners' awareness of such structures over time, as the level of awareness is likely to increase with frequent exposure to the relevant form. Qualitative investigations of learner output or stimulated recall might be appropriate research methodologies to investigate such issues.

This dissertation research opens up the crucial question as to whether the learners were *aware* of the stem-changes. More specifically, it remains unknown whether the learners were aware of the differences in the stem vowel between the 2nd and the 3rd person singular and the *ich*-form or the infinitive, or whether the learners were aware of the relevance of the vowel in the stem without being able to explain this irregularity. While, according Schmidt (1990, 1994a, 1995, 2001), attention and awareness are isomorphic, that is, two parts of the same coin or two

concepts that are part of the noticing construct, several scholars do not support the premise that awareness coexists with attention (e.g., Lamme, 2003). In order to tease apart the construct of attention and awareness and to better understand the roles these cognitive mechanisms assume in noticing, qualitative awareness measures that complement quantitative attention measures are needed (Godfroid, 2010). In her dissertation conclusion, Godfroid (2010) suggested ways in which attention and awareness can be measured in one research study: “our understanding of noticing would benefit tremendously from triangulating the eye-tracking data with some qualitative awareness measure, such as retrospective verbal reports [...] or stimulated recall” (p. 214). To reiterate this quote, future research should implement a carefully thought-out, mixed methods or multi-method design that enables the researcher to distinguish between the constructs of attention and awareness as separate components of noticing. Such an approach would incorporate quantitative (e.g., eye-fixation counts) and qualitative (e.g., introspective) measures. Because of the reactivity issue (Chapter 2), the possibility of using qualitative, introspective measures on awareness to complement attention data (gathered via eye-movement measures) warrants some elaboration. With introspective methods (e.g., stimulated recall or retrospective reports) the data come from participants' own statements, that is, the data reflect participants' subjective reports about the way they perceive and understand information. As such, introspective methods are frequently used as supplements for data gathered based on learners' behavior (Bowles, 2010).

Simulated recall methodology is used to explore learners' thought processes (or strategies) at the time of an activity or task. This is achieved by asking learners to report those thoughts after they have completed a task or activity. In regard to an eye-tracking experiment, where attention is measured via eye fixations on words, the stimuli that the participant has seen

during the eye-tracking experiment could serve as the stimulus for stimulated recall.

Alternatively, a video showing the eye-movements of the relevant learner while he or she was reading could also serve as a strong and fine-grained stimulus for recall,⁵⁷ however, one important concern is that the use of a video of the eye-movements during reading might bias the recall of the thought processes the learners had during the treatment task, resulting in *prompted* rather than *stimulated* recall. Whether this is the case or not is yet to be explored. A future study should investigate the effects of using such a fine-grained stimulus (eye-movement videos) as a support for recall.

If the eye-movement videos could be shown not to affect the validity of stimulated recall methodology, this technique would have important advantages. The appeal of using such a fine-grained stimulus relies on the eye-mind assumption (e.g., Godfroid, 2010; Just & Carpenter, 1980; Reichle, Pollatsek & Rayner, 2006). As explained above, eye fixations during reading can be taken as a reliable index of the learner's focus of attention at a given time: There is a close link between what learners look at during reading and what they are actually attending to at that moment. On this basis, the use of the eye-movement videos as a support for stimulated recall should truly stimulate the learner to recall their thought processes. In fact the psycholinguistic eye-tracking literature cited in this dissertation is built on the assumption that the eye-tracking data is a window to the mind, and that cognitive processes *did* take place during reading. These arguments notwithstanding, there is a small possibility that eye-movements can be guided by

⁵⁷ Although video clips of learners' eye-movements of learners have not been implemented as a stimulus for recall to date, traditional stimuli for retrospective recall have also consisted of some auditory or visual recording of learners performing a given task. Examples of recall prompts include a video of a learner writing an essay or underlining words while reading a text. The eye-movement videos would likewise be a rendering of a learner carrying out a given task (i.e. reading) – nothing more, nothing less. What differs is the amount of detail that is provided by an eye-movement video stimulus as compared to, say, a video of someone writing an essay.

automatic (i.e. unconscious) processes.⁵⁸ So, a risk is that by showing learners videos of their eye-movements may make increase their awareness of some feature in the input. Whether this might be the case it has yet to be explored. There is a need for future mixed methods research that evaluates the stimulated recall protocols received from two stimuli: (1) eye-movement videos, and (2) the eye-tracking stimulus without the eye-movement showing. If a difference in learners' responses is evidenced, careful decisions need to be made as to which stimulus to use for future eye-tracking research that also aims to implement a measure of awareness.

Another possible methodological pitfall with regard to the joint use of eye-tracking and stimulated recall lies in the implementation of a posttest. If the stimulated recall session follows the post-test, learners' responses are likely to be influenced by their testing experience given its overt focus on the target structure. Learners' might report on what they believe the researcher wants to hear rather than what they were truly thinking or were aware of at the moment when they were performing the original language task. If the post-test was implemented after the stimulated recall session, the post-test would also measure the learning that happened due to the stimulated recall session. Thus, employing a stimulated recall session and a post-test in succession is less than ideal because of potential design flaws. Godfroid (2010) suggested a mixed within- and between subjects design for investigating the relative contributions of attention and awareness to learning: Half the learners take the post-test and the other half participate in the stimulated recall session. With such a design, the stimulated recall data (i.e., measure of awareness) could be triangulated with the eye-movement data (measure of attention), and the post-test data (measure of learning) could also be related to the attention measure. The

⁵⁸ Some aspects of eye-movements (e.g., planning, executing saccades, or making regressions) are unconscious. However, the eye-movement records are full with information

administration of the stimulated recall sessions and the post-tests would not interact because the data from each would be collected from different individuals. However, they could be linked via a common frame of reference: the attention data.

In conclusion, I am citing Robinson (2008) who stressed the necessity for methodologically solidified research to gain a deeper understanding of the roles of attention and awareness:

“Future research will likely adopt increasingly sensitive measures of the contents of awareness, and explore new methodologies for operationalizing these. Neurophysiological measures of physical changes in brain states will also be used increasingly to complement the behavioral and introspective methods for studying the relationship of attention and awareness to learning.”
(p. 142)

CHAPTER 7 CONCLUSION

This dissertation examined what it means to attend to and “notice” a specific linguistic form, a fundamental issue that has captivated many researchers over the past two decades. This study showed that L2 learners of German do attend to novel linguistic stimuli. The amount of attention and the accompanying level of awareness are very likely to be influenced by the learners’ L1, their prior knowledge, their attentional capacity, their developmental level, and the grammar targeted. This overall conclusion is much in line with what Schmidt, in 2001, stated:

[A]ttention is a crucial concept for SLA. The allocation of attention is the pivotal point at which learner-internal factors (including aptitude, motivation, current L2 knowledge, and processing ability) and learner-external factors (including the complexity and distributional characteristics of input, discoursal and interactional context, instructional treatment, and task characteristics) come together (p. 12).

One unique finding in this study was the identification of *noticing outliers*; that is, instances in which learners attend to new or unfamiliar features in the input, which I take to reflect more extensive, deeper processing and, possibly learning. Generally, however, this study did not evidence an association between fixation time and learning of *irregular* verb morphology. Yet, there was a relationship between longer fixation times, regardless of whether they occurred on *stem-changing* verb forms or not and the acquisition of *regular* verb morphology, as reflected in the overall improvement in written verb production. Building on Robinson (1995, 2003) and Godfroid (2010), I propose that increased processing times indicate that an item is registered and processed in working memory. This cognitive activity is accompanied by some level of awareness. This stresses the importance of attention—be it attention to some specific morphological or syntactic form or a vocabulary item as a whole—in learning a foreign language as an adult.

Pedagogical Suggestions

This conclusion lets me turn to pedagogic concerns. Teaching goals and practices should include didactic strategies that enhance learners' attention and awareness via incidental focus on form (Loewen, 2005) or forms (Long, 1991), or by implementing consciousness-raising tasks (Fotos, 1994; Sharwood-Smith, 1997). Beginning adult learners need to learn to pay attention and, likewise, they need to be taught *how to pay attention and notice* and *what to pay attention to* if they are to succeed in learning a foreign language (N. Ellis & Sagarra, 2011). Skehan (1998) reviewed several psychometric studies in which different groups were taught a specific grammatical feature via different methodologies: (a) explicit rule giving versus highlighting of the target grammatical feature, (b) encouraging effective input processing versus requiring much output practice, (c) rule oriented instruction versus meaning oriented instruction, and (d) traditional (explicit rule giving) versus consciousness-raising treatments (p. 62-66). Based on his review, Skehan concluded that foreign language instructors should put all efforts into developing learning tasks that are both meaningful and which encourage *attention* to specific forms. Merely having to comprehend something is not sufficient for successful language learning at the adult age (Schmidt, 1983). Ellis (2003) has described such tasks as *focused* in contrast to unfocused tasks that are designed purely for communication. He describes focused activities as (a) structure-based production tasks, (b) comprehension tasks, and (c) consciousness-raising tasks (p. 151). Pedagogical activities that help the learners pay attention to a certain form need to be *meaningful*; for example, communicative tasks that require the use of the target form help develop the learners' awareness of grammar structures (e.g., R. Ellis, 2003; Robinson, 1996). Thus, foreign language instructors need to know of the important role that attention plays in language acquisition: Any contact their learners make with the target language has potential to

trigger cognitive activity that is crucial for acquisition, and much can be done to increase the likelihood that such cognitive processing will indeed occur.

The Contributions of This Study to SLA Theory

In addition to the implications of my findings for the development of pedagogic materials and instructional techniques, the results of this study also make an important contribution to SLA theory. Most importantly, this study supports the notion that learners have a *selective* attentional filter. The selection of stimuli that is attended to is determined by the learners' L1 processing habits, the stimuli available for processing, the learners' prior knowledge, and contextual cues that may determine what stimuli pass through the selective filter. Furthermore, the findings in this study give much prominence to the following questions: (a) How much, if any, of the learners' awareness is involved when they pay attention to form? (b) What is it exactly that learners are noticing about a language form and what triggers their noticing? There are no ready answers to these questions. It is an important task of future SLA research to tease apart the noticing construct to determine how attention and awareness come together in noticing (Godfroid, 2010). My suggestion is to triangulate the eye-tracking data (measure of attention) with measures of awareness and learning (for example, stimulated recall). While the design of such a research study poses a methodological challenge to SLA researchers, it is a very important task to accomplish. The field of SLA needs to work with a more nuanced and empirically-based understanding of the noticing construct. Based on this refined knowledge, we will be able to better link acquisition theory to pedagogical practice.

APPENDICES

APPENDIX A: Background Questionnaire

1. Participant Number: _____ 2. Gender: M____ F____ 3. Age: _____
4. Native Language (i.e., language spoken at home): _____
5. Are you of German heritage? (circle one) YES NO
If yes, do you speak German with any family members? (Please state who: parents, grandparents, aunts, uncles, etc.) _____
6. How many years have you studied German in the following environments?
Elementary/Middle School _____ College _____
High School _____ Private Language School/Tutoring _____
7. List ALL of the college German courses that you have previously taken or are currently taking. _____
8. At what age did you start to learn German? _____
9. Besides German, what other languages have you studied? (Please write languages, number of years studied, and proficiency.) _____
10. Have you ever been to Germany? (circle one) YES NO
If YES, how long did you stay there? _____
Did you study there? Please explain. _____
Did you live with German people? _____
11. How many hours a week do you spend using German outside of the class to...
do homework: _____ listen to German (music, films, language cassettes): _____
study for tests: _____ read for pleasure in German (books, stories, etc.): _____
speak in German: _____ write in German (journals, letters) _____
12. How would you assess your proficiency in German? (Circle one for each category.)
Listening: Beginner Low Intermediate Intermediate High Int. Advanced
Speaking: Beginner Low Intermediate Intermediate High Int. Advanced
Reading: Beginner Low Intermediate Intermediate High Int. Advanced
Writing: Beginner Low Intermediate Intermediate High Int. Advanced

APPENDIX B

Complete List of Irregular Verbs Used in this Study

Table 7: Verbs with Stem Changes from e → i/ie

	sehen (to see)	lesen (to read)	essen (to eat)	geben (to give)	Nehmen (to take)	sprechen (to speak)
ich	sehe	lese	esse	gebe	nehme	spreche
du	s <u>i</u> ehst	l <u>i</u> est	isst	g <u>i</u> bst	n <u>i</u> mmst	spr <u>i</u> chst
er/sie*	s <u>i</u> eht	l <u>i</u> est	isst	g <u>i</u> bt	n <u>i</u> mmst	spr <u>i</u> cht
wir	sehen	lesen	essen	geben	nehmen	sprechen
ihr	seht	lest	esst	gebt	nehmt	sprecht
sie	sehen	lesen	essen	geben	nehmen	sprechen

*es (it) is also part of the third person but not included here for spacing reasons.

Table 8: Verbs with Stem Changes from a → ä

	fahren (to drive)	fangen (to catch)	schlafen (to sleep)	waschen (to wash)	laufen (to run)	Tragen (to carry)
ich	fahre	fangen	schlafe	wasche	laufe	trage
du	f <u>ä</u> hrst	f <u>ä</u> ngst	schl <u>ä</u> fst	w <u>ä</u> schst	l <u>ä</u> ufst	tr <u>ä</u> gst
er/sie*	f <u>ä</u> hrt	f <u>ä</u> ngst	schl <u>ä</u> ft	w <u>ä</u> scht	l <u>ä</u> uft	tr <u>ä</u> gst
wir	fahren	fangen	schlafen	waschen	laufen	tragen
ihr	fahrt	fangt	schlaft	wascht	lauft	tragt
sie	fahren	fangen	schlafen	waschen	laufen	tragen

*es (it) is also part of the third person but not included here for spacing reasons.

APPENDIX C.

Table 9: Sentence Stimuli by Group

Group I:	baseline	critical
	Ich sehe das Meer.	Du siehst die Natur.
	Ich spreche mit dem Lehrer.	Du sprichst am Telefon.
	Ich esse gern Brot.	Du isst oft Müsli.
	Ich nehme ein Buch.	Du nimmst das Geschenk.
	Ich lese die Zeitung.	Du liest ein Buch.
	Ich gebe dir den Pass.	Du gibst mir das Geschenk.
	Ich fahre mit dem Zug.	Er fährt mit dem Auto.
	Ich wasche meine Haare.	Er wäscht das Fenster.
	Ich schlafe auf der Couch.	Er schläft in dem Bett.
	Ich laufe im Sand.	Er läuft einen 5km Lauf.
	Ich trage mein Gepäck.	Er trägt das Fahrrad.
	Ich fange die Frisbee.	Er fängt einen Fisch.
	Ich wandere gern allein.	Er wandert am Nachmittag.
	Ich koche um sechs Uhr	Sie kocht das Mittagessen.
	Ich studiere gern Medizin.	Er studiert immer Mathematik.
	Ich trinke öfters Kaffee.	Sie trinkt viel Wasser.
	Ich segle mit meinem Freund.	Sie segelt manchmal allein.
	Ich zeige auf die Tafel	Sie zeigt auf die Uhr
	Ich singe nicht so gern.	Du singst ein schönes Lied
	Ich kaufe das Gemüse.	Du kaufst ein Brot.
	Ich lache über meine Schwester.	Du lachst über den Lehrer.
	Ich begrüße den Professor.	Du begrüßt deine Familie.
	Ich schreibe eine Email.	Du schreibst an die Tafel.
	Ich gehe in die Stadt.	Du gehst zur Universität.

Table 10: Sentence Stimuli by Group (a)

Group II:	baseline	Critical
	Ich sehe die Natur.	Du siehst das Meer.
	Ich spreche am Telefon.	Du sprichst mit dem Lehrer.
	Ich esse oft Müsli	Du isst gern Brot.
	Ich nehme das Geschenk.	Du nimmst ein Buch.
	Ich lese ein Buch	Du liest die Zeitung.
	Ich gebe dir das Geschenk	Du gibst mir den Pass.
	Ich fahre mit dem Auto.	Sie fährt mit dem Zug.
	Ich wasche das Fenster.	Sie wäscht die Haare.
	Ich schlafe in dem Bett.	Sie schläft auf der Couch.
	Ich laufe einen 5km Lauf.	Sie läuft im Sand.
	Ich trage das Fahrrad.	Sie trägt das Gepäck.
	Ich fange einen Fisch.	Sie fängt die Frisbee.
	Ich wandere am Nachmittag.	Er wandert gern allein.
	Ich koche das Mittagessen.	Sie kocht um sechs Uhr.
	Ich studiere immer Mathematik.	Er studiert gern Medizin.
	Ich trinke viel Wasser.	Er trinkt öfters Kaffee.
	Ich segle manchmal allein.	Sie segelt mit dem Freund.
	Ich zeige auf die Uhr.	Du zeigst auf die Tafel.
	Ich singe ein schönes Lied.	Du singst nicht so gern.
	Ich kaufe ein Brot.	Du kaufst das Gemüse.
	Ich lache über den Lehrer.	Du lachst über deine Schwester.
	Ich begrüße meine Familie.	Du begrüßt den Professor.
	Ich schreibe an die Tafel.	Du schreibst eine Email.
	Ich gehe zur Universität.	Du gehst in die Stadt.

Table 11: Sentence Stimuli by Group (b)

Group III: baseline	critical
Ich sehe das Meer.	Er sieht die Natur.
Ich spreche mit dem Lehrer.	Er spricht am Telefon.
Ich esse gern Brot.	Sie isst oft Müsli.
Ich nehme ein Buch.	Er nimmt das Geschenk .
Ich lese die Zeitung.	Er liest ein Buch.
Ich gebe dir den Pass.	Er gibt mir das Geschenk.
Ich fahre mit dem Auto.	Du fährst mit dem Zug.
Ich wasche das Fenster.	Du wäschst deine Haare.
Ich schlafe in dem Bett.	Du schläfst auf der Couch.
Ich laufe einen 5km Lauf.	Du läufst im Sand.
Ich trage das Fahrrad.	Du trägst das Gepäck.
Ich fange einen Fisch.	Du fängst die Frisbee.
Ich wandere am Nachmittag.	Du wanderst gern allein.
Ich koche das Mittagessen.	Du kochst um sechs Uhr.
Ich studiere immer Mathematik.	Du studierst gern Medizin.
Ich trinke viel Wasser.	Du trinkst öfters Kaffee.
Ich segle manchmal allein.	Du segelst mit deinem Freund.
Ich zeige auf die Uhr.	Du zeigst auf die Tafel.
Ich singe nicht so gern.	Sie singt ein schönes Lied
Ich kaufe das Gemüse	Sie kauft ein Brot.
Ich lache über meine Schwester.	Sie lacht über den Lehrer.
Ich begrüße den Professor.	Sie begrüßt die Familie.
Ich schreibe eine Email.	Sie schreibt an die Tafel.
Ich gehe in die Stadt.	Sie geht zur Universität.

Table 12: Sentence Stimuli by Group (c)

Group IV:	baseline	critical
	Ich sehe die Natur.	Sie sieht das Meer.
	Ich spreche am Telefon.	Sie spricht mit dem Lehrer.
	Ich esse oft Müsli	Er isst gern Brot.
	Ich nehme das Geschenk.	Sie nimmt ein Buch.
	Ich lese ein Buch.	Sie liest die Zeitung.
	Ich gebe dir das Geschenk.	Sie gibt mir den Pass.
	Ich fahre mit dem Zug.	Du fährst mit dem Auto.
	Ich wasche meine Haare.	Du wäschst das Fenster .
	Ich schlafe auf der Couch.	Du schläfst in dem Bett .
	Ich laufe im Sand.	Du läufst einen 5km Lauf.
	Ich trage mein Gepäck.	Du trägst das Fahrrad.
	Ich fange die Frisbee.	Du fängst einen Fisch.
	Ich wandere gern allein.	Du wanderst am Nachmittag.
	Ich koche um sechs Uhr	Du kochst das Mittagessen.
	Ich studiere gern Medizin.	Du studierst immer Mathematik.
	Ich trinke öfters Kaffee.	Du trinkst viel Wasser.
	Ich segle mit meinem Freund.	Du segelst manchmal allein.
	Ich zeige auf die Tafel.	Du zeigst auf die Uhr.
	Ich singe ein schönes Lied.	Sie singt nicht so gern.
	Ich kaufe ein Brot.	Er kauft das Gemüse.
	Ich lache über den Lehrer.	Er lacht über seine Schwester.
	Ich begrüße meine Familie.	Er begrüßt den Professor.
	Ich schreibe an die Tafel.	Er schreibt eine Email.
	Ich gehe zur Universität.	Er geht in die Stadt.

APPENDIX D.

Table 13: Frequency of the Verbs Used in this Study

Verb	Frequency*
essen (<i>to eat</i>)	6559
geben (<i>to give</i>)	134276
lesen (<i>to read</i>)	14947
nehmen (<i>to take</i>)	66822
sehen (<i>to see</i>)	115440
fahren (<i>to drive</i>)	35278
fangen (<i>to catch</i>)	6842
schlafen (<i>to sleep</i>)	6589
tragen (<i>to carry</i>)	29714
waschen (<i>to wash</i>)	2200
kochen (<i>to cook</i>)	2478
segeln (<i>to sail</i>)	2988
studieren (<i>to study</i>)	4263
trinken (<i>to drink</i>)	6178
wandern (<i>to hike</i>)	2422
zeigen (<i>to point at sth.</i>)	42623
begrüßen (<i>to greet</i>)	3398
gehen (<i>to go/walk</i>)	95361
kaufen (<i>to buy</i>)	6724
lachen (<i>to laugh</i>)	9586
schreiben (<i>to write</i>)	29985
singen (<i>to sing</i>)	7213

*(number of hits out of 100 million tokens)

**calculated via the DWDS (Digitales Wörterbuch der deutschen Sprache = *Digital Dictionary of the German Language*)
(<http://www.dwds.de/>)

REFERENCES

REFERENCES

- Adams, R. (2003). L2 output, reformulation and noticing: Implications for IL development. *Language Teaching Research*, 7, 347-376.
- Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 259-302). Honolulu, HI: University of Hawai'i.
- Altarriba, J., Kroll, J., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory and Cognition*, 24, 477-492.
- Anderson, J. (2005). *Cognitive psychology and its implications* (6th ed.). New York, NY: Worth Publishers.
- Baars, B. (1997). In the theatre of consciousness: Global Workspace theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies*, 4, 292-309.
- Baars, B. (2002). The conscious access hypothesis: Origins and recent evidence. *TRENDS in Cognitive Sciences*, 6, 47-52.
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Reviews Neuroscience*, 4, 829-839.
- Bardovi-Harlig, K. (1987). Markedness and salience in second language acquisition. *Language Learning*, 37, 385-407.
- Bell, F. (2006). *Comprehension aids, internet technologies, and the reading of authentic materials by adult second language learners*. (Doctoral dissertation, Florida State University) Retrieved from <http://search.proquest.com/>
- Bell, P., & Collins, L. (2009). 'It's vocabulary'/'it's gender': Learner awareness and incidental learning. *Language Awareness*, 18, 277-293.
- Bird, K. (2004). *Analysis of Variance via Confidence Intervals*. London: Sage.
- Block, N. (1995). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18, 227-247.
- Block, N. (2005). Two neural correlates of consciousness. *TRENDS in Cognitive Sciences*, 9, 46-51.
- Bordag, D., & Pechmann, T. (2007). Factors influencing L2 gender processing. *Bilingualism: Language and Cognition*, 10, 299-314.

- Bowles, M. (2010). *The think-aloud controversy in second language research*. New York, NY: Routledge.
- Bowles, M. (2008). Task type and reactivity of verbal reports in SLA: A first look at an L2 task other than reading. *Studies in Second Language Acquisition*, 30, 359-387.
- Bowles, M., & Leow, R. (2005). Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation. *Studies in Second Language Acquisition*, 27, 415-440.
- Broderick, M. (2005). *Phonological awareness and second language decoding in FLES students learning German* (Doctoral Dissertation, University of Pittsburgh). Retrieved from <http://search.proquest.com/>
- Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye-movement control in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 53-78). Oxford: Oxford University Press.
- Bybee, J., & Newman, J. (1995). Are stem changes as natural as affixes? *Linguistics*, 33, 633-654.
- Calvo, M., & Meseguer, E. (2002). Eye-movements and processing stages in reading: Relative contributions of visual, lexical, and contextual factors. *The Spanish Journal of Psychology*, 5, 66-77.
- Carr, T. & Curran, T. (1994). Cognitive factors in learning about structured sequences: Applications to syntax. *Studies in Second Language Acquisition*, 16, 205-230.
- Carrell, P. (1992). Awareness of text structure: Effects on recall. *Language Learning*, 42, 1-20.
- Carroll, S. (2006a). Autonomous induction theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An Introduction* (pp. 155-174). Mahwah, NJ: Lawrence Erlbaum Associates.
- Carroll, S. (2006b). The 'micro-structure' of a learning problem: Prosodic prominence, attention, segmentation and word learning in a second language. In C. Gurski & M. Radisic (Eds.), *Proceedings of the 2006 annual conference of the Canadian Linguistic Association: Cascadilla Proceedings Project*.
- Carroll, S. (2006c). Salience, awareness and SLA. In M. G. O'Brien, C. Shea & J. Archibald (Eds.), *Proceedings of the 8th generative approaches to second language acquisition conference* (pp. 17-24). Somerville, MA: Cascadilla Proceedings Project.
- Carroll, S. (2001). *Input and evidence: The raw material of second language acquisition*. Amsterdam: John Benjamins.

- Clahsen, H. (1999). Lexical entities and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences*, 22, 199-1060.
- Clahsen, H. (2007). Psycholinguistic perspectives on grammatical representations. In S. Featherston & W. Sternefeld (Eds.), *Roots: Linguistics in Search of its Evidential Base* (pp. 97-132). Mouton de Gruyter: Berlin.
- Clahsen, H., Felser, C. Neubauer, K., Mikako, S. & Silva, R. (2010). Morphological structure in native and non-native language processing. *Language Learning*, 60, 21-43.
- Clahsen, H. & Neubauer, K. (2010). Morphology, frequency, and the processing of derived words in native and non-native speakers. *Lingua*, 120, 2627-2637.
- Cleermans, A., & Destrebecqz, A. (2005). Real rules are conscious. *Behavioral and Brain Sciences*, 28, 19-20.
- Clifton, C., Staub, A., & Rayner, K. (2007). Eye-movements in reading words and sentences. In R. Van Gompel (Ed.), *Eye-movements: A window on mind and brain* (pp. 241-372). Oxford: Elsevier Science.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education*. New York, NY: Routledge.
- Cook, R. (1977). Detection of influential observations in linear regression. *Technometrics American Statistical Association*, 19, 15-18.
- Corbetta, M., & Schulman, G. (2002). Control of goal directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3, 201-215.
- Cowan, N. (1995). *Attention and memory: An integrated framework*. Oxford: Oxford University Press.
- Cowan, N. (2005). *Working memory capacity*. Hove, UK: Psychology Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 97-185.
- Dehaene, S., & Changeux, J. (2004). Neural mechanisms for access to consciousness. In M. Gazzaniga (Ed.), *The cognitive neurosciences III* (pp. 1145-1158). Cambridge, MA: MIT Press.
- Dehaene, S., Changeux, J., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *TRENDS in Cognitive Sciences*, 10, 204-211.

- Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J., Poline, J., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, 4, 752-758.
- DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, 19, 195-221.
- DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). New York, NY: Blackwell Publishing.
- DeKeyser, R. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, 55, 1-25.
- DeKeyser, R. (2009). Cognitive-psychological processes in second language learning. In M. Long & C. Doughty (Eds.), *Handbook of Second Language Teaching*. Oxford: Blackwell.
- DeKeyser, R., Salaberry, R., Robinson, P., & Harrington, M. (2002). What Gets Processed in Processing Instruction? A Commentary on Bill VanPatten's "Processing Instruction: An Update". *Language Learning*, 52, 805-823.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NY: Lawrence Erlbaum Associates.
- Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.
- Doughty, C., & Williams, J. (1998). *Focus on form in classroom and second language acquisition*. Cambridge: Cambridge University Press.
- Doughty, C., & Williams, J. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 206-257). Cambridge: Cambridge University Press.
- Drieghe, D., Brysbaert, M., Desmet, T., & De Baecke, C. (2004). Word skipping in reading: On the interplay of linguistic and visual factors. *European Journal of Cognitive Psychology*, 16, 79-103.
- Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. (2007). Visual word recognition by bilinguals in a sentence context: Evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 663-679.
- Egi, T. (2004). Verbal reports, noticing and SLA research. *Language Awareness*, 13, 243-264.
- Ellis, N. (1994a). Implicit and explicit processes in language acquisition: An introduction. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1-32). London: Academic Press.

- Ellis, N. (1994b). Vocabulary acquisition: The implicit ins and outs of explicit cognitive mediation. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 211-282). London: Academic Press.
- Ellis, N. (2002). Frequency effects in language processing. A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143-188.
- Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 305-352.
- Ellis, N. (2006a). Language acquisition as rational contingency learning. *Applied Linguistics*, 27, 1-24.
- Ellis, N. (2006b). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, 27, 164-194.
- Ellis, N. (2007). Implicit and explicit knowledge about language. In J. Cenoz & N. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed., Vol. 6, pp. 119-132). New York, NY: Springer.
- Ellis, N. (2011). Implicit and explicit SLA and their interface. In C. Sanz & R. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA and bilingualism* Washington DC: Georgetown University Press.
- Ellis, N., & Sagarra, N. (2010). Learned attention effects in L2 temporal reference: The first hour and the next eight semesters. *Language Learning*, 60, 85-108.
- Ellis, N., & Schmidt, R. (1998). Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language and Cognitive Processes*, 13, 307-336.
- Ellis, N. & Sinclair, S. (1996). Working Memory in the Acquisition of Vocabulary and Syntax: Putting Language in Good Order. *The Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*, 49, 234-250.
- Ellis, R. (1994). Factors in the incidental acquisition of second language vocabulary from oral input: A review essay. *Applied Language Learning*, 5, 1-32.
- Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge? A review of research. *Studies in Second Language Acquisition*, 24, 223-236.
- Ellis, R., Basturkmen, H., & Loewen, S. (2002). Doing focus on form. *System*, 30, 419-435.
- Elman, J. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48, 71-99.

- Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, 112, 777-813.
- Eysenck, M., & Keane, M. (2005). *Cognitive psychology: A student's handbook* (5th ed.). New York, NY: Psychology Press.
- Eysenck, M. & Keane, M. (2010). *Cognitive psychology: A student's handbook*: New Yorkk: Psychology Press.
- Field, A. (2009). *Discovering statistics using SPSS - and sex and drugs and rock 'n' roll* (3rd ed.). London: Sage.
- Fitzmaurice, G., Laird, N. & Ware, J. (2004). *Applied longitudinal analysis*. Hoboken, NJ: Wiley.
- Fotos, S. (1993). Consciousness raising and noticing through focus on form: Grammar task performance versus formal instruction. *Applied Linguistics*, 14, 385-407.
- Fotos, S. (1994). Integrating grammar instruction and communicative language use through grammar consciousness-raising tasks. *TESOL Quarterly*, 28, 223-351.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye-movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, 14, 178-210.
- French-Mestre, C. (2005). Eye-movement recording as a tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research*, 21, 175-198.
- Frisson, S., Rayner, K., & Pickering, M. (2005). Effects of contextual predictability and transitional probability on eye-movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 862-877.
- Gass, S. (1997). *Input, interaction and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gass, S., & Alvarez-Torres, M. L. (2005). Attention when?: An investigation of the ordering effect of input and interaction. *Studies in Second Language Acquisition*, 27, 1-31.
- Gass, S., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Gass, S., Mackey, A., & Pica, T. (1998). The role of input and interaction in second language acquisition – Introduction to the special issue. *The Modern Language Journal*, 28, 299-305.

- Gass, S., Svetics, I., & Lemelin, S. (2003). Differential affects of attention. *Language Learning*, 53, 497-545.
- Godfroid, A. (2010). *Cognitive processes in second language acquisition: The role of noticing, attention and awareness in processing words in written L2 input*. University of Brussels.
- Godfroid, A., Housen, A., & Boers, F. (2010). A procedure for testing the Noticing Hypothesis in the context of vocabulary acquisition. In M. Pütz & L. Scola (Eds.), *Inside the learner's mind: Cognitive processing and second language acquisition* (pp. 169-197). Amsterdam: John Benjamins.
- Goldschneider, J., & DeKeyser, R. (2001). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning*, 51, 1-50.
- Goo, J. (2010). Working memory and reactivity. *Language Learning*, 60, 712-752.
- Grigorenko, E., Sternberg, R., & Ehrman, M. (2000). A theory based approach to the measurement of foreign language learning ability: The Canal-F theory and test. *Modern Language Journal*, 84, 390-405.
- Hama, M., & Leow, R. (2010). Learning without awareness revisited: Extending Williams (2005). *Studies in Second Language Acquisition*, 32(3), 465-491.
- Hampton, J. (2005). Rules and similarity - a false dichotomy. *Behavioral and Brain Sciences*, 28, 26.
- Hanaoka, O. (2007). Output, noticing, and learning: An investigation into the role of spontaneous attention to form in a four-stage writing task. *Language Teaching Research*, 11, 459-479.
- Han, Z., Park, E. S., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics*, 29, 597-618.
- Hardin, J. & Hilbe, J. (2003). *Generalized estimating equations*. London: Chapman & Hall.
- Hasegawa, R., Carpenter, P., & Just, M. (2002). An fMRI study of bilingual sentence comprehension and workload. *NeuroImage*, 15, 647-660.
- Haznedar, B. (2001). The acquisition of the IP system in child L2 English. *Studies in Second Language Acquisition*, 1-39.
- Hoover, M., & Dwivedi, V. (1998). Syntactic processing by skilled bilinguals. *Language Learning*, 48, 1-29.
- Hosmer, D., & Lemeshow, S. (2000). *Applied logistic regression*. New York: Wiley.
- Howell, D. (2009). *Statistical methods for psychology* (3rd ed.). Wadsworth: Cengage Learning.

- Hulstijn, J. (in press). Incidental learning in second language acquisition. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics*: Blackwell.
- Hulstijn, J. (1989). Implicit and incidental second language learning: Experiments in the processing of natural and partly artificial input. In H. Dechert & M. Raupach (Eds.), *Interlingual processing* (pp. 49-73). Tübingen: Gunter Narr.
- Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, 18, 193-223.
- Hulstijn, J. (2003). Incidental and intentional learning. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 349-381). New York, NY: Blackwell Publishing.
- Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition*, 27, 129-140.
- O'Brien, I., Freed B., & Collentine, J. (2007). Phonological memory predicts second language oral fluency gains in adults. *Studies in Second Language Acquisition* 29, 557 -581.
- Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL Revitalization. *Studies in Second Language Acquisition*, 24, 541-577.
- Izumi, S., & Bigelow, M. (2000). Does output promote noticing in second language acquisition? *TESOL Quarterly*, 34, 239-278.
- Izumi, S., Bigelow, M., Fujiwara, F., & Fearnow, S. (1999). Testing the output hypothesis. *Studies in Second Language Acquisition*, 21, 421-452.
- Izumi, Y., & Izumi, S. (2004). Investigating the effects of oral output on the learning of relative clauses in English: Issues in the psycholinguistic requirements for effective output tasks. *Canadian Modern Language Review*, 60, 587-609.
- Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, 57, 1-33.
- Jourdenais, R., Ota, M., Stauffer, S., Boyson, B., & Doughty, C. (1995). Does textual enhancement promote noticing? A think-aloud analysis. In R. W. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 183-216). Honolulu, HI: University of Hawai'i.
- Just, M., & Carpenter, P. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87, 329-354.
- Just, M., & Carpenter, P. (1992). A theory of working memory: Individual differences in comprehension. *Psychological Review*, 99, 122-149.

- Kennedy, A., & Pynte, J. (2008). The consequences of violations to reading order: An eye-movement analysis. *Vision Research*, 48, 2309-2320.
- Kennedy, A., Pynte, J., & Ducrot, S. (2002). Parafoveal-on-foveal interactions in word recognition. *Quarterly Journal of Experimental Psychology*, 55, 1307-1338.
- Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI perspective. *Discourse Processes* 39, 125-128.
- Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye-movements in reading. *European Journal of Cognitive Psychology*, 16, 262-284.
- Kliegl, R., Nuthmann, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, 135, 12-35.
- Koch, C. (2004). *The Quest for Consciousness: A neurobiological Approach*. CO: Roberts and Publishers.
- Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain processes. *TRENDS in Cognitive Sciences*, 11, 16-22.
- Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon Press.
- Krashen, S. (1985). *The Input Hypothesis: Issues and implications*. London: Longman.
- Krashen, S. & Terrell, T. (1983). *The Natural Approach: Language acquisition in the classroom*. Hayward, CA: Alemany Press.
- Kurby, C., Britt, M., & Magliano, J. (2005). The role of top-down and bottom-up processes in between-text integration. *Reading Psychology*, 26, 335-362.
- Lamme, V. (2003). Why visual attention and awareness are different. *TRENDS in Cognitive Sciences*, 7, 12-18.
- Lardiere, D. (2007). *Ultimate attainment in second language acquisition: A case Study*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Larsen-Freeman, D. (2010). Not so fast: A discussion of L2 morpheme processing and acquisition. *Language Learning*, 60, 221-230.
- Lee, J. (1998) The relationship of verb morphology to second language reading comprehension and input processing. *The Modern Language Journal*, 82, 33-48.

- Lee, S., & Huang, H. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition*, 30, 307-331.
- Leon-Guerrero, A., & Frankfort-Nachmias, C. (2010). *Social statistics for a diverse society* (6 ed.). California: Pine Forge Press.
- Leow, R., Johnson, E. & Zárte-Sández, G. (2011). Getting a grip on the slippery construct of awareness: Toward a finer-grained methodological perspective" In C. Sanz & R. Leow. *Implicit and explicit conditions, processes and knowledge in SLA and bilingualism*. Washington, DC: Georgetown University Press, 2010.
- Leow, R., Hsieh, H., & Moreno, N. (2008). Attention to form and meaning revisited. *Language Learning*, 58, 665-695.
- Leow, R., & Morgan-Short, K. (2004). To think-aloud or not to think-aloud: The issue of reactivity in SLA research methodology. *Studies in Second Language Acquisition*, 26, 35-57.
- Leow, R. (1997a). Attention, awareness and foreign language behavior. *Language Learning*, 47, 467-505.
- Leow, R. (1998). Toward operationalizing the process of attention in SLA: Evidence for Tomlin and Villa's (1994) fine-grained analysis of attention. *Applied Psycholinguistics*, 19, 133-159.
- Leow, R. (2000). A study of the role of awareness in foreign language behavior. Aware versus unaware learners. *Studies in Second Language Acquisition*, 22, 557-584.
- Leow, R. (2002). Models, attention, and awareness in SLA. A response to Simard and Wong's "alertness, orientation, and detection: The conceptualization of attentional focus in SLA". *Studies in Second Language Acquisition*, 24, 113-119.
- Liversedge, S. & Findlay, J. (2000). Saccadic eye-movements and cognition. *TRENDS in Cognitive Sciences*, 4, 6-14.
- Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, 27, 361-386.
- Long, M. (1996). The role of linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of Second Language Acquisition*. San Diego: Academic Press.
- Lotto, L., & De Groot, A. (1998). Effects of learning method and word type on acquiring vocabulary in a new language. *Language Learning*, 48, 31-69.
- Mackey, A. (2006). Feedback, noticing and instructed second language learning. *Applied Linguistics*, 27, 405-430.

- Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback. *Studies in Second Language Acquisition*, 22, 471-497.
- Mackey, A., & Philp, J. (1998). Conversational interaction and second language development: Recasts, responses, and red herrings? *The Modern Language Journal*, 82, 338–356.
- Maindonald, J., Braun, J. (2010). *Data Analysis and Graphics Using R: An example -based approach* (3 ed.). Cambridge: Cambridge University Press.
- Mathews, R., Buss, R., Stanley, W., Blanchard-Fields, F., Cho, J., & Druhan, B. (1989). Role of implicit and explicit processes in learning from examples: A synergistic effect. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 1083-1100.
- McConkie, G., Kerr, P., & Dyre, B. (1994). What are “normal” eye-movements during reading: Toward a mathematical description. In J. Ygge & G. Lennerstrand (Eds.), *Eye-movements in reading* (pp. 331-343). Amsterdam: Pergamon Press.
- McConkie, G., Kerr, P., Reddix, M., Zola, D., & Jacobs, A. (1989). Eye-movement control in reading: Frequency of refixating a word. *Perception & Psychophysics*, 46, 245–253.
- McLaughlin, B. (1965). ‘Intentional’ and ‘incidental’ learning in human subjects: The role of instructions to learn and motivation. *Psychological Bulletin*, 63, 359-376.
- McLaughlin, B. (1990). ‘Conscious’ versus ‘unconscious’ learning. *TESOL Quarterly*, 24, 617-634.
- McLaughlin, B., & Heredia, R. (1996). Information-processing approaches to research on second language acquisition and use. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 213-228). San Diego, CA: Academic Press.
- Merikle, P., Smilek, D., & Eastwood, J. (2001). Perception without awareness: Perspectives from Cognitive Psychology. *Cognition*, 79, 115-134.
- Meyers, L., Gamst, A., & Guarino, J. (2006). *Applied multivariate research: Design and interpretation*. Austin, TX: Sage.
- Miyake, A., & Friedman, D. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. Healy & L. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Morrison, R. (1984). Manipulation of stimulus onset delay in reading: Evidence for parallel programming of saccades. *Journal of Experimental Psychology: Human Perception and Performance*, 10, 667-682.
- Naccache, L., Blandin, E., & Dehaene, S. (2002). Unconscious masked priming depends on temporal attention. *Psychological Science*, 13, 416-424.

- Neubauer, K., & Clahsen, H. (2009). Decomposition of inflected words in a second language: An experimental study of German participles. *Studies in Second Language Acquisition*, 31, 403-435.
- Norris, J. & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417-528.
- Ortega, L. (2009). *Understanding second language acquisition*. London: Hodder Education.
- Overgaard, M., Fehl, K., Mouridsen, K., Bergholt, B., & Cleermans, A. (2008). Seeing without seeing? Degraded conscious vision in a blindsight patient. *Public Library of Science One*, 3, e3028.
- Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 393-419). London: Academic Press.
- Parodi, T., Schwartz, B., & Clahsen, H. (2004). On the L2 acquisition of the morphosyntax of German nominals. *Linguistics*, 42, 669-705.
- Perruchet, P., & Pacteau, C. (1990). Synthetic grammar learning: Implicit rule abstraction or explicit fragmentary knowledge? *Journal of Experimental Psychology: General*, 119, 264-275.
- Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: one phenomenon, two approaches. *Trends in Cognitive Sciences*, 10, 233-223.
- Philp, J. (2003). Constraints on noticing the gap. *Studies in Second Language Acquisition*, 25, 99-126.
- Pienemann, M. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition*, 6, 186-214.
- Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics*, 10, 52-79.
- Pollatsek, A., & Rayner, K. (2009). Reading. In L. Squire (Ed.), *Encyclopedia of Neuroscience* (Vol. 8, pp. 29-34). Oxford: Academic Press.
- Prévost, P., & White, L. (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, 16, 103-133.
- Qi, D., & Lapkin, S. (2001). Exploring the role of noticing at a three-stage second language writing task. *Journal of Second Language Writing*, 10, 277-303.

- Rayner, K. (1998). Eye-movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.
- Rayner, K. (2007). *Attention and eye-movements in reading, scene perception, and visual search*. Paper presented at the Barlett Lecture to the Experimental Society, Edinburgh.
- Rayner, K. (2009). Eye-movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62, 1457-1506.
- Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E. (2004). The effects of frequency and predictability on eye fixations in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 720-732.
- Rayner, K., Liversedge, S., White, S., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye-movements. *Psychological Science*, 14, 385-389.
- Rayner, K., Liversedge, S., & White, S. (2006). Eye-movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research*, 46, 310-323.
- Rayner, K., Liversedge, S., White, S., & Vergilino-Perez, D. (2003). Reading disappearing text: Cognitive control of eye-movements. *Psychological Science*, 14, 385-389.
- Rayner, K., Pollatsek, A., & Reichle, E. (2003). Eye-movements in reading: Models and data. *Behavioral and Brain Sciences*, 26, 507-526.
- Rayner, K., Raney, G., & Pollatsek, A. (1995). Eye-movements and discourse processing. In R. Lorch & E. O'Brien (Eds.), *Sources of coherence in reading* (pp. 9-36). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rayner, K., Sereno, S., & Raney, G. (1996). Eye-movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1188-1200.
- Rayner, K., & Well, A. (1996). Effects of contextual constraint on eye-movements in reading: A further examination. *Psychonomic Bulletin & Review*, 3, 504-509.
- Reber, A., Allen, R., & Reber, P. (1999). Implicit versus explicit learning. In R. J. Sternberg (Ed.), *The nature of cognition* (pp. 475-513). Cambridge, MA: MIT Press.
- Reichle, E., Pollatsek, A., & Rayner, K. (2006). E-Z Reader: A cognitive-control, serial-attention model of eye-movement behavior during reading. *Cognitive Systems Research*, 7, 4-22.
- Reichle, E., Rayner, K., & Pollatsek, A. (1999). Eye-movement control in reading: Accounting for initial fixation locations and refixations within the E-Z Reader model. *Vision Research*, 39, 4403 - 4411.

- Reichle, E. Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and brain sciences*, 26, 445-526.
- Roberts, M. (1995). Awareness and the efficacy of error correction. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 162-182). Honolulu, HI: University of Hawaii Press.
- Roberts, L. Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 Influence and General Learner Effects. *Studies in Second Language Acquisition*, 30, 333-357.
- Robinson, P. (1995). Attention, memory, and the "Noticing" Hypothesis. *Language Learning*, 45, 283-331.
- Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition*, 18, 27-67.
- Robinson, P. (1997a). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition*, 19, 223-247.
- Robinson, P. (1997b). Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, 47, 45-99.
- Robinson, P. (2002). Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition. *Second Language Research*, 17, 368-392.
- Robinson, P. (2003). Attention and memory during SLA. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 633-678). Oxford: Blackwell Publishing.
- Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition*, 27, 235-268.
- Robinson, P. (2008). Attention and awareness. In J. Cenoz & N. Hornberger (Eds.), *Encyclopedia of Language and Education: Knowledge about Language* (Vol. 6, pp. 133-142). New York: Springer Academic.
- Rosa, E., & Leow, R. (2004). Awareness, different learning conditions, and L2 development. *Applied Psycholinguistics*, 25, 269-292.
- Rosa, E., & O'Neill, M. (1999). Explicitness, intake and the issue of awareness. *Studies in Second Language Acquisition*, 21, 511-556.

- Rossomondo, A. (2007). The role of lexical temporal indicators and text interaction format in the incidental acquisition of the Spanish future tense. *Studies in Second Language Acquisition*, 29, 39-66.
- Rutherford, A. (2001). *Introductin ANOVA and ANCOVA: A GLM Approach*. London: Sage.
- Ruz, M., & Nobre, A. (2008). Attention modulates initial stages of visual word processing. *Journal of Cognitive Neuroscience*, 20, 1727-1736.
- Sachs, R., & Polio, C. (2007). Learners' uses of two types of written feedback on an L2 writing revision task. *Studies in Second Language Acquisition*, 29(67-100).
- Sachs, R., & Suh, B. R. (2007). Textually enhanced recasts, learner awareness, and L2 outcomes in synchronous computer-mediated interaction. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 197-227). Oxford: Oxford University Press.
- Sagarra, N. (2007). Working memory and L2 processing of redundant grammatical forms. In Z. Han (Ed.), *Second language processing and instruction: Broadening the scope of inquiry* (pp. 142-159). Clevedon, UK: Multilingual Matters.
- Sall, J., Creighton, L., & Lehman, A. (2005). *JMP start statistics: A guide to statistics and data analysis using JMP and JMP IN software* (3 ed.). Cary, NC: SAS Institute Inc.
- Sanz, C., Lin, H., Lado, B., Bowden, H., & Stafford, C. (2009). Concurrent verbalizations, pedagogical conditions, and reactivity: two CALL studies. *Language Learning*, 59, 33-71.
- Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instructional design. In P. Robinson (Ed.), *Cognition and second language acquisition*. New York: Cambridge University Press.
- Schachter, J. (1998). Recent Research in Language Learning Studies: Promises and Problems. *Language Learning*, 48, 557-583.
- Schierloh, M. (in press). Re-investigating the noticing function of output. *Language Learning*, 62.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-158.
- Schmidt, R. (1993a). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206-226.
- Schmidt, R. (1993b). Consciousness, learning and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage Pragmatics* (pp. 21-42). New York: Oxford University Press.

- Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11, 11-26.
- Schmidt, R. (1994b). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165-209). London: Academic Press.
- Schmidt, R. (1995). Consciousness and foreign language learning: a tutorial on the role of attention and awareness in learning R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1-63). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.
- Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istatnto, M. Nagami, J. W. Sew, T. Suthiwan & I. Walker (Eds.), *Proceedings of CLaSiC*. National University of Singapore: Center for Language Studies.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-322). Rowley, MA: Newbury House.
- Segalowitz, N. (2003). Automaticity and second languages. In M. Long & C. Doughty (Eds.), *Handbook of Second Language Acquisition*. Malden, MA: Blackwell.
- Shanks, D., & St. John, M. (1994). Characteristics of dissociable human learning systems. *Behavioral and Brain Sciences* 17, 367-395.
- Shanks, D. (2005). Implicit learning. In K. Lamberts & R. Goldstone (Eds.), *Handbook of Cognition* (pp. 202-220). London: Sage.
- Sharwood-Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research*, 7, 118-132.
- Sharwood-Smith, M., & Truscott, J. (2010). Consciousness and language: A processing perspective. In E. Perry (Ed.), *New horizons in the neuroscience of consiousness*. Amsterdam: John Benjamins.
- Sharwood-Smith, M., & Truscott, J. (in press). Input, intake, and consciousness: The quest for a theoretical foundation. *Studies in Second Language Acquisition*, 33.
- Sherman, K. (2007). *A meta-analysis of interventions for phonemic awareness and phonics instruction for delayed older readers*. Dissertation, University of Oregon.

- Simard, D., & Wong, W. (2001). Alertness, orientation, and detection. The conceptualization of attentional functions in SLA. *Studies in Second Language Acquisition*, 23, 103-124.
- Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research* 27, 251-272.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P., & Foster, P. (2001). Cognition and tasks P. Robinson (ed.) *In cognition and second language instruction* (pp. 183-205). Cambridge: Cambridge University Press.
- Smith, B. (2010). Employing eye-tracking technology in researching the effectiveness of recasts in CMC. In: Hult, F. M. (ed.), *Directions and Prospects for Educational Linguistics* (pp. 79-97). Springer.
- Smith, E. (2005). Rule and similarity as prototype concepts. *Behavioral and Brain Sciences*, 28, 34-35.
- Smith, L., Gratz, Z., & Bousquet, S. (2008). *The art and practice of statistics*. California: Wadsworth.
- Song, M., & Suh, B. (2008). The effects of output task types on noticing and learning of the English past counterfactual conditional. *System*, 36, 295-312.
- Spada, N. & Lightbrown, P. (1999). Instruction, first language influence, and developmental readiness in second language acquisition. *Modern Language Journal*, 38, 1-22.
- Starr, M., & Rayner, K. (2001). Eye-movements during reading: Some current controversies. *TRENDS in Cognitive Sciences*, 5, 156-163.
- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125-144). Oxford: Oxford University Press.
- Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom and second language acquisition* (pp. 64-81). New York: Cambridge University Press.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471-484). Mahwah, NJ: Lawrence Erlbaum Associates.
- Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.

- Swain, M., & Lapkin, S. (2002). Talking it through: Two French immersion learner's responses to reformulation. *International Journal of Educational Research*, 37, 285-304.
- Taylor, J. (2003). *Linguistic categorization* (3 ed.). Oxford: Oxford University Press.
- Todeva, E. (2010). Multilingualism as a kaleidoscopic experience: The mini universes within. In E. Todeva & J. Cenoz (Eds.), *The multiple realities of multilingualism*. Berlin: Mouton de Gruyter.
- Tokowicz, N., & Warren, T. (2010). Beginning adult L2 learners' sensitivity to morphosyntactic violations: A self-paced reading study. *The European Journal of Cognitive Psychology*, 22, 1092-1106.
- Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16, 183-204.
- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second Language Research*, 24, 103-135.
- Truscott, J. (2000). Unconscious second language acquisition: Alive and well. *Studies in English Literature and Linguistics*, 25, 115-131.
- Tsui, A., & Fullilove, J. (1998). Bottom-up or top-down processing as a discriminator of L2 reading performance. *Applied Linguistics*, 4, 432-451.
- van den Noort, M., Bosch, P., & Hugdahl, K. (2006). Foreign Language Proficiency and Working Memory Capacity. *European Psychologist*, 11, 289-296.
- Van Hell, J. G., & De Groot, A. M. (1999). Explicitness, intake, and the issue of awareness. Another piece to the puzzle. *Studies in Second Language Acquisition*, 21, 511-556.
- van Heuven, W., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, 39, 458-483.
- VanPatten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition*(12), 287-301.
- VanPatten, B. (1994). Evaluating the role of consciousness in second language acquisition: Terms, linguistic features and research methodology. *AILA Review*, 11, 27-36.
- VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Norwood, NJ: Ablex.
- VanPatten, B. (2003). Input processing in SLA. In B. VanPatten (Ed.), *Processing Instruction: Theory, Research and Commentary* (pp. 5-32). Mahwah, NJ: Lawrence Erlbaum Associates.

- VanPatten, B. (2004). *Form-meaning connections in second language acquisition*. New York, NY: Routledge.
- VanPatten, B., & Cadierno, T. (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition*, 15, 225–243.
- VanPatten, B., Keating, G., & Leiser, M. (in press). Missing verbal inflections as a representational problem: evidence from online methodology. *Linguistic Approaches to Bilingualism*.
- VanPatten, B., & Oikarinen, S. (1996). Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition*, 18, 495–510.
- Vergilino, D., & Beauvillain, C. (2000). The planning of refixation saccades in reading. *Vision Research*, 40, 3527–3538.
- Vergilino, D., & Beauvillain, C. (2001). Reference frames in reading: Evidence from visually guided and memory-guided saccades. *Vision Research*, 41, 3547 – 3557.
- Vergilino-Perez, D., Collins, T., & Dore-Mazars, K. (2004). Decision and metrics of refixations in reading isolated words. . *Vision Research*, 44, 2009–2017.
- Weiskrantz, L. (1997). *Consciousness lost and found: A neuropsychological exploration*. New York, NY: Oxford University Press.
- White, J. (1998). Getting the learner's attention. In C. Doughty and J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 85-113). Cambridge: Cambridge University Press.
- White, S. (2008). Eye-movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 205-223.
- Wickens, C. (1992). *Engineering psychology and human performance* (2nd ed.). New York: HarperCollins.
- Wickens, C. (2007). Attention to the second language. *IRAL - International Review of Applied Linguistics in Language Teaching*, 177–191.
- Wickens, C., Gordon, S. E., & Liu, Y. (1997). *An introduction to human factors engineering*. New York, NY: Longman.
- Williams, J. (2001). Learner generated attention to form. *Language Learning*, 51, 303-346.
- Williams, J. (2005). Form-focused instruction. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 671-692). Mahwah, NJ: Lawrence Erlbaum Associates.

- Williams, J. N. (2004). Implicit learning of form-meaning connections. In J. Williams, B. VanPatten, S. Rott & M. Overstreet (Eds.), *Form-meaning connections in second language acquisition*. Mahwah, NJ: Erlbaum.
- Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition*, 27, 269-304.
- Williams, J. N. (2009). Implicit Learning. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition*. Bingley, UK: Emerald Group Publishing Limited.
- Williams, J. N. & Lovatt, P. (2003). Phonological memory and rule learning. *Language Learning*, 53, 67-121.
- Williams, R., & Morris, R. (2004). Eye-movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology*, 16, 312-339.
- Woodman, G. & Luck, S. (2003). Dissociations among attention, perception, and awareness during object-substitution masking. *Psychological Science*, 14, 605-611.
- Wright, R., & Ward, L. (2008). *Orienting of attention*. Oxford: Oxford University Press.
- Yantis, S. (2008). The neural basis of selective attention: Cortical sources and targets of attentional modulation. *Neurological Research*, 30, 356-360.
- Yoshida, M. (2008). *Think-aloud protocols and type of reading task: The issue of reactivity in L2 reading research*. Paper presented at the 2007 Second Language Research Forum, Somerville, MA.