THE EFFECT OF PRIOR L1 KNOWLEDGE ON THE IMPLICIT AND EXPLICIT LEARNING OF L2 SYNTAX FROM READING A NOVEL

By

Jieun Ahn

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Second Language Studies – Doctor of Philosophy

2019

ABSTRACT

THE EFFECT OF PRIOR L1 KNOWLEDGE ON THE IMPLICIT AND EXPLICIT LEARNING OF L2 SYNTAX FROM READING A NOVEL

By

Jieun Ahn

A growing number of researchers have considered the acquisition of L2 grammar under incidental, or meaning-focused, learning conditions in order to understand the scope and possibilities of naturalistic adult L2 learning. However, the effect of the first language (L1) in the incidental learning of L2 syntax has not been directly studied. To address this gap, in the present dissertation, I investigated the effect of prior L1 knowledge on the implicit and explicit learning of L2 syntax under incidental learning conditions.

Forty L1 English (head-initial, right-branching) and forty L1 Korean (head-final, left-branching) speakers read the novel *The Mysterious Affair at Styles* (Christie, 1920) rewritten in a semi-artificial language, *Koreanish*, which consisted of English vocabulary and head-final Korean syntax. The participants' eye movements were recorded during reading by using an EyeLink Portable Duo eye tracker (SR Research, Canada). From their eye movements, I derived two processing measures: changes in sentence reading times over time and the participants' real-time responses to word order violations. After reading, the participants were immediately tested with a surprise grammaticality judgment test (GJT) with source attributions, which was followed by the first part of the debriefing interview. Two weeks later, the participants completed a delayed GJT and the second part of the debriefing interview.

Triangulation of the online and offline measures exhibited the significant and pervasive effects of prior L1 knowledge on the incidental acquisition of L2 syntax. During the exposure task (novel reading), the Korean experimental (KE) group exhibited a faster initial decrease in

sentence reading times than did the English experimental (EE) group. Furthermore, only the KE group showed online grammatical sensitivity. Particularly, the Korean L1-aware participants, who later became aware of the cross-linguistic similarity between the target language and their L1, showed robust sensitivity effects, even before L1 awareness emerged. Subsequently, on the GJT, the KE group exhibited stronger evidence of implicit and explicit knowledge than did the EE group. The KE group's syntactic knowledge was significantly enhanced over time, which prompted them to outperform the EE group on the delayed GJT. These findings indicated that the Korean speakers certainly had advantages in acquiring the target word order, which followed their L1 Korean, and L1 awareness gave them an edge in knowledge development. In terms of online processing data, although both the KE and EE groups sped up over time while reading the exposure text, their learning curves had a different form. The EE group's reading data followed the power law of practice (Anderson, 1982), whereas those of the KE group did not follow a clearly discernable pattern.

The findings of this dissertation suggest that L2 learners with a different L1 background perform on an unequal footing because of their prior L1 experience, especially in naturalistic learning contexts. The two L1 groups, in fact, presented opposing directions of knowledge development. The Korean participants—who had relevant prior L1 knowledge—showed a progression from implicit to explicit knowledge under incidental exposure, whereas the English participants—who lacked relevant prior L1 knowledge—began with explicit knowledge and failed to reach implicit knowledge. This dissertation therefore elucidated the multifaceted aspects of L1 influence on L2 development, uncovering the complex nature of linguistic transfer, variability in L2 development, and the moderating role of awareness in meaning-focused, incidental learning conditions.

ACKNOWLEDGEMENTS

Regarding my ability to achieve this wonderful accomplishment, I am greatly indebted to my dedicated advisor Dr. Aline Godfroid for going above and beyond anything I could have expected from an advisor. I would like to express my deepest appreciation for her persistent guidance and support through each stage of the process. Her enthusiasm and love for research are contagious. She opened a door to the world of eye-tracking for me, and it was like riding a magic carpet. Under her guidance, I successfully overcame many difficulties. Her faith in me strengthened my resolve, and she has shown me what a good researcher and person should be. I have been extremely lucky to have her as my mentor.

I would also like to extend my deepest gratitude to all my committee members for providing me with extensive personal and professional guidance. Without their constructive and intensive feedback, this dissertation would not have been possible. My earnest thanks go to Drs. Sue Gass, Patti Spinner, and Shawn Loewen. Dr. Gass is an amazing role model of what it means to be a superior mentor, teacher, and leader. She is indeed a multitasking genius. Dr. Spinner nurtured my enthusiasm for syntax and always shared her expertise willingly. In her morphosyntax class, I was able to learn the beauty of syntax. Dr. Loewen believed in my abilities and provided me with continuous support and encouragement. He is very approachable and always gives care for his students. My sincere thanks also go to Drs. Zoltan Dienes and Patrick Rebuschat for their willingness to serve as external dissertation committee members, which meant a great deal to me. Dr. Dienes is a statistics guru who provided a new angle for looking at my complicated eye-tracking data. I am extremely appreciative of his insightful and valuable advice. I am also grateful to Dr. Rebuschat for his support throughout my research project. His

expertise in implicit and explicit learning helped me develop my research design.

Throughout this arduous work, I always had my family and friends to count on when times were tough. Special thanks are due to my husband Dr. Joon Mahn Lee for his unparalleled patience and support. Without him, I might never have reached this point. He was always beside me during the good and bad moments to push and motivate me. He has shown and taught me what true love is. I must also thank my lovely daughter Jiyoung Joan Lee for giving me the strength and patience to work through every obstacle with which I was presented. She is the most beautiful gift that this world has given me. Special thanks are extended to my parents for their unconditional love and care. Their support was the source of my perseverance over the years. I am forever indebted to my parents for giving me the opportunities and experiences that have made me who I am. Many thanks also go to my parents-in-law for their incredible support and encouragement. They took care of Jiyoung three times a week so that I could focus on writing my dissertation. Had it not been for their support, I could not have completed this work. My thanks also go to everyone in the SLS family—a group of genuinely nice people who want to help one another. I cherish all the good times that we have spent together. I also wish to thank all my friends (especially Saeyoon Kim, Soojin Kim, and Juwon Eum) who supported me during my Ph.D. journey. They were always around when I thought that it was impossible to continue.

This journey would not have been possible if not for my family and friends, and I dedicated this milestone to them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

THE EFFECT OF PRIOR L1 KNOWLEDGE ON THE INCIDENTAL LEARNING OF L2 SYNTAX

A growing number of researchers have considered the acquisition of L2 grammar under incidental, or meaning-focused, learning conditions in order to understand the scope and possibilities of naturalistic adult L2 learning (Denhovska, Serratrice, & Payne, 2016; Godfroid, 2016; Grey, Williams, & Rebuschat, 2014; Kerz, Wiechmann, & Riedel, 2017; Kim & Godfroid, 2019; Leung & Williams, 2012, 2014; Miller & Godfroid, 2019; Morgan-Short, Deung, Brill-Schuetz, Farreta-Stutenberg, Wong, & Wong, 2015; Morgan-Short, Steinhauer, Sanz, & Ullman, 2012; Rebuschat & Williams, 2012; Robinson, 1995, 2005; Rogers, Revesz, & Rebuschat, 2015; Tagarelli, Ruiz, Moreno, & Rebuschat, 2016; Williams & Kuribara, 2008; Williams, 2011). This increased interest in incidental learning conditions reflects one of the core questions in Second Language Acquisition (SLA): How and to what extent is adult L2 acquisition comparable to child L1 acquisition? There is a broad consensus that incidental exposure to a first language (L1) in naturalistic settings uniformly results in children's successful L1 acquisition, which consists of a complete set of intuitive, tacit, and implicit linguistic knowledge. Specifically, children incidentally pick up regularities from the L1 input without intending to learn and eventually master the complex system of L1 without an awareness of what has been learned. In child L1 acquisition, implicit knowledge, which is naturally developed outside the classroom, underlies communicative competence and facilitates fluent language use. Humans' ability to develop such unconscious, implicit knowledge can be remarkably powerful (Dienes, 2008; Reber, 1993) and represents an elementary aspect of human cognition.

In this sense, many SLA researchers have paid special attention to the possibility of

adults' naturalistic L2 learning because the goal of L2 instruction is to develop implicit knowledge for fluent L2 use (Doughty, 2003). Of primary interest is the question of whether adult learners can also acquire a L2 without awareness under meaning-focused learning conditions, as children do in their L1 acquisition. Although some evidence denies the possibility of learning without awareness (e.g., Hama & Leow, 2010; Leow, 2015), and not all studies report evidence for learning without awareness (e.g., Hamrick, 2013; Kim & Godfroid, 2019; Miller & Godfroid, 2019), there is accumulating evidence to support the opposite opinion: that adult learners can develop some unconscious, implicit knowledge of L2 grammar under incidental learning conditions (e.g., Godfroid, 2016; Grey et al., 2014; Leung & Williams, 2011, 2012, 2014; Rebuschat & Williams, 2012; Williams, 2005, 2011; Williams & Kuribara, 2008). This body of work has influenced the field by providing valuable information on adults' capacity to acquire L2 knowledge under meaning-focused, incidental learning conditions.

Although child L1 acquisition sets a good benchmark for adult L2 acquisition, it is imperative to note the differences between child L1 acquisition and adult L2 acquisition. Clearly, adult L2 learners bring prior L1 knowledge to their L2 learning. They already possess L1 representations, which involve symbolic knowledge and linguistic notions, such as subject and predicate; hence, L2 learners are not empty vessels. Given that adult L2 acquisition is built on well-established L1 knowledge and representations, the dominant L1 knowledge may give rise to L1-based habits during L2 acquisitional processes, especially in the case of naturalistic L2 learning. Leung and Williams (2014), for example, demonstrated the effect of L1 on the implicit learning of L2 form-meaning mappings (classifiers) under meaning-focused conditions.

However, the effect of L1 on other linguistic units, such as syntax, has not been fully examined under incidental learning conditions. The role of prior L1 knowledge in the incidental

learning of L2 syntax therefore remains an open question in the field. By examining whether and how L1 knowledge influences L2 acquisition under meaning-focused conditions, this exploration has the potential to make a significant contribution to SLA. It adds to our understanding of whether and how prior L1 knowledge and the implicit and explicit L2 learning of syntax interact. To close this research gap in the present dissertation, I aim to address how prior L1 knowledge affects the development of implicit and explicit syntactic knowledge from a purely meaning-focused task: reading a novel. Reading a novel had advantages over performing a highly controlled, sentence-level task in the context of incidental learning. Specifically, it provided prolonged exposure to target patterns in a more ecologically valid research context and ensured that the participants focused more closely on meaning.

The following general questions guide the present dissertation research: Does prior L1 knowledge affect the incidental learning of L2 syntax consciously, unconsciously, or both? Does prior L1 knowledge affect learning processes as well as learning outcomes under incidental learning conditions? To answer these questions, I compared two L1 groups—English (head initial, right-branching) and Korean (head-final, left-branching)—that were learning a semi-artificial language under incidental learning conditions. The semi-artificial language, *Koreanish*, consisted of English vocabulary and head-final Korean syntax. While the English speakers' L1 syntax had a pattern that was reverse to the target Koreanish syntax (i.e., a cross-linguistic difference), the Korean speakers' L1 syntax mirrored the target Koreanish syntax (i.e., a cross-linguistic similarity). Further, I used multiple measures to track two L1 groups' learning trajectories: eye movements as an online measure of real-time processing and knowledge and grammaticality judgments as an offline knowledge measure. These eye-movement measures were advantageous, in that they enabled me to employ a within-subject design (revealing how

individuals' eye movements change over time) in conjunction with a between-subjects design (showing how much extra reading time the experimental groups spend compared to the control groups). I triangulated online measures derived from eye movements (residualized sentence reading times and grammatical sensitivity) with other offline knowledge measures (grammaticality judgments) and with awareness measures (source attributions, retrospective verbal reports). Together, these methodological components have the potential to reveal dynamic relationships among prior L1 knowledge, awareness, learning, and knowledge under incidental learning conditions.

1.1 The Present Dissertation

For the present dissertation, I investigated the effect of prior L1 knowledge on the implicit and explicit learning of nonnative syntax under incidental exposure. I focused on how prior L1 knowledge not only constrains but also facilitates the development of implicit and explicit L2 knowledge. To do so, I introduced methodological advances into this study. First, I used an authentic novel to incidentally expose participants to the target semi-artificial language. Second, I recorded participants' eye movements while they were reading a novel (Godfroid, Ahn, et al., 2018) and from their eye movement records I extracted online measures of learning and knowledge. A decrease in sentence reading times during training was used as an online measure of learning; an increased sentence reading time to word order violations was used as an online measure of integrated knowledge. I was further interested in the shape of the learning trajectory (i.e., the decrease in sentence reading time) and whether or not it followed the power law of practice (Anderson, 1982; DeKeyser, 1997, 2015; Newell & Rosenbloom, 1981). I triangulated the online measures derived from eye movements with an offline knowledge

measure (GJT) as well as two offline measures of awareness (Rebuschat, 2013; Rebuschat et al., 2015), namely, source attributions (Dienes, 2004; Dienes & Scott, 2005; Dienes, 2008) and retrospective verbal reports. Third, this study is one of few to include a delayed offline GJT to capture the durability of the learning effect (Grey et al., 2014). These methodological devices shed light on the multifaceted effects of L1 transfer during the incidental learning of syntax.

1.2 Overview of the Dissertation

This dissertation is organized into four chapters. In chapter 1, I present the background and motivation for the dissertation and provide definitions of key terms. Chapter 2 focuses on the offline measure of grammaticality judgments, and chapter 3 concerns the online measures derived from the eye movements. Each chapter presents a literature review, research questions, methods, results, and a discussion. Finally, in chapter 4, I summarize the findings of the study and present a general discussion.

1.3 Definition of Key Terms

In this section, I provide the definitions of key constructs in the dissertation and the corresponding measurement variables used to operationalize them (Table 1.1). The key constructs include *learning, power law of practice, knowledge, integrated knowledge, grammatical sensitivity, awareness, incidental learning, implicit knowledge*, and *explicit knowledge*.

Table 1.1

Definition of key terms

	Definition	Measurement variable
Learning	The process of acquiring new knowledge from the input.	A decrease in sentence reading times during the exposure task.
Power law of practice	The ubiquitous decrease in the time required to perform a task that follows a power function ($Y = X^{-n}$) involving the process of automatization.	A decrease in the sentence reading times that follows the power function.
Knowledge	The outcome of the learning process.	(a) Online: grammatical sensitivity An increased reading time in response to syntactic violations during the exposure task. (b) Offline: grammaticality judgments Above-chance performance on grammaticality judgments after the exposure task.
Integrated knowledge	The mental representation of the knowledge that underlies automatic competence.	An increased reading time in response to syntactic violations during the exposure task.
Grammatical sensitivity	The ability to apply integrated knowledge of syntax in an automatic manner.	An increased reading time in response to syntactic violations during the exposure task.

Table 1.1 (cont'd)

Awareness	The conscious perception of what is being learned.	(a) Offline: source attributions in the GJT Above-chance performance on grammaticality judgments attributed to recollection and rule knowledge. (b) Offline: retrospective verbal reports An ability to verbalize the target rules.
Incidental learning	Learning conditions in which participants are exposed to target rules through a meaning-focused activity without a prior notice of testing.	Reading for comprehension a novel in which the target rules are embedded throughout the text.
Implicit knowledge	Unconscious knowledge that exists outside of the participants' awareness.	(a) Online: sensitivity to violations An unconscious slowdown in sentence reading times in response to syntactic violations during the exposure task. (b) Offline: Unaware learners' grammaticality judgments; grammaticality judgments based on guessing and intuition A comparison of the unaware group's GJT performance with that of the control group; the unaware group's above-chance performance. Above-chance performance on grammaticality judgments attributed to guesses and intuition.

Table 1.1 (cont'd)

Explicit knowledge	Conscious knowledge that the participants are aware that they know.	(a) Offline: retrospective verbal reports An ability to verbalize the target rules. (b) Offline: grammaticality judgments based on recollection and rule knowledge Above-chance performance on grammaticality judgments attributed to recollection and rule knowledge.

CHAPTER 2

OFFLINE MEASURE: GRAMMATICALITY JUDGMENTS

2.1 Background

2.1.1 Development of implicit and explicit knowledge in incidental learning conditions

Many studies targeting the incidental learning of non-native grammar have shown that learners can acquire some knowledge of the grammar without awareness (Godfroid, 2016; Grey, Williams, & Rebuschat, 2014; Kerz, Wiechmann, & Riedel, 2017; Rebuschat & Williams, 2012; Robinson, 1995; Rogers, Revesz, & Rebuschat, 2015; Williams, 2011; Williams & Kuribara, 2008; but for an opposing view, see Hama & Leow, 2010; Leow, 2015b). In general, this body of work has exposed participants to a semi-artificial language through a meaning-focused task (the *training phase*) and then tested their knowledge of the target grammar without prior notice (the *testing phase*). After measuring the participants' awareness of the target grammar rules, the researchers investigated whether the unaware participants showed learning effects on the knowledge measure, such as GJTs, to investigate the possibility of implicit learning. The findings from previous research have suggested that adult learning of non-native syntax can take place in the absence of awareness of the target rules under incidental exposure conditions (but see Andringa, in press; Curcic, Andringa, & Kuiken, 2019; Hamrick, 2013; Kim & Godfroid, 2019; Miller & Godfroid, 2019), and that implicit learning without awareness is just slightly above chance on the knowledge measure. The latter result leads one to ask why the observed effect of implicit learning was so small. One answer may lie in the multiple mediating factors within the experimental design.

One such factor that possibly affects incidental learning is the nature of exposure that learners receive during training. For example, recent studies found that the frequency

(Denhovska et al., 2016) and linguistic complexity (Tagarelli, Ruiz, Moreno, & Rebuschat, 2016) of the input given to learners play a role in implicit learning. A related factor that has yet to receive ample attention is the type of exposure task. In general, researchers have used meaning-focused, psycholinguistic tasks as exposure activities, such as plausibility judgments (e.g., Rebuschat & Williams, 2012; Williams, 2011) or sentence-picture matching tasks (e.g., Godfroid, 2016; Leung & Williams, 2011). Such tasks can be advantageous because they allow researchers to exert tight control over the stimuli in terms of sentence type or word frequency. However, the amount of input learners receive from such tasks may not be enough to engender robust learning. Furthermore, processing isolated single sentences is arguably a somewhat artificial task. As a growing body of language-processing research (Brennan, Hasson, Malach, Heeger, & Pylkkänen, 2012; Speer, Reynolds, Swallow, & Zacks, 2009; Nijhof & Willems, 2015; Willems, Frank, Nijhof, Hagoort, & Van den Bosch, 2015) and incidental vocabulary learning research (Elgort, Brysbaert, Stevens, & Van Assche, 2018; Elgort & Warren, 2014; Godfroid, Ahn, et al., 2018; Mohamed, 2018; Pellicer-Sanchez & Schmidt, 2010) suggests, using more naturalistic stimuli in longer texts could overcome the problem of artificiality in psycholinguistic experiments.

Reading a novel, for example, has three advantages for incidental learning over performing a highly controlled, sentence-level task. First, a novel is more engaging and contains more meaningful content, which ensures a stronger participant focus on meaning. Second, a novel provides exposure to target patterns over a greater length of time, and such rich experiences of the target patterns may increase the probability of learning. Third, reading a novel represents a more natural reading task and resembles an everyday language learning activity; thus, it is ecologically valid (Godfroid et al., 2018). Hence, a longer text from a novel can be

considered a promising medium to convey the targeted syntax for the purpose of exploring the implicit and explicit learning of syntax. In light of this, I opted to use novel reading as an exposure task in this dissertation to train participants incidentally on target syntactic patterns in a more ecologically valid research context.

Another factor that has been overlooked is the passage of time, that is, the delayed effect of incidental exposure (Grey et al., 2014; Morgan-Short et al., 2012; Robinson, 2002). Although Norris and Ortega (2000), in their meta-analysis on the effectiveness of L2 instruction, reported a decline in the L2 learning effect from immediate tests to delayed tests (but for a contrasting view, see Spada & Tomita, 2010), this tendency apparently does not always occur, particularly in the incidental learning of L2 grammar. Morgan-Short et al. (2012), in an event-related potentials study, found positive evidence for L2 grammar development from implicit training, which was reflected in the consolidation of knowledge over several months. Both instructional groups showed increased nativelike neural processing, even after a substantial, three-to-five-month, delay with no further exposure. Further, the implicit instruction group, who received incidental exposure to Brocanto2 only, demonstrated greater nativelike syntactic processing than the explicit learning group, who received incidental exposure and a metalinguistic explanation, before and after the delay. This indicated that the incidental exposure was likely linked to a more nativelike neurocognitive processing.

Grey et al. (2014) demonstrated the delayed effects of incidental exposure on the implicit learning of L3 case marking and word order. In their study, the researchers found that learning from incidental exposure was not only retained, but also showed improvement, in the delayed test, which was administered two weeks after the immediate test. Additionally, awareness of the rules seemed to play a crucial role in acquiring word order. Specifically, awareness of the correct

rules appeared to increase the knowledge of word order, as demonstrated in the delayed test. Based on these results, the authors pointed out the importance of delayed testing to explore the effect of time on the incidental learning of grammar. Building on Morgan-Short et al. (2012) and Grey et al. (2014), in this dissertation I too incorporate delayed testing to capture the changes in L2 grammar development over time.

2.1.2 The effect of L1 on implicit and explicit knowledge development in incidental learning conditions

The role of L1 in L2 acquisition has been studied extensively over three decades of SLA research (e.g., Clahsen & Felser, 2006; Gass, 1979, 1984; Ellis & Sagarra, 2011; Jarvis & Pavlenko, 2008; MacWhinney, 2005; Ringbom, 2007; Tolentino & Tokowicz, 2011; White, 2000). For example, scholars taking a generative approach based on the linguistic theory of universal grammar (UG) (Chomsky, 1981) have had a longstanding interest in the interaction between UG and L1 transfer (for a review, see White, 2000). Recently, not only UG researchers but also usage-based researchers have begun to emphasize the role of L1 prior knowledge in implicit and explicit L2 learning (Brooks & Kempe, 2013; Godfroid, 2016; Leung & Williams, 2014; Onnis & Thiessen, 2013; Tolentino & Tokowicz, 2014; Williams, 2005; Williams & Kuribara, 2008). Usage-based researchers have increasingly acknowledged the importance of the L1 effect on L2 learning because L2 learners, especially in the case of natural L2 learning, bring prior L1 knowledge to their L2 learning, which may yield L1-based habits (e.g., Ellis & Sagarra, 2011; MacWhinney, 2008). In this regard, L2 learners are not empty vessels; that is, learners' domain-specific, prior L1 knowledge could have an impact on learning that relies on domain-general learning mechanisms. The role of L1 in L2 syntactic acquisition is a fruitful avenue for research because it could open a productive dialogue between the usage-based and generative

approaches (VanPatten & Rothman, 2015; Williams & Kuribara, 2008). This exploration has the potential to make a significant contribution to SLA theory because it helps researches to gain a better understanding of how and to what extent domain-general and domain-specific learning mechanisms can account for L2 learning.

From the usage-based perspective, Ellis, Sagarra, and their colleagues (Cintrón-Valentín & Ellis, 2015; Ellis & Sagarra, 2010, 2011; Ellis, Hafeez, Martin, Chen, Boland, & Sagarra, 2014) viewed learned attention as a form of L1 transfer in L2 processing and acquisition. The central idea of their argument is that a prior L1 experience can direct L2 attentional processing, either positively or negatively. On one hand, L2 learners would pay more attention to familiar cues based on their prior L1 experience, processing them more fluently and automatically. On the other hand, they would block their attention to unfamiliar and foreign cues, resulting in L2 processing biases. For example, Ellis and Sagarra (2010) Experiment 2 and Ellis and Sagarra Experiments 2 and 3 consistently demonstrated how L1 backgrounds have led to attentional biases in L2 processing. They showed that when equal amounts of adverbial and verbal cues were presented in an input, Chinese speakers, whose L1 lacked morphological markings, experienced difficulties in acquiring verbal inflectional cues compared with Russian and Spanish speakers, whose respective L1s are rich in morphological markings. These findings indicated that learned attention in L1 may exert a major influence on subsequent L2 processing and acquisition.

Researchers have shown that L1 grammatical concepts that are relevant to L1 form-meaning mappings can be transferred in L2 implicit learning (Brooks & Kempe, 2003; Leung & Williams, 2014; Williams, 2005; Williams & Lovatt, 2003). Williams (2005), in a seminal SLA study on learning without awareness, provided initial evidence of L1 influence in implicit learning. Participants whose L1 marked noun genders with determiners tended to be better at

learning targeted determiner-noun mappings in a semi-artificial language than participants whose L1 did not have a gender-marking system. Extending Williams (2005), Leung and Williams (2014) probed the L1 influence on the implicit learning of determiner-noun mapping based on semantic relationships. In their study, native speakers of English and Chinese participated in three reaction-time experiments that targeted semantic concepts such as animacy, the number of capital letters and strokes, and the long/flat distinction. Four artificial determiners were introduced as target forms: *gi*, *ro*, *ul*, and *ne*. *Gi* and *ro* were used with nearby objects, while *ul* and *ne* were used with faraway objects, and this mapping was explicitly described to participants. However, participants were not informed about an additional, hidden mapping rule. In particular, the determiners also varied by animacy (i.e., animate *vs*. inanimate) (Experiment 1), the number of capital letters and strokes (Experiment 2), and the long/flat distinction (Experiment 3). An interesting crosslinguistic influence was found in Experiment 3. In Experiment 3, the learning target (i.e., the long/flat distinction) was a concept exemplified in the Chinese classifier system. The Chinese speakers, whose L1 encoded this regularity in their grammar, slowed down their responses to article usage based on distance (near/far) when the hidden long/flat rule was violated. However, the English speakers, who did not have this distinction in their L1, did not show such sensitivity to the long/flat violation. Leung and Williams interpreted this as evidence for the role of L1 constraints in the implicit learning of semantics. More generally, their results evidenced the importance of the L1 effect in implicit learning. There is evidence for L1 transfer in the acquisition of semantics (Leung & Williams, 2014) and morphology (Cintrón-Valentín & Ellis, 2015, Ellis, Hafeez, Martin, Chen, Boland & Sagarra, 2014; Ellis & Sagarra, 2011; Sagarra & Ellis, 2013; Williams, 2005) but not in other areas such as phonology, syntax, and pragmatics. To help close this gap, in the present study, I investigate the effect of prior L1 syntactic

knowledge on the incidental learning of L2 syntax.

To demonstrate L1 syntactic transfer, many researchers targeted word order and head-direction (Onnis & Thiessen, 2013; Williams & Kuribara, 2008; Williams, 2011). Williams and Kuribara (2008) used word order and head direction to investigate the characteristics of the initial stage of L2 UG-guided and L2 frequency-guided learning. In their study, English native speakers, whose L1 is a head-initial language, were exposed to Japlish, a semi-artificial language consisting of English vocabulary and Japanese syntax. The participants completed a meaning-focused task without any instruction or feedback. The Japlish syntactic structures in the exposure phase comprised canonical SOV sentences primarily, along with some scrambled OSV sentences that involved optional movement. After exposure, participants took a surprise GJT on new sentences containing canonical and scrambled word orders. There was evidence for the generalizability of knowledge of canonical patterns without awareness, but the picture regarding scrambled patterns was not so clear. Although a subset of participants who accepted trained scrambled structures also accepted unscrambled complex sentences, they failed to accept untrained scrambled simple sentences. This performance did not support the clustering effects that parameter resetting should entail. If the head direction parameter had been reset under the guide of a UG, all the underlying related structures should have been acquired, including scrambled structures. Based on these results, researchers concluded that the initial stage of adult L2 learning is not UG-guided.

Within the field of cognitive psychology, Onnis and Thiessen (2013) documented English and Korean speakers' L1-induced biases on their statistical learning of sequential information. Statistical learning involves humans' sensitivity to frequencies, probabilities, and regularities in an environment. Implicit learning and statistical learning share many similarities, in terms of

research area (i.e., how humans unconsciously gain information from the input) and research methodology (i.e., the use of the artificial grammar paradigm), and thus the two approaches are sometimes equated (Conway & Christiansen, 2006; Kuhn & Dienes, 2008; Monaghan, Schoetensack, & Rebuschat, 2019). Drawing on the fact that Korean is head-final and English is head-initial, Onnis and Thiessen (2013) conducted a corpus analysis on Korean and English corpora, which confirmed their prediction that the opposite head direction tendencies in the Korean and English word orders could give rise to reverse patterns in the two groups' syntactic parsing behavior. With this crosslinguistic evidence established, the authors compared how Korean and English native speakers sequence complex linguistic and non-linguistic stimuli. The results suggested that participants showed a parsing preference for patterns that matched their L1 word-order patterns—in other words, statistical learning was constrained by the participants' L1 characteristics. A question that remains is to what extent these findings from the statistical learning literature also apply to more natural linguistic materials (for cautionary results, see Kim & Godfroid, 2019). To expand our understanding of these issues, in the present study I compare the performance of English and Korean speakers in their learning of a semi-artificial language consisting of Korean word order and English vocabulary.

2.1.3 Research questions

In this chapter, I will present the answers to the following research questions:

RQ1 (incidental exposure)

Does incidental exposure facilitate the acquisition of syntactic knowledge? Do the experimental group and the control group differ in their performance on the offline knowledge measure?

RQ2 (prior L1 knowledge)

Does L1 affect syntactic knowledge development under incidental exposure conditions? Do the

English experimental group and the Korean experimental group differ in their performance on the offline knowledge measure?

RQ3 (time)

Is the acquired knowledge durable? Does the experimental groups' performance on the offline knowledge measure change over time?

RQ 4 (awareness reflected in retrospective verbal reports)

4-1. Is there evidence of learning without awareness? Can unaware learners acquire syntactic knowledge from incidental exposure?

4-2. Do aware and unaware, English and Korean subgroups differ in their performance on the offline knowledge measure?

RQ 5 (awareness reflected in source attributions)

5-1. What is the nature of the acquired knowledge? To what extent is it implicit or explicit?

5-2. Do the English and Korean speakers differ in the development of implicit and explicit knowledge?

2.2 Methods

2.2.1 Participants

In this study, I compared two language groups: English speakers, whose L1 has a head-initial structure, and Korean speakers, whose L1 has head-final word order. English speakers (n = 40, M_{age} = 23.71, SD = 3.90) and Korean speakers (n = 40, M_{age} = 25.25, SD = 4.42) were recruited from three universities: one large Mid-western university in the United States and two large universities in South Korea). All participants held at least bachelor's degree or were currently enrolled as bachelor's students in a university.

Each language group was subdivided into an experimental and a control group. This resulted in four subgroups: (a) an English experimental group (EE, *n* = 25), (b) an English control group (EC, *n* = 15), (c) a Korean experimental group (KE, *n* = 25), and (d) a Korean control group (KC, *n* = 15). The experimental groups, EE and KE, read the novel *The Mysterious Affair at Styles* (Christie, 1920) in a semi-artificial language that consisted of English vocabulary and Korean word order. The control groups, EC and KC, read the English version of the novel; that is, the same novel with English vocabulary and English word order. The use of control groups was important because they yielded baseline data for the GJT scores. As Hamrick and Sachs (2017) rightly pointed out, the use of statistical chance as a comparison in incidental learning research had some limitations due to "the potential for participants to show preexisting biases or to learn during the test phase" (p. 15). We cannot exclude the possibility that using chance as a baseline will pose a threat to the internal validity of the study. In this study, therefore, the control group's grammaticality judgments were employed as the baseline, which allowed a more accurate assessment of learning effects.

The English speakers had no background in Korean or any other head-final languages, such as Japanese or Turkish. Further, the experimental groups did not have any knowledge of German, a language that was used to create violation blocks in the testing phase during the exposure task (see below). The Korean speakers were unbalanced Korean-English bilinguals who were highly proficient in English (see Table 1). At the time of the research, 19 Korean speakers resided in the United States, whereas the remaining 31 Korean participants resided in South Korea. They had a mean iBT TOEFL score of 111.67 (*SD* = 4.65) and a mean length of residence in an English-speaking country of 7.89 years (*SD* = 3.27). All Korean speakers reported that their L1 was Korean and their L2 was English; however, they also reported that

they had no difficulty in reading English novels or communicating in English. Table 2.1 summarizes information about the Korean speakers' English learning backgrounds and proficiency levels. There was no difference between the KE and KC groups in their age of exposure (U = 180.00, p = .847, r = 0.01), length of residence (U = 174.50, p = .720 r = 0.05), TOEFL score (U = 127.00, p = .474 r = 0.16), and self-rated proficiency (U = 170.00, p = .639, r = 0.10). The participants were recruited by flyers, web postings, and word of mouth. They were compensated with $30 at the end of the experiment.

Table 2.1.

Korean speakers' English learning backgrounds and English proficiency

		KE (n = 25)		KC (n = 15)	
		M	*SD*	*M*	*SD*
Age at testing		25.32	4.61	25.13	4.24
Age of exposure		5.80	1.76	5.93	1.67
Length of residence (years)		7.74	3.26	8.12	3.38
TOEFL score		112.22	4.99	110.69	3.97
Self-evaluation of proficiency	Total	34.96	1.79	34.53	2.50
	Listening	8.96	0.68	8.80	1.01
	Reading	8.88	0.78	8.60	0.78
	Speaking	8.56	0.82	8.47	1.13
	Writing	8.56	0.87	8.67	0.98

2.2.2 Targeted semi-artificial language: Koreanish

The present study adopted a semi-artificial language paradigm, in which vocabulary from

the participants' native language was rearranged following the patterns in the target language. Specifically, the semi-artificial language in this study consisted of English vocabulary and Korean syntax. To create this hybrid language, which I will call Koreanish, I rearranged the English words in the English version of the Agatha Christie novel *The Mysterious Affair at Styles* (Christie, 1920) according to Korean word order. Therefore, from the participants' perspective, the words were familiar, but the syntax was an unknown system that posed comprehension difficulty.

The word orders in Korean and English often look like mirror images of each other because the two languages order heads and complements differently. Syntactically, English is a head-initial language. The heads of phrases precede their complements, thus forming a right-branching structure. Korean is a head-final language. Heads follow complements, thus forming a left-branching structure. To illustrate, the English sentence "John ate an apple" is glossed as "John apple ate" in Korean because the verb (head) comes at the end of the phrase, and its complements are naturally located to the left. Likewise, "at church" is glossed as "church at" because the preposition (head) comes at the end of the phrase; that is, Korean is postpositional.

Based on Korean word order, I used four syntactic rules regarding Korean's head-finality to generate the sentence stimuli. First, in the verb phrase (VP), the order of elements is complement-verb. Second, in the postpositional phrase (PP), the order of elements is noun phrase (NP)-postposition. Third, in the complementizer phrase (CP), the order of elements is clause-complementizer (e.g., relative pronoun, *that*, *whether*). Fourth, in the NP, the order of elements is complement (e.g., appositive clause, relative clause)-noun. The following sentences are an actual example from the novel and show how an English sentence (1), (3) can be rearranged according to Koreanish rules, as shown in (2), (4), respectively.

(1) English: A vague suspicion of everyone and everything filled my mind.

(2) Korean: Everyone and everything of a vague suspicion my mind filled.

First, the PP “of everyone and everything”” is transformed into “Everyone and everything of ” because “of,” the apposition, is the head of the phrase and comes in final position. Second, the NP “a vague suspicion everyone and everything of” is reordered into “Everyone and everything of a vague suspicion” because the NP “a vague suspicion” serves the head of the phrase. The NP “Everyone and everything of a vague suspicion” becomes the subject of the sentence. Third, the verb “filled” in the VP “filled my mind” should be moved to the end, which results in SOV word order.

(3) English: “I have a cousin who is a nurse,” I remarked.

(4) Koreanish: “I a nurse is who a cousin have,” I remarked.

First, in the VP “is a nurse,” the verb “is,” the head, should come at the end, thus forming “a nurse is.” Second, the CP “who is a nurse” should be “a nurse is who” because “who,” the relative complementizer, is the head of the CP and should be placed at the end of the phrase. Third, “a nurse is who” should modify the head, “a cousin” by preceding rather than following it. This results in the relative clause – NP sequence “a nurse is who a cousin.” Fourth, the complex NP “a nurse is who a cousin” should merge with the verb “have,” thus forming the VP. In the VP, the verb “have,” the head, should be placed at the end of the phrase. Hence, the full sentence reads “I a nurse is who a cousin have.” In this manner, Koreanish consistently features the head,

the main element of the phrase, at the final position.

2.2.3 Materials

The exposure task: Novel reading. The exposure text for training was the novel *The Mysterious Affair at Styles* by Agatha Christie. This novel provided incidental exposure to the semi-artificial word order system, Koreanish, to participants. The participants read the first two chapters, Chapters 1 and 2 (approximately 6,500 words), of *The Mysterious Affair at Styles*. The same novel was used in a series of eye-tracking studies that examined various aspects of English monolinguals' and Dutch-English bilinguals' reading processes (Cop, Dirix, Drieghe, & Duyck, 2017; Cop, Dirix, Van Assche, Drieghe, & Duyck, 2017; Cop, Drieghe, & Duyck, 2015; Cop, Keuleers, Drieghe, & Duyck, 2015). Cop and colleagues noted that this novel was selected based on its appropriate difficulty level for college level L2 English speakers (Flesch Reading Ease = 81.3, SMOG grade = 7.4)[1] as well as its similarity in word frequency distribution to natural language based on the Subtlex database (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014) In other words, this novel is a suitable reading material for a L2 experiment because it is easy to read and reflects real-life, natural language use.

The participants were told that they were going to read the first two chapters of a detective novel (92 screens), and that—for some of the participants—the sentences would be presented with the words scrambled. The experimental groups read the Koreanish version of the novel, consisting of English sentences with Korean word order, while the control groups read the English version. Before reading, the participants were given a brief description of the novel's

[1] Flesch Reading Ease gives a scale between 0 (difficult to read) and 100 (easy to read); higher scores reflect greater readability. The SMOG grade is an estimation of how many years of educations are required to understand the text.

main characters to help them understand the story. The participants were informed that there would be a practice session at the beginning and a short break approximately 15 screens later (6 breaks in total). At every other break, the participants were asked to answer comprehension check questions. The purpose of the comprehension test was to keep participants engaged while reading the scrambled sentences for about an hour. These comprehension tests, which contained eight simple true or false statements regarding plot-specific information, were administered three times throughout the reading. The total number of questions was 24. One point was assigned for correct answers, and the total score of overall comprehension was 24. The overall reliability coefficients (Cronbach's alpha) were 0.78 for the EE group and 0.71 for the KE group. Three counterbalanced versions of the exposure text were created to control for lexical and topical confounds (Table 2.2).

Table 2.2

The counterbalanced versions of the exposure task

Phase	Syntax	*N* of sentences	Sentence ID		
			Version A	Version B	Version C
Practice	English	44	544 – 583, 1 – 4	564 – 583, 1 – 24	1 – 44
Training	Koreanish	479	5 – 483	25 – 503	45 – 523
Testing: Control_pre	Koreanish	20	484 – 503	504 – 523	524 – 543
Testing: Violation	German	20	504 – 523	524 – 543	544 – 563
Testing: Control_post	Koreanish	20	524 – 543	544 – 563	564 – 583

The testing task: GJT. To test the participants' acquired knowledge of Koreanish, a GJT with source attributions was implemented after the exposure task. Importantly, the participants were not informed ahead of time that they would be tested. After reading the novel, they were told that the scrambled order was not arbitrary but followed a complex system. They were asked to judge the grammaticality of the new sentences based on the system presented in the novel (see Appendix C). No feedback was provided regarding their answer.

In the GJT, 80 new sentences were presented to the participants, distributed evenly between 40 grammatical and 40 ungrammatical items (Table 2.3). Five grammatical Koreanish patterns were created, including two simple and three complex patterns:

(a) simple sentence: SOV;

(b) simple sentence with postpositional phrase: SPP(*postposition*)V,

(c) complex sentence with a that-clause: S[SOV*that*]V,

(d) complex sentence with a relative clause: S[OV*relative pronoun*]OV;

(e) complex sentence with a subordinate clause: SOV*subordinator*, SOV.

For each grammatical structure, two ungrammatical patterns were created as direct counterparts:

(a) *SVO and (b) *VSO for the SOV pattern;

(c) *S[*that*SOV]V and (d) *SV[SOV*that*] for the S[SOVthat]V pattern;

(e) *S[PP*preposition*]V and (f) *SV[PP*postposition*] for the S[PP*postposition*]V pattern;

(g) *S[*relative pronoun*OV]OV and (h)*SO[OV*relative pronoun*]V for the S[OVrelative pronoun]OV pattern;

(i) **subordinator*SOV, SOV and (j) *SOV, SOV*subordinator* for the SOVsubordinator, SOV pattern.

In the case of the complex sentences, one ungrammatical pattern contained the clause-level error

(within the dependent clause) and the other ungrammatical pattern included the sentence-level error (within the independent clause).

First, the SOV pattern is a simple sentence with basic Koreanish word order and reflects the VP rule. Both *SVO and *VSO contrast with the SOV pattern. The *SVO pattern follows English word order, so it allowed me to test whether the participants knew that English-like word order was not possible in Koreanish. The *VSO pattern also enabled me to assess whether participants could reject the cases that did not conform to the VP rule.

Second, the S[PP(*postposition*)]V pattern represents a simple sentence that features a postpositional phrase. It tests knowledge of the PP and the VP rules. The ungrammatical patterns for this structure are *S[PP(*preposition*)]V and *SV[PP(*postposition*)]. The *S[PP(*preposition*)]V pattern tests the PP rule by allowing me to determine whether participants knew that the apposition, the head of the PP, should come at the end of the phrase. The *S[PP(*preposition*)]V pattern allowed me to test the VP rule that the verb should come in final position in a verb phrase.

Third, the S[SOV*that*]V pattern is a complex sentence with a *that*-clause involving the CP and VP rule. The ungrammatical patterns for this structure are *S[*that*SOV]V and *SV[SOV*that*]. *S[*that*SOV]V assesses the CP rule applied at the clause level. It tests whether participants knew that the complementizer *that* should come at the end of the complement clause. *SV[SOV*that*] measures the VP rule at the sentence level, by testing whether participants kneow that the verb should be in sentence-final position.

Fourth, the S[OV*relative pronoun*]OV pattern is a complex sentence that features a relative clause, which reflects the CP and the NP rules. The ungrammatical patterns for this structure are *S[*relative pronoun*OV]OV and *SO[OV*relative pronoun*]V. The *S[*relative*

*pronoun*OV]OV pattern allowed me to measure knowledge of the CP rule at the clause level. It tests whether participants knew that the relative pronoun, the head of the CP, should follow complements and come at the final position in the relative clause. The *SO[OV*relative pronoun*]V assesses the NP rule at the sentence level. It assesses whether participants knew that the relative clause, as a complement, should precede the head noun it modifies.

Fifth, the SOV*subordinator*, SOV pattern is a complex sentence with a subordinate clause. It involves the CP and VP rules. The ungrammatical patterns for this structure are **subordinator*SOV, SOV and the *SOV, SOV*subordinator*. The **subordinator*SOV, SOV pattern tests the CP rule at the clause level. Within the subordinate clause, the subordinator is the head that should follow complements. The *SOV, SOV*subordinator* pattern assesses the knowledge of the VP rule at the sentence level. It enabled me to see whether participants knew that the verb in the main clause should come at the end of the sentence as the head of the VP.

In addition to grammaticality judgments, participants were asked to indicate the basis of their decision: *guess*, *intuition*, *recollection*, or *rule knowledge*. This source ratings were useful to probe the nature of participants' knowledge, whether it is implicit or explicit (Dienes, 2004; Dienes & Scott, 2005; Rebuschat, 2013; Spinner & Gass, 2019). Participants were informed to select the "guess" category if they believed their decision was a true guess, meaning they might as well have flipped a coin. If a participant had some confidence in their decision and knew, to some degree, that the judgment was correct but could not describe why, they were told to opt for the "intuition" category. Participants were asked to use the "recollection" category if the decision was based on the memory of specific sentences (or parts of the sentences) that they read in the exposure phase. Finally, they were asked to select the "rule knowledge" category if they followed a verbalizable rule when making their decision. Eight lists of the GJT were created,

counterbalanced for the grammaticality and the presentation order of the stimuli. The reliability of the GJT was good, with Cronbach's alpha ranging from $\alpha = 0.84$ on the immediate test to $\alpha = 0.89$ on the delayed test.

Table 2.3

Counterbalanced grammatical and ungrammatical stimuli used in the testing set

Sentence type	Grammatical Pattern	Ungrammatical pattern	Target rule
Simple, basic	SOV (*k* = 8) e.g., Joon the paper revised.	*SVO (*k* = 4) e.g., Joon revised the paper.	VP rule
		*VSO (*k* = 4) e.g., Revised Joon the paper.	VP rule
Simple, postposition	S[PP*postposition*]V (*k* = 8) e.g., Max the classroom in studied.	*S[*preposition*PP]V (*k* = 4) e.g., Max in the classroom studied.	PP rule
		*SV[PP*postposition*] (*k* = 4) e.g., Max studied the classroom in.	VP rule
Complex, that-clause	S[SOV*that*]V (*k* = 8) e.g., The dean he his salary donated that lied.	* S[*that*SOV]V (*k* = 4) e.g., The dean that he his salary donated lied.	CP rule
		*SV[SOV*that*] (*k* = 4) e.g., The dean lied he his salary donated that.	VP rule
Complex, relative clause	S[OV*relative pronoun*]OV (*k* = 8) e.g., Jessie a blue coat wore who the man found.	*S[*relative pronoun*OV]OV (*k* = 4) e.g., Jessie who a blue coat wore the man found.	CP rule
		*SO[OV*relative pronoun*]V (*k* = 4) e.g., Jessie the man a blue coat wore who found.	NP rule

Table 2.3 (cont'd)

Complex, subordinate clause	SOV*subordinator*, SOV (k = 8) e.g., We dinner enjoyed while, the band music played.	**subordinator*SOV, SOV (k = 4) e.g., While we dinner enjoyed, the band music played.	CP rule
		*SOV, SOV*subordinator* (k =4) e.g., The band music played, we dinner enjoyed while.	VP rule

Debriefing interview. The purpose of the debriefing interview was to probe participants' levels of awareness of the targeted Koreanish syntactic patterns (see Appendix D). The interview was carried out over two sessions. At the end of Session 1, the first part of the debriefing interview was held to examine participants' awareness of the violation block. I asked the participants whether they had noticed anything odd while reading the novel. In addition, I asked them to report whether they had noticed any particular rule or regularity, to specify when they might have noticed it (i.e., during reading or on the test), and to describe what they believed they had noticed.

In Session 2, the second part of the debriefing interview was conducted. I began by asking whether they ever indicated recollection of rule knowledge as a source of their grammaticality judgement. If so, they were asked to explain why and what they were thinking. In addition, I asked participants how they attempted to read the sentences with the words scrambled and whether they had tried to search for a pattern in the scrambled word order while reading. Moreover, I asked them to reflect specifically on the placement of words within the sentences and to recall any specific rule or regularity in an attempt to tap into participants' lower awareness levels. Finally, I asked them questions about their usual reading experiences.

2.2.4 Procedure

The experiment was conducted in two sessions, Session 1 (90 minutes) and Session 2 (30 minutes). Session 1, which was carried out in a quiet, dimly-illuminated study room, included the exposure phase, the immediate testing phase, and the first part of the debriefing interview (short version). Session 2 comprised the delayed testing phase, the second part of the debriefing interview (long version), and the background questionnaires (see Figure 2.1).

In Session 1, the participants first signed the consent form and then filled out a language

background questionnaire. After being told that they were going to read two chapters from a detective novel, a brief description of the main characters was provided. They were informed that there would be several comprehension questions after reading each part (Appendix B); however, they were not told that there would be a GJT about the word order when they completed reading the novel. They silently read the first two chapters of *The Mysterious Affair at Styles* on a laptop screen while their eye movements were recorded by the Eyelink Portable Duo eye-tracking system (SR Research, Ottawa, Ontario, Canada). The experimental groups read the novel in Koreanish and the control groups read the English version of the novel. After they finished reading the novel, a surprise GJT with source attributions was administered for immediate testing. This test was followed by the first part of the debriefing interview regarding awareness of the violation block in the reading (please see Table 2.2, and the reading time data will be presented in chapter 3) and Koreanish word order.

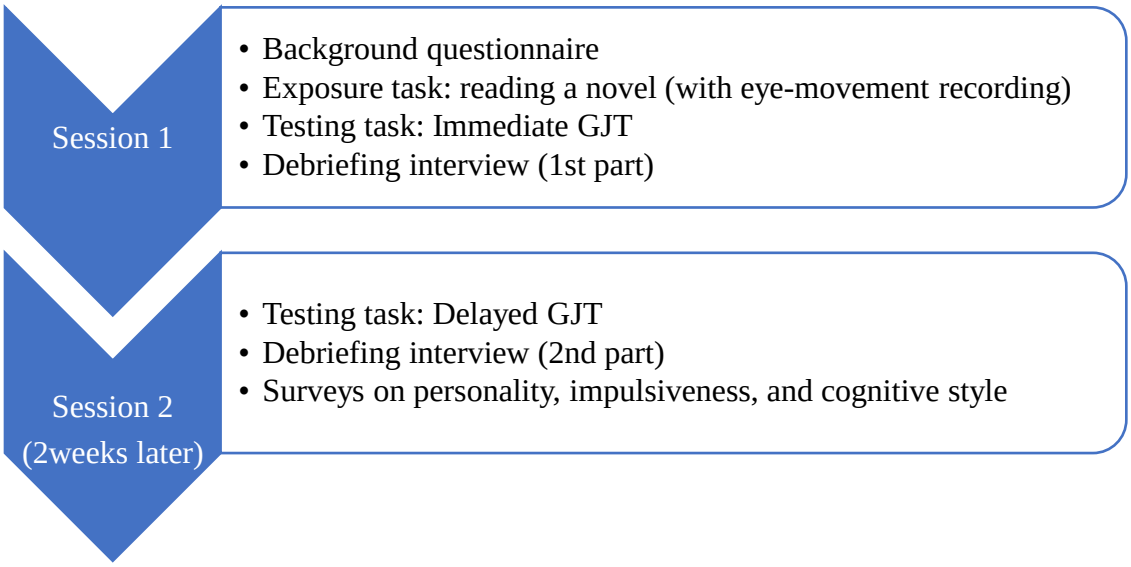

Figure 2.1

Procedure in the experiment

Session 2 took place two weeks following Session 1 for delayed testing. To each participant, the researcher sent an e-mail containing the links to the delayed test, the second part of the debriefing interview, and the surveys regarding personality (Big Five personality test, De Young, Quility, & Peterson, 2007), impulsiveness (the UPPS, Whiteside & Lynam, 2001), and cognitive style (Rational Experimental Inventory, Pacini & Epstein, 1999). The participants were asked to complete the test within three days.

2.2.5 Analysis

All the data from the testing task (grammaticality judgements) were entered into the Statistical Package for Social Sciences, version 25. Two participants, one from the EE group and one from the KE group, were excluded from the analyses, leaving 24 EE and 24 KE. The excluded participants stated in their retrospective verbal reports that they intentionally searched for word order rules during reading, in violation of the incidental nature of the study. For the delayed GJT, analyses were conducted on 72 participants (23 EE, 14 EC, 22 KE, 13 KC) since 6 participants (1 EE, 1 EC, 2 KE, 2 KC) did not participate in Session 2.

Participants' responses were transformed to d-prime (*d'*) scores—a sensitivity index that reflects participants' ability to discriminate between grammatical and ungrammatical sentences. The *d'* score is known to be a more accurate measure than raw accuracy in that it takes response bias into account based on hits and false alarm rates (MacMillan & Creelman, 2005); therefore, where appropriate, *d'* scores will be reported[2]. A *d'* score of zero is equivalent to chance performance and a *d'* of four is interpreted as near-perfect performance. That is, a positive *d'*

[2] With hit rates and false alarms of 0 and 1, I did a standard correction to compute the *d'* scores. Given the maximum number of false alarms of 40, the extreme values (0 and 1) were strategically replaced with $1/(2 \times 40) = 0.0125$ and $1-1(2 \times 40) = 0.9875$.

score indicates above chance-level performance, whereas a negative *d'* indicates below chance-level performance. In this study, a higher *d'* score indicates participants' superior ability to discriminate target-like word order in Koreanish.

To investigate the effect of prior L1 knowledge on the incidental acquisition of syntactic knowledge over time (RQs 1, 2, and 3), I ran a 2 × 2 × 2 mixed-design ANOVA on *d'* scores, with Time (Immediate, Delayed) as the within-subject factor and with Condition (Experimental, Control) and L1 background (English, Korean) as the between-subject factors. To confirm the possibility of the unaware group's implicit learning (RQ 4-1), I ran Mann–Whitney *U* tests to compare the respective *d'* scores of the unaware group and the control group. Additionally, I ran one-sample t-tests on the unaware groups' *d*' scores, with 0 (i.e., chance) as the test value. To investigate the effect of verbal awareness on the incidental acquisition of syntactic knowledge, I conducted a mixed-design ANOVA for each experimental group separately. For the EE group, I ran a 2 × 2 mixed-design ANOVA with Time (Immediate, Delayed) and Awareness (VP-aware, Unaware); for the KE group, I performed a 2 × 3 mixed-design ANOVA with Time (Immediate, Delayed) and Awareness (L1-aware, VP-aware, Unaware). To probe the nature of the acquired syntactic knowledge (RQ 4-2), I ran one-sample t-tests on mean accuracy (%) by source attributions, with 0.5 (i.e., chance) as the test value. Further, to investigate whether two L1 groups differed in terms of implicit and explicit knowledge development (RQ 4-2), I ran a 2 × 2 × 2 mixed-design ANOVA on mean accuracy with Attributions (Implicit, Explicit), L1 (English, Korean), and Time (Immediate, Delayed). To clarify the nature of any significant interactions, I followed up on the main model by using stepdown ANOVAs and comparisons of simple effects. An alpha level of 0.05 was used for all statistical tests.

2.3 Results

2.3.1 Comprehension test

The overall comprehension scores of the EE and KE groups were 18.17 (SD = 2.51) and 18.13 (SD = 3.17), respectively, which indicated that their comprehension level of the scrambled text was acceptable. The difference between the two L1 groups' comprehension scores was not significant [t(46) = 0.31, p = .759, d = 0.09]. The overall comprehension scores for the control groups, the EC and KC groups, were 23.13 (*SD* = 0.96) and 22.80 (*SD* = 1.05), respectively.

2.3.2 Overall GJT performance

Descriptive statistics of GJT accuracy scores (%) for the EE, EC, KE, and KC groups are presented in Table 2.4.

Table 2.4

Mean accuracy (%) of grammaticality judgments for English and Korean speakers

	Immediate GJT				Delayed GJT			
	n	*M* (%)	*SD*	95% *CI*	*n*	*M* (%)	*SD*	95% *CI*
				L1: English				
Experimental	24	56.93	9.80	[52.79, 61.07]	23	58.91	8.18	[55.37, 62.45]
Control	15	45.08	4.64	[42.51, 47.65]	14	45.54	4.59	[42.88, 48.19]
				L1: Korean				
Experimental	24	69.64	11.70	[64.70, 74.58]	22	77.61	11.81	[72.38, 82.85]
Control	15	49.04	9.44	[43.81, 54.27]	13	48.75	7.87	[43.99, 53.51]

Note: Accuracy scores for the EE, EC, KE, KC groups were normally distributed according to the Shapiro-Wilk test ($p > .05$)

The EE group's mean accuracy was moderate [Immediate: M = 56.93, SD = 9.80; Delayed: M =

58.91, SD = 8.18] but exceeded a baseline of 50% [Immediate: t(23) = 3.46, p = .002, d = 0.71; Delayed: t(22) = 5.22, p < .001, d = 1.09]. The KE group's mean accuracy was large [Immediate: M = 69.64, SD = 11.70; Delayed: M = 77.63, SD = 11.81] and was significantly greater than chance level [Immediate: t(23) = 8.22, p < .001, d = 1.68; Delayed: t(21) = 10.98, p < .001, d = 2.33]. The difference between the EE and the EC groups [Immediate: t(35.05) = 5.079, p < .001, d = 1.55; Delayed: t(34.862) = 6.365, p < .001, d = 2.02] and between the KE and the KC groups [Immediate: t(37) = 5.74, p < .001, d = 1.94; Delayed: t(33) = 7.82, p < .001, d = 2.88] was significant.

Descriptive statistics of *d*' scores are reported in Table 2.5, and the *d'* scores by group are depicted in Figure 2.2. The difference between the EE and the EC groups was significant [Immediate: $U = 50.00$, $p < .001$, $r = 0.60$; Delayed: $U = 28.50$, $p < .001$, $r = 0.68$], and the EE group's *d'* scores across the testing sessions [Immediate: $M = 0.37$, $SD = 0.74$; Delayed: $M = 0.54$, $SD = 0.62$] were greater than 0 (i.e. chance) [Immediate: $t(23) = 2.43$, $p = .023$, $d = 0.50$; Delayed: $t(22) = 4.15$, $p < .001$, $d = 0.87$). The difference between the KE and KC groups was also significant [immediate: $U = 32.50$, $p < .001$, $r = 0.68$; delayed: $U = 1.00$, $p < .001$, $r = 0.82$), and the KE (immediate: $M = 1.13$, $SD = 0.81$; delayed: $M = 1.80$, $SD = 1.10$) group's *d'* scores were also significantly above 0 [Immediate: $t(23) = 6.84$, $p < .001$, $d = 1.40$; Delayed: $t(21) = 7.70$, $p < .001$, $d = 1.64$).

Table 2.5

d' scores of the grammaticality judgments for English and Korean speakers

	Immediate GJT				Delayed GJT			
	n	*d'*			*n*	*d'*		
		M	*SD*	95% *CI*		*M*	*SD*	95% *CI*
				L1: English				
Experimental	24	0.37	0.74	[0.06, 0.68]	23	0.54	0.62	[0.27, 0.81]
Control	15	-0.65	0.60	[-0.98, -0.31]	14	-0.57	0.48	[-0.85, -0.30]
				L1: Korean				
Experimental	24	1.13	0.81	[0.79, 1.47]	22	1.80	1.10	[1.32, 2.29]
Control	15	-0.17	0.66	[-0.53, 0.19]	13	-0.07	0.42	[-0.33, 0.18]

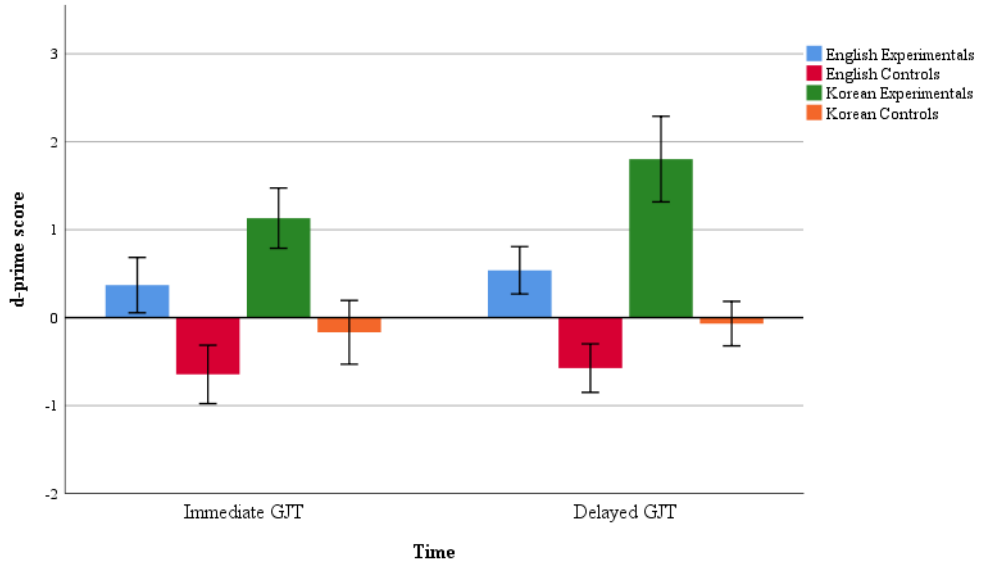

Figure 2.2

Mean d' scores for Korean and English speakers
Note: Error bars represent 95% confidence intervals.

To investigate whether *d'* scores on the GJT differed as a function of Condition (Experimental, Control), L1 (English/Korean), and Time (Immediate, Delayed), I ran a 2 (Condition) × 2 (L1) × 2 (Time) mixed-design ANOVA. According to the Shapiro–Wilk test, the distribution of *d'* scores was normal except for the EC group's and the KE group's delayed *d'* scores. As can be seen in Table 2.6, this analysis revealed significant main effects of Condition [$F(1,70) = 70.06$, $p < .001$, $\eta_p^2 = .51$], L1 [$F(1,70) = 27.36$, $p < .001$, $\eta_p^2 = .29$], and Time [$F(1,70) = 11.28$, $p = .001$, $\eta_p^2 = .14$]. A Time by L1 interaction was significant [$F(1,68) = 4.83$, $p = .031$, $\eta_p^2 = .07$], which was qualified by a borderline significant Time by L1 by Condition interaction [$F(1,68) = 3.84$, $p = .054$, $\eta_p^2 = .05$]. Figure 2.3 represents the three-way interaction visually.

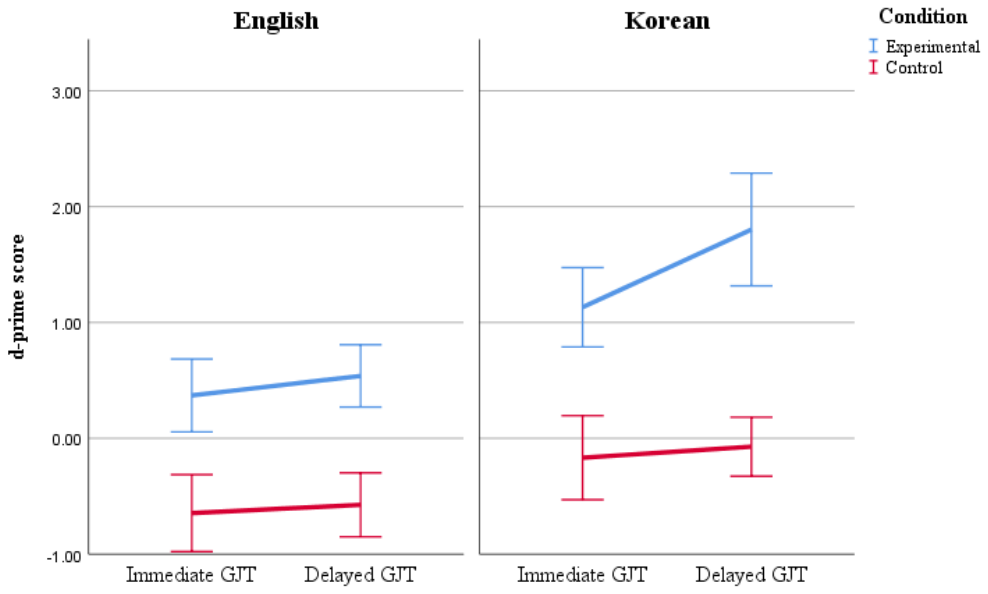

Figure 2.3

Three-way interactions between Condition, Time, and L1

Table 2.6

Mixed-design ANOVA for the d' scores on GJTs.

	SS	*df*	*MS*	*F*	*p*	η_p^2	power
Main effect (within-subject variable)							
Time	2.56	1	2.56	11.28	.001**	.14	.91
Main effect (between-subject variable)							
Condition	59.52	1	59.52	70.06	< .001***	.51	1.0
L1	23.24	1	23.24	27.36	< .001***	.29	1.0
2-way interaction							
Time x Condition	0.80	1	0.80	3.51	.065	.05	.46
Time x L1	1.10	1	1.10	4.83	.031*	.07	.58
Condition x L1	1.62	1	1.62	1.91	.172	.03	.28
3-way interaction							
Time x Condition x L1	.87	1	.87	3.84	.054+	.05	.49

+ $p < .06$; * $p < .05$; ** $p < .01$; *** $p < .0001$

To follow up on the three-way interaction, I carried out a 2 (Time) × 2 (Condition) mixed-design ANOVA for each L1 group separately. The analysis for the L1 English group showed that the main effect of Condition was significant [$F(1,35) = 40.69$, $p < .001$, $\eta_p^2 = .54$], but the main effect of Time [$F(1,35) = 0.79$, $p = .379$, $\eta_p^2 = .02$] and the Time by Condition interaction [$F(1,35) = 0.00$, $p = .986$, $\eta_p^2 = .00$] were not significant. This indicated that incidental exposure through a novel was effective for English speakers to develop syntactic knowledge, but the amount of knowledge remained steady over time.

The analysis for the Korean speakers revealed significant main effects of Time [$F(1,33) =$

14.75, p = .001, η_p^2 = .31] and Condition [F(1,33) = 33.21, p < .001, η_p^2 = .50] and, crucially, a significant Time by Condition interaction [F(1,33) = 6.95, p = .013, η_p^2 = .17]. To understand the nature of this interaction, I ran simple effects of Condition for each time point as well as simple effects of Time for each Condition. First, the effect of Condition was significant for both the immediate [F(1,33) = 20.95, p < .001, η_p^2 = .39] and the delayed tests [F(1,33) = 34.55, p < .001, η_p^2 = .51]. Second, there was no effect of Time for the KC group [F(1,33) = 0.06, p = .814, η_p^2 = .00], while there was a strong effect for the KE group [F(1,33) = 21.49, p < .001, η_p^2 = .39], which pointed to different amounts of improvement in the KE and KC groups' knowledge over time. These results indicated that incidental exposure through a novel was effective for the Korean speakers to develop syntactic knowledge, and, importantly, there was a significant increase only in the KE group's knowledge from the immediate to the delayed test.

To test the effect of L1 (RQ2), I ran additional 2 (L1) × 2 (Condition) factorial ANOVAs for each Time separately. Another two-way interaction was found only in the delayed test [F(1,68) = 2.45, p = .042, η_p^2 = .059], and here, I only report the results of simple effects of L1 for each Time. For the immediate test, the significant main effect of L1 [F(1,74) = 15.77, p < .001, η_p^2 = .18] and the non-significant L1 by Condition interaction [F(1,74) = 0.70, p > .05, η_p^2 = .01] suggested that the Korean speakers outperformed the English regardless of condition. In other words, even without being exposed to Koreanish sentences, the Korean speakers were better at judging the grammaticality of Koreanish sentences than the English speakers. However, in the delayed test, there was a significant effect of L1 in the experimental condition [F(1,68) = 31.39, p < .001, η_p^2 = .32], but in the control condition, there was no effect of L1 [F(1,68) = 2.97, p = .089, η_p^2 = .04]. This reflected the additional benefit that L1 Korean speakers derived from their native-language grammar for developing long-term syntactic knowledge of a new head-

final language.

These findings suggest that the training, i.e., incidental exposure through a novel, facilitated the development of new grammatical knowledge for both the English and the Korean speakers (RQ1). The Korean experimental group's knowledge was significantly enhanced two weeks later, as reflected in their increased scores, whereas the English experimental group's knowledge did not show evidence of enhancement (RQ3). Accordingly, the Korean experimental group outperformed the English experimental group at the delayed test (RQ2). This diverging pattern between the two L1 groups suggests that the Korean experimental group's prior L1 knowledge led to a clear advantage in the acquisition of Koreanish word order over time.

2.3.3. Awareness reflected in retrospective verbal reports and GJT performance

Analysis of retrospective verbal reports. The retrospective verbal report data from the debriefing interviews from Sessions 1 and 2 were analyzed to gauge participants' level of awareness. The retrospective verbal report data from the debriefing interviews were analyzed to gauge the participants' levels of awareness. To enhance the inter-rater reliability in the coding, two independent raters (the researcher and a trained rater) coded the verbal report data, which was gathered from two debriefing interview sessions. The raters identified 1) what type of awareness the participants developed (unaware, VP-aware, L1-aware), 2) when they became aware of the rules (exposure and testing), and 3) when they reported awareness (Session 1 and Session 2). The reliability coefficient (kappa) for the identification of awareness type was $\kappa = 0.94$, for the identification of the timing of awareness was $\kappa = 0.99$, and for the identification of the timing of reporting was $\kappa = 1.00$.

This analysis enabled me to identify subgroups of participants based on their reported awareness of the target rules. Note that only the experimental groups who read the Koreanish

version of the novel participated in the debriefing interview. From the verbal report data, I attempted to confirm the internal validity of the incidental learning condition (cf. Leow, 1997, Leow & Hama, 2013) by asking participants whether they tried to search for patterns while they were reading. This inspection led me to exclude two participants—one from each L1 group. Their answers, presented in Example (5), clearly showed that they intentionally searched for a pattern while reading, which means that their attention was oriented towards forms.

Example (5)

P119 (English participant)

"At first, I did not think there was a pattern. But after the first break, I started to notice a pattern, so I would try to figure it out and unscramble the sentences. I realized that the verb at the end of a dependent clause was the action of the first noun in the sentence."

P304 (Korean participant)

"I have noticed that the order of noun, verb, and etc. were messed up regularly. I did try to find the order of the words in the sentence but could not find a clear pattern."

An analysis of the English speakers' verbal report data (see Table 2.7) indicated that half of the participants—12 of 24 participants—became aware of the VP rule. This was a partial awareness of Koreanish word order rules, since participants did not mention any other rules regarding head finality. Based on their awareness of the VP rule, I categorized them as VP-aware. Among the 12 participants, 10 participants reported awareness of the VP rule at Session 1, during the short version of the interview, and two participants reported awareness at Session 2 (two weeks after Session 1), during the long version of the interview. All participants said they

became aware of the VP rule during the GJT. The other 12 participants did not report awareness of any target rules and were categorized as unaware.

Table 2.7

English speakers' verbal report data

	n	Timing of Awareness	Comment
Unaware	12	-	I thought the words were randomly jumbled.
VP-aware	12	Exposure phase	I noticed the subject or noun was normally at the beginning of the sentence, and the verb was at the end.

Unlike the analysis of the English speakers' data, an analysis of the Korean speakers' verbal report data (see Table 2.8) exhibited three types of awareness: L1 aware (n = 9), VP-aware (n =10), and unaware (n = 5).

Table 2.8

Korean speakers' verbal report data

	n	Timing of Awareness	Comment
Unaware	5	-	The rule I noticed is to place nouns at the beginning of the sentence.
VP-aware	10	Exposure phase	I recognized that one of the types is to place the verb in the end of the sentence.
L1-aware	9	Testing phase	I've noticed that the order of some sentences was similar to my native language, Korean, at test.

First, I found an interesting subgroup, the L1-aware group, who reported their awareness of the

relationship between the target language and their L1 (Korean). These nine participants reported that they became aware that the language used in the novel followed Korean word order while they were taking the GJT. It is important that they were unaware of any rules (n = 7) or only discovered the VP rule (n = 2) during the exposure phase; it was in the testing phase that they figured out that the Koreanish sentences in the novel matched their L1. In terms of the timing of reported awareness, among 9 L1-aware participants, 2 participants reported their L1 awareness after taking the immediate GJT at Session 1; the rest, 7 participants, reported their awareness after taking the delayed GJT at Session 2. Second, 10 participants became VP-aware. They reported noticing the VP rule but did not notice other rules regarding head finality in the Koreanish text. All VP-aware participants reported their VP-awareness at Session 1 after taking the immediate GJT. Lastly, five participants remained verbally unaware and were categorized as unaware.

Can learners acquire knowledge without awareness? Table 2.9 presents the descriptive statistics for GJT performance by participant awareness as reflected in the participants' retrospective verbal reports. To find evidence of implicit learning, the unaware groups' d-prime scores were analyzed using one-sample t-tests and Mann–Whitney's U test. Although the English unaware group (n = 12) outperformed the EC group [Immediate: U = 25.50, p = .002, d = 1.52; Delayed: U = 12.00, Z = -3.57, p < .001, d = 2.03], their performance on the immediate test was not significantly above chance [$t(11)$ = 1.19, p = .259, d = 0.34]. The Korean unaware group (n = 5), on the other hand, outperformed not only the KC group [Immediate: U = 5.50, Z = -2.79, p = .005, d = 1.60; Delayed: U = .00, Z = -2.94, p = .003, d = 2.04] but also the English unaware group [Immediate: U = 10.00, Z = -2.11, p = .035, d = 1.19; delayed: U = .00, Z = -2.87, p = .004, d = 2.21]. This finding indicated that the Korean unaware

group was better at learning without awareness than the English unaware group under incidental exposure.

Table 2.9

The EE and KE groups' GJT performance by awareness reflected in retrospective verbal reports

			Accuracy (%)			*d'*		
		n	*M*	*SD*	95% CI	*M*	*SD*	95% *CI*
L1: English								
Unaware	Immediate	12	54.27	7.32	[49.62, 58.92]	0.21	0.62	[-0.18, 0.61]
	Delayed	11	58.07	8.22	[52.54, 63.59]	0.43	0.44	[0.14, 0.72]
VP-aware	Immediate	12	59.58	11.49	[52.29, 66.88]	0.53	0.85	[-0.01, 1.06]
	Delayed	12	59.69	8.43	[54.33, 65.04]	0.64	0.76	[0.16, 1.12]
L1: Korean								
Unaware	Immediate	5	66.25	9.84	[54.03, 78.47]	1.02	0.66	[0.20, 1.84]
	Delayed	4	72.19	1.20	[70.28, 74.09]	1.20	0.08	[1.07, 1.33]
VP-aware	Immediate	10	65.63	11.12	[57.67, 73.58]	0.80	0.79	[0.24, 1.36]
	Delayed	9	70.44	7.44	[64.69, 76.14]	1.17	0.49	[0.79, 1.54]
L1-aware	Immediate	9	75.97	11.56	[67.09, 84.85]	1.56	0.79	[0.95, 2.17]
	Delayed	9	87.22	11.56	[78.34, 96.10]	2.70	1.17	[1.80, 3.60]

Further, considering that the VP-aware group reported awareness only regarding the VP rule, I tested their performance on other rules, for which they did not mention in their verbal reports. The relative clause sentence, one of the target structures which involved the NP and CP rules, was a suitable structure to test the acquisition of the CP and NP rules outside of verbal

awareness. The KE group [Immediate: *M* = 0.40, *SD* = 0.75, *95% CI* = 0.09, 0.72; Delayed: *M* =1.15 , *SD* = 1.93, *95% CI* = 0.29, 2.01] outperformed the KC group [Immediate: *M* = -0.24, *SD* = 0.92, *95% CI* = -0.75, 0.27] [Immediate: *U* = 112.50, *p* = .049, *r* = 0.31; Delayed: *U* = 55.00, *p* =.003, *r* = 0.51], which reflected their learning of complex structure for which they did not verbalize a rule. In contrast, there was no significant difference between the EE's [Immediate: *M* = 0.06, *SD* = 1.02, *95% CI* = -0.37, 0.49; Delayed: *M* = -0.29, *SD* = 0.75, *95% CI* = -0.61, 0.04] and EC's *d'* scores [Immediate: *M* = -0.09, *SD* = -0.79, *95% CI* = -0.53, 0.35; Delayed: *M* = -0.50, *SD* = 0.88, *95% CI* = -1.00, 0.01] [Immediate: *U* = 173.50 *p* = .849, *r* = 0.06; Delayed: *U* = 149.00, *p* = .701, *r* = 0.06]. Taken together, the Korean speakers showed stronger evidence of learning without awareness than the English speakers.

Relationship between awareness reflected in retrospective verbal reports, L1, and GJT performance. To answer Research Question 4-2, which investigates the relationship between awareness based on retrospective verbal reports and grammatical knowledge development under incidental learning conditions, I carried out a mixed-design ANOVA with Time as a within-subject variable and Awareness as a between-subject variable on the *d'* scores for each L1 group separately. The main effect and interactions with Awareness are of interest here. The Kolmogorov-Smirnov test showed that all subgroups' *d'* scores were normally distributed ($p > .05$).

For the EE group, a 2 × 2 mixed-design ANOVA with Awareness (VP-aware, Unaware) and Time (Immediate, Delayed) was carried out. The analysis showed that the main effect of Awareness [$F(1,21) = 0.68$, $p = .420$, $\eta_p^2 = .03$] and the main effect of Time [$F(1,21) = 0.44$, $p = .514$, $\eta_p^2 = .02$] were not significant. The Awareness by Time interaction was also not significant [$F(1,21) = 0.02$, $p = .895$, $\eta_p^2 = .00$]. This indicated that for English speakers, there

was no evidence showing that awareness affected grammatical knowledge development over time.

For the KE group, a 3 × 2 mixed-design ANOVA with Awareness (L1-aware, VF aware, Unaware) and Time (Immediate, Delayed) was performed. The analysis revealed significant main effects of Awareness [$F(2,19) = 6.95$, $p = .005$, $\eta_p^2 = .42$] and Time [$F(1,19) = 18.05$, $p < .001$, $\eta_p^2 = .49$]. The interaction between Time and Awareness was not significant [$F(2,19) = 2.83$, $p = .084$, $\eta_p^2 = .23$]. Post hoc comparisons using Tuckey's contrasts found a statistical difference between the L1-aware and the unaware groups (mean difference = 1.13, *95% CI* = 0.06, 2.21, $p = .038$) as well as the L1-aware and the VF-aware groups (mean difference = 1.14, *95% CI* = 0.30, 1.98, $p = .007$). This indicated that the L1-aware group excelled in the GJT across the testing sessions, which demonstrated that L1-awareness boosted test performance across time points.

To delve into the special role of conscious L1 transfer, I recategorized the Korean experimental group into the L1-aware and L1-unaware groups. The L1-unaware group combined participants from the VF-aware and the Unaware groups into a single category. A 2 × 2 mixed-design ANOVA with Time (Immediate, Delayed) and L1-awareness (L1-aware, L1-unaware) revealed significant main effects of Time [$F(1,20) = 18.99$, $p < .001$, $\eta_p^2 = .23$] and L1-awareness [$F(1,20) = 14.62$, $p = .001$, $\eta_p^2 = .42$] and a significant Time by L1-Awareness interaction [$F(1,20) = 5.93$, $p = .024$, $\eta_p^2 = .23$]. I then analyzed the two-way interaction (Figure 2.4) by examining the simple effects of Time for each awareness group. There was no effect of Time in the L1-unaware group [$F(1,20) = 3.21$, $p = .088$, $\eta_p^2 = .14$], while there was in the L1-aware group [$F(1,20) = 21.71$, $p < .001$, $\eta_p^2 = .52$]. The results indicated that the L1-aware group improved significantly in the two-week time period following the training intervention, which

suggested that L1-awareness led to improved knowledge consolidation.

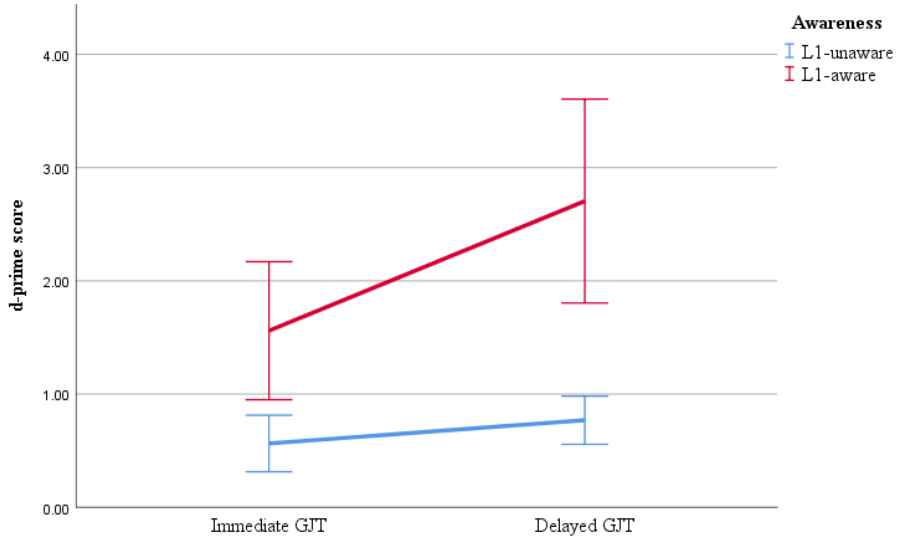

Figure 2.4

Two-way interaction between Time and L1-awareness for the KE group

2.3.4 Awareness reflected in source attributions and GJT performance

Is acquired knowledge implicit or explicit? The EE and KE groups' proportions and accuracy (%) across source attributions are presented in Table 2.10. For the analysis of source attributions, accuracy scores were used instead of *d'* scores because most of the participants did not have complete data for sixteen cells (source x grammaticality x correctness) to get the information (hits and false alarm rates) necessary to calculate *d'* scores. Therefore, mean accuracy scores by source attributions will be used in all subsequent analyses. Also, the untrained control groups' performance was not used as a baseline here because the controls could not attribute their decision on recollection or rule without being exposed to the target language.

I combined Guess and Intuition attributions to calculate the total proportion of responses

based on unconscious, implicit knowledge; I also added Recollection and Rule knowledge attributions to generate the total proportion of responses based on conscious, explicit knowledge. If a participant showed above chance performance even when they judged the grammaticality of a given sentence based on Guess or Intuition, this was considered evidence of unconscious, implicit knowledge (Dienes & Scott, 2005). Table 2.11 and Figure 2.5 present the descriptive information on the GJT accuracy by implicit and explicit attributions. The English speakers' implicit knowledge on the immediate test [M = 53.79, SD = 11.35] was not significantly above chance [$t(23) = 1.64$, p =.115, d = 0.33] but on the delayed test [M = 61.59, SD = 17.70] it was significantly above chance [$t(21) = 3.07$, p = .006, d = 0.66]. Unlike the English speakers, the Korean speakers' implicit source attributions [Immediate: M = 61.83, SD = 19.00; Delayed: M = 66.80, SD = 21.41] were significantly above chance for both the immediate [$t(23) = 3.05$, p = .006, d = 0.62] and delayed [$t(19) = 3.51$, p = .002, d = 0.78] tests. In case of the explicit knowledge, both the English and Korean speakers' accuracy was significantly above chance across the two testing sessions. Therefore, the results indicated that the Korean experimental group exhibited both implicit and explicit knowledge. Meanwhile, the English experimental group's implicit knowledge developed over time, as there was only evidence of implicit knowledge in the delayed test.

Table 2.10

The EE and KE groups' proportions and accuracy (%) across source attributions

		Proportion (%)		Accuracy (%)		
		M	*SD*	*M*	*SD*	*95% CI*
		L1: English				
Guess	Immediate	16.51	17.11	59.03	20.20	[49.58, 68.48]
	Delayed	24.67	21.62	55.96	22.61	[45.07, 66.86]
Intuition	Immediate	45.47	25.36	54.09	13.98	[48.18, 59.99]
	Delayed	35.92	23.78	61.17	16.77	[53.54, 68.81]
Recollection	Immediate	12.97	11.41	61.50	21.65	[51.42, 71.68]
	Delayed	13.26	20.93	73.56	18.92	[63.83, 83.29]
Rule knowledge	Immediate	25.05	27.50	59.90	24.18	[49.18, 70.62]
	Delayed	26.14	34.12	61.33	33.94	[44.45, 78.21]
		L1: Korean				
Guess	Immediate	23.19	27.21	58.38	28.80	[45.62, 71.15]
	Delayed	26.22	33.60	68.46	25.65	[54.26, 82.66]
Intuition	Immediate	28.39	24.86	60.74	24.34	[49.95, 71.53]
	Delayed	32.35	33.57	66.77	25.24	[54.22, 79.32]
Recollection	Immediate	25.50	25.31	77.46	16.41	[69.98, 84.94]
	Delayed	14.44	23.73	82.20	20.11	[70.60, 93.81]
Rule knowledge	Immediate	22.92	27.90	85.03	13.96	[77.30, 92.76]
	Delayed	26.99	38.57	84.19	11.82	[76.25, 92.13]

Table 2.11

The EE and KE groups' accuracy (%) for implicit and explicit attributions

	Implicit attributions (based on Guess or Intuition)				Explicit attributions (based on Recollection or Rule)			
	n	*M*	*SD*	*95% CI*	*n*	*M*	*SD*	*95% CI*
				L1: English				
Immediate	24	53.79	11.35	[49.00, 58.58]	24	61.83	14.22	[55.83, 67.84]
Delayed	22	61.59	17.70	[53.75, 69.44]	22	67.77	18.99	[59.35, 76.19]
				L1: Korean				
Immediate	24	61.83	19.00	[53.81, 69.86]	22	79.59	13.32	[73.68, 85.50]
Delayed	20	66.80	21.41	[56.78, 76.82]	18	83.78	12.12	[77.15, 89.80]

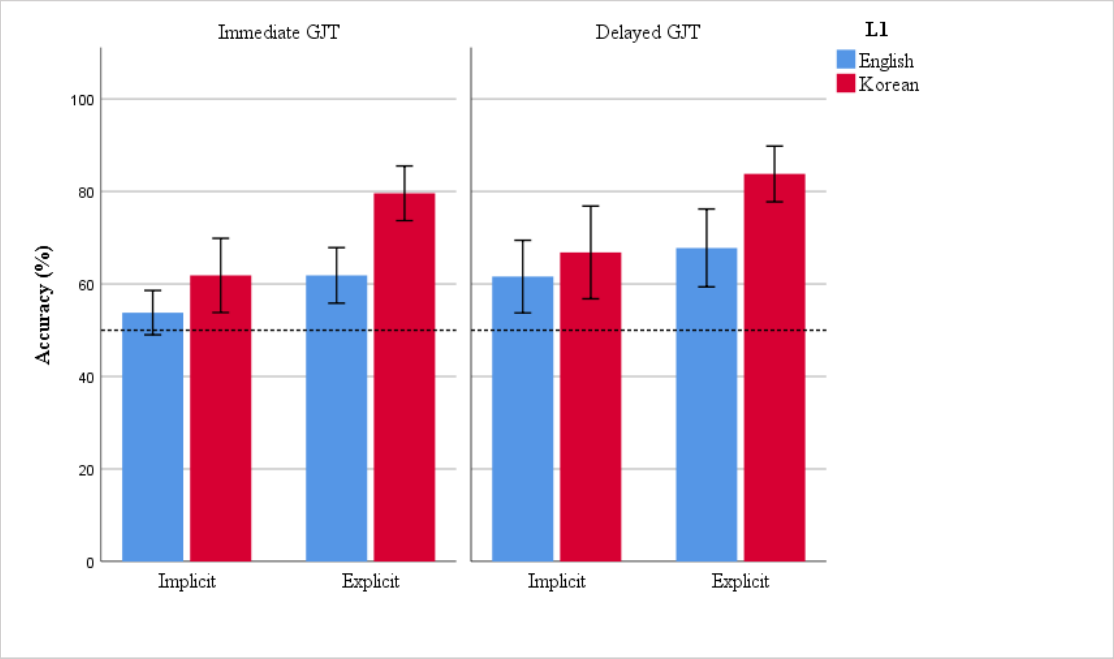

Figure 2.4

The EE and KE groups' accuracy (%) for implicit and explicit attributions

Relationship between awareness reflected in source attributions, L1, and GJT performance. To answer Research Question 4-4, which investigates whether the English and the Korean experimental groups differ in their development of implicit and explicit knowledge, I conducted a 2 x 2 X 2 mixed-design ANOVA on GJT mean accuracy with Attributions (implicit, explicit) and Time (immediate, delayed) as a within-subject variable and L1 (English, Korean) as a between-subject variable. According to the Shapiro-Wilk test, the distribution of accuracy scores by implicit and explicit attributions was normal except for the English and the Korean experimental groups' implicit attributions at the delayed test.

The main effects of Attributions [$F(1,35) = 32.68$, $p < .001$, $\eta_p^2 = .48$], L1 [$F(1,35) = 4.97$, $p = .032$, $\eta_p^2 = .12$], and Time [$F(1,35) = 5.73$, $p = .022$, $\eta_p^2 = .14$] were significant, as were the Attributions by L1 interaction [$F(1,35) = 4.70$, $p = .037$, $\eta_p^2 = .14$]. This Attributions by L1 interaction is illustrated in Figure 2.5.

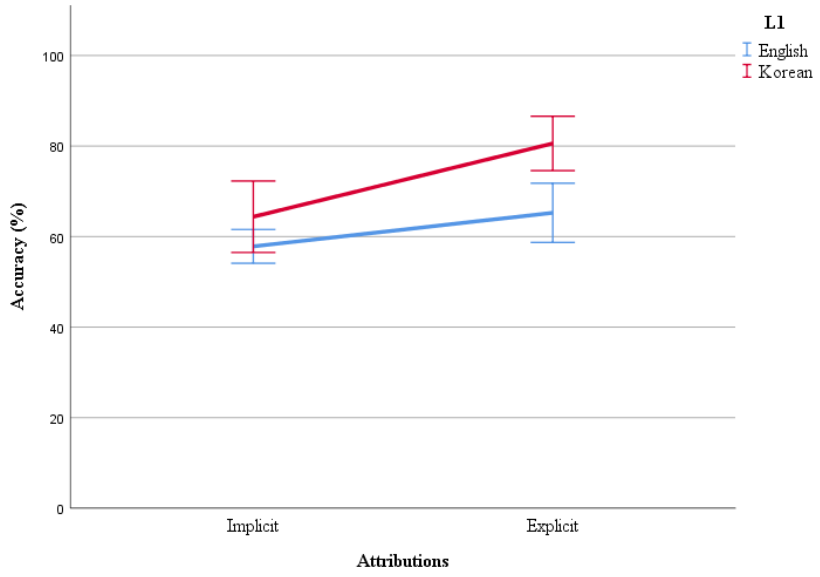

Figure 2.5

Two-way interaction between Attributions and L1

To follow up on the two-way interaction, the simple effects of L1 were analyzed for each attribution separately. The Korean experimental group outperformed the English experimental group when they relied on explicit sources [$F(1,35) = 8.59$, $p = .006$, $\eta_p^2 = .20$], but there was no difference between the two L1 groups' scores when they relied on implicit sources [$F(1,35) = 0.78$, $p = .383$, $\eta_p^2 = .02$].

The results thus showed that the Korean experimental group developed more explicit knowledge than the English experimental group under incidental exposure, whereas there was no difference between the two L1 groups in the development of implicit knowledge. Therefore, this suggested that the Korean experimental group's prior L1 knowledge had a facilitative effect on the development of explicit knowledge under incidental exposure.

2.3.5 Summary of results

- Both the English and Korean experimental groups reliably learned the target grammar after incidental exposure through a novel.
- After two weeks, the Korean experimental group's knowledge was significantly enhanced, which prompted them to outperform the English experimental group more at the delayed GJT.
- This enhanced knowledge consolidation was led by the Korean L1-aware group, who noticed at test (not during reading) that the target grammar matched their L1 grammar, Korean.
- The Korean experimental group provided stronger evidence of implicit learning and implicit knowledge than the English experimental group.
- The English experimental group's implicit learning and implicit knowledge was comparatively limited.
- The Korean experimental group also developed more explicit knowledge than the English

experimental group.

2.4 Discussion

Multiple lines of evidence in the present study demonstrated the impact of prior L1 knowledge on the implicit and explicit learning of non-native syntax under incidental learning conditions. Although the English and Korean experimental groups consistently outperformed the control groups, the differential patterns of the two groups' GJT performance indicated that cross-language similarity facilitated the development of implicit and explicit knowledge. First, the Korean experimental group showed strong evidence of implicit knowledge over time, while the English experimental group did not provide reliable evidence of implicit knowledge. Second, the Korean experimental group acquired more explicit knowledge than the English experimental group, largely through their awareness of the connection between their L1 and the target language. These findings extend previous literature in important ways, being the first to reveal a dynamic between L1 prior knowledge, awareness, and time in incidental syntax learning.

The present investigation provided not only the supporting evidence of the possibility of learning L2 grammar without awareness (Godfroid, 2016; Kerz, Wiechman, & Riedel, 2017; Robinson, 1995; Rogers et al., 2016; Williams, 2011; Williams & Kuribara, 2008; Williams & Rebuschat, 2012; Tagarelli et al., 2016), but also the compelling evidence for syntactic transfer of L1 word order knowledge in incidental learning conditions.

Interestingly, L1 transfer occurred at both conscious and unconscious levels. First, the Korean experimental group's implicit knowledge development was indicative of unconscious L1 transfer. Two awareness measures—retrospective verbal reports and source attributions—suggested that while the English experimental group only showed evidence of implicit

knowledge at the delayed test, the Korean experimental group reliably exhibited a considerable amount of implicit knowledge across both testing sessions. A follow-up investigation of performance on the relative clause, a complex structure that combined two rules of which most participants remained unaware (i.e., NP and CP), showed that only the Korean speakers judged these structures with above-chance accuracy, whereas the English speakers did not. In other words, although no Korean speaker mentioned rules or patterns regarding relative clauses, they still learned them. These findings point to the Korean speakers' performance advantage over the English speakers for implicit learning of head-final syntax, most likely due to unconscious transfer of L1 word order rules. The Korean speakers may have unconsciously relied on their deeply entrenched L1 word order, boosting their GJT accuracy on responses based on guess and intuition. One possible account of the Korean speakers' results is that domain-general learning mechanisms (at play in the current learning task) interacted with their domain-specific L1 knowledge (cf. Leung & Williams, 2014) and produced superior learning outcomes as a result.

Second, another noteworthy finding of this study concerns the conscious transfer of L1 knowledge. We can reasonably assume that the L1-aware group, those who noticed the connection between the target language and their L1, used their unconscious L1 knowledge strategically in the present task. This conscious L1 transfer provided the L1-aware participants a critical advantage over their L1-unaware peers. It also led them to produce significant gains in GJT scores from the immediate to the delayed test. On the debriefing interview, L1-aware participants reported experiences such as "I rewrote the sentence in Korean in my mind" or "I read the given test sentences in the way I read Korean sentences." This reliance on L1 knowledge to test hypotheses and formulate rules may have contributed to the Korean speakers' developing more explicit knowledge than the English speakers.

L1 awareness appears to be very unique in its nature. They developed conscious judgmental knowledge based on their unconscious structural knowledge (Dienes & Scott, 2005). L1 awareness involves the deliberate activation of a full set of implicit L1 knowledge, but it certainly cannot be equated to the development of L2 implicit knowledge. Rather, it is the activation of a given body of knowledge for strategic use. More importantly, L1 awareness also cannot be viewed as the development of explicit knowledge; that is, although the participants were aware of the importance of their L1, they could not verbalize the specific rules in the L1 and the target language. One possibility is that the participants did not have the metalinguistic terminology needed to describe the regularities in the target language (none of the participants were linguistic majors), particularly regarding the rules for complex structures. Hence, some caution is needed when interpreting the non-verbalization as a lack of explicit linguistic knowledge. One could imagine a situation in which a tennis player tries playing badminton. Although he or she has not swung a badminton racket before, he will be able to hit the shuttlecock quickly when his body unconsciously remembers the body movement used in tennis. Once he figures out the similarities between tennis and badminton, he would be able to master badminton with the help of his experience in tennis, even if he has never taken a single tennis lesson in his life. In this way, it is probable that the Korean speakers did better than the English speakers on the GJT with the help of their deeply entrenched Korean grammar.

Traditionally, SLA researchers have examined two levels of awareness by means of debriefing interviews (Schmidt, 1995, 2001): noticing (i.e., a low level of awareness that involves a conscious registration of stimuli) and understanding (i.e., a high level of awareness that requires the recognition of a rule). Such a dichotomy may be too simple to capture the multifaceted cognitive layers of awareness in some cases. Keeping this in mind, some

researchers recently attempted to establish more detailed categories of awareness through a thorough debriefing interview (Curcic, Andringa, & Kuiken, 2019; Godfroid, 2016; Leow, 2015a, 2015b; Rebuschat et al., 2015). For example, Godfroid (2016) observed that a "verbwise" subgroup (i.e., participants who noticed the important role of verbs) showed mixed characteristics of aware and unaware learners. Using a visual-world paradigm, Curcic et al. (2019) found that only a "prediction aware" subgroup (i.e., participants who knew that determiners helped them comprehend at test) showed successful L2 predictive processing. It should be highlighted that these types of awareness commonly enabled learners to engage in a strategic and deliberate behavior that contributed to their excellent test performance, similarly to the L1-awareness observed in this study. However, unlike Godfroid (2016) and Curcic et al. (2019), in this study I used an untimed, written GJT as assessment and L1-awareness emerged only during the GJT.

The question for the Korean speakers, then, is how was the representation of Koreanish established? The shared syntax model may provide an explanation for this representational question (e.g., Hartsuiker, Pickering, & Veltkamp, 2004; Bernolet, Hartsuiker, & Pickering, 2013; Hartsuiker & Bernolet, 2017). Using evidence of crosslinguistic priming, this model proposed that bilinguals have shared syntax representations, in which all L2 structures similar to L1 merge with the corresponding L1 node and thus establish one final state of abstract representations. Based on this model, the extensive amount of Koreanish input may have resulted in shared representations of Korean and Koreanish in the Korean speakers, given that the word orders of the two languages were identical. The Korean speakers could have accessed their existing L1 syntactic nodes, combined the Koreanish structures with the L1 nodes, and then connected the English vocabulary to the shared syntax.

From a methodological perspective, this study presented two advantages: the use of a naturalistic task—reading a novel—during training and the use of a delayed test. First, the training effect of the naturalistic task—reading a novel—showed an interaction with prior L1 knowledge. The amount of input (497 sentences) achieved through novel reading during the training was almost four times greater than that achieved in previous studies [Grey et al. (2014): 128 sentences; Kim & Godfroid, in press: 120 sentences; Miller & Godfroid, (2019): 120 sentences; Rogers et al. (2016): 144 sentences; Tagarelli et al. (2016): 120 sentences; Williams & Rebuschat (2012): 120 sentences]. Moreover, reading an authentic novel might have simulated more natural reading processes than reading isolated sentences or unconnected, short texts. However, depending on the L1 background, such abundant exposure involved in a naturalistic task resulted in differential learning rates. On one hand, the English experimental group showed similar learning rates (immediate: 56.93%; delayed: 58.91%) as those in previous studies that utilized shorter and more controlled exposure tasks [e.g., Grey et al. (2014): immediate 57.5%, delayed 59.8%; Kim & Godfroid (in press): 55.6 %; Miller & Godfroid (2019): 53.4 %; Rogers et al. (2016): 55.44%; Tagarelli et al. (2016): 55.53%; Williams & Rebuschat (2012): 54.6%]. On the other hand, the learning effects were stronger (immediate: 69.64%; delayed: 77.61%) in the Korean experimental group than in the study groups in previous studies. This discrepancy indicated that the four-time input in fact did not produce superior learning gains when participants were not aided by prior L1 knowledge. That is, the extensive input under incidental exposure was likely to be beneficial only when a positive L1 transfer occurred. As to why more input did not lead to greater learning without the help of L1 knowledge remains a question. Plausibly, the naturalistic input from the novel was cognitively too demanding for beginner learners because of the complexity of the unmodified sentences. In terms of maintaining the

ecological validity of the study, I had no control over the sentences from the original material (i.e. the novel); for example, the type and token frequency of the target structures, which were shown to mediate the effectiveness of incidental learning (Denhovska, Serratrice, & Payne, 2016), were not controlled. There might have been too much syntactic complexities in the input (e.g., too many rules and examples), and such input was introduced at once without any instructions, which might have caused cognitive burden for beginner learners who could not rely on prior L1 knowledge.

Second, the use of delayed tests revealed an intriguing pattern of knowledge development over time. Consistent with Grey et al. (2014) and Morgan-Short et al. (2012), incidental learning of syntax was durable regardless of L1 background. Moreover, the Korean speakers showed gains in GJT scores over time without additional exposure, and these gained stemmed primarily from the L1-aware group. This finding could be accounted by the benefits of sleep in the enhancement of implicit and explicit memories (Batterink, Oudiette, Reber, & Paller, 2014; Fischer, Drosopolous, Tsen, & Born, 2006; Plihal & Born, 1997; Wagner, Gais, Haider, Verleger, & Born, 2004). Psychological evidence suggests that sleep has facilitative effects on the conversion of implicitly acquired information into explicitly available knowledge, possibly through an interaction between implicit and explicit memory processes. For example, Batternink et al. (2014) provided neurolinguistic evidence showing that sleep contributes to the stabilization of new linguistic rule knowledge due to the reactivation of linguistic regularities during sleep. Likewise, in the present study, during the two-week span between the immediate and delayed tests, the participants had multiple nights of sleep, which presumably benefitted their memory consolidation of newly acquired syntactic knowledge. Given that the majority of Korean L1-aware participants (seven out of nine) became aware at delayed testing, there is a possibility that

the Korean speakers' reorganization of their implicit structural knowledge of Koreanish during sleep could have given rise to L1 awareness, resulting in their improved performance at delayed testing. Had it not been for the delayed tests, such interaction between time, L1, and awareness could not have been demonstrated. To obtain a more complete picture of the time-course of implicit and explicit learning, more research is needed that uses delayed testing, with varying degrees of delay built into the research design.

CHAPTER 3

ONLINE MEASURE: READING TIMES FROM EYE-TRACKING

3.1 Background

3.1.1 Theoretical background on L2 developmental processes

A wide array of SLA theories share a goal of explaining how L2 develops. One way to observe the acquisitional processes in L2 development is to draw on a cognitive psychological perspective, particularly through the lens of domain-general learning mechanisms. From this cognitive perspective, "language is cut of the same cloth as other cognitive processes" (Ellis, 1998, p. 637), such as reasoning, motor activity, and visual perception. This approach has enriched the field by making it possible to explore empirically whether L2 development is comparable to other cognitive development as well as the ways in which it is unique (Ellis, 2006; DeKeyser, 2007; MacWhinney, 1997; Ullman, 2005). In this section, I will briefly introduce four cognitive-psychological frameworks that provide insights into the present study, with regard to L2 knowledge development and representation: the two-dimensional model of language proficiency (Bialystok, 1994), the representational redescription model (Karmiloff-Smith, 1992), the radical plasticity thesis (Cleeremans, 2007, 2011), and skill acquisition theory (Anderson, 1993; DeKeyser, 2003, 2015).

Bialystok (1994) put forward a two-dimensional framework for explaining cognitive aspects of L2 proficiency development. Specifically, she posited analysis and control as two processing constructs that jointly lead to an increase in L2 proficiency. First, analysis is the process "by which mental representations that were loosely organized around meanings (knowledge of the world) become rearranged into explicit representations" (Bialystok, 1994, p. 159). Through analysis, implicit knowledge becomes explicit knowledge, and this increases the accessibility to knowledge. Second, control is the real-time "processing choice about where

attention should be best spent" (p.160). A high level of control, an ability to selectively allocate attentional resources, is necessary for learners to develop fluency or automaticity (also see Segalowitz, 2010). According to Bialystok's framework, when these two processing constructs are applied, mental representations undergo a qualitative change, creating more analyzed representations that can be attended more efficiently and selectively. However, these two abilities might not always go hand in hand. It is possible some learners may have a higher level of analysis but a lower level of control.

Karmiloff-Smith's representational redescription model (1992) pointed out the importance of representational changes for a child's cognitive development. In this model, implicit knowledge can be gradually redescribed into explicit knowledge, going through four stages of representational redescription. The first stage is the Implicit level, which involves data-driven learning and procedural representations. Next, at Explicit Level 1, knowledge is abstracted from the procedural representation. This may yield inflexible behavior because the child makes error corrections to adapt his or her knowledge to a new code. Knowledge is still unconscious at this point. The third stage is Explicit Level 2, in which the conscious manipulation of knowledge is possible but verbalization of knowledge is not available yet. Due to increased understanding, performance may show improvement during this period. Finally, at Explicit Level 3, a full set of explicit knowledge is formed, which appears to be conscious, verbalizable, and flexible. According to Karmiloff-Smith's model, proceeding from implicit to explicit knowledge is likely to show a U-shaped pattern, because learners make errors and error corrections at Explicit Level 1 and subsequently show improvement at Explicit Level 2.

The radical plasticity thesis (Cleeremans, 2007, 2011) regards awareness as evidence of learning. Drawing on a connectionist approach, Cleeremans accounted for how the brain

formulates awareness through learning. The crux of his thesis is that awareness is the outcome of the brain's unconscious learning from the external environment and its interactions with internal representational system. By processing information over long time scales, learners accumulate implicit knowledge, which results in a weak knowledge representation. Then, as learning progresses, learners' implicit cognition will gradually gain access to awareness (see also, Dienes and Perner, 1999), resulting in a stronger mental representation with explicit knowledge. Awareness plays an essential role in learning because it gives adaptive, flexible control over behavior. The involvement of awareness will decrease until the learner reaches a very high-quality of representation; this is the point where automatic behavior emerges. Cleeremans pointed out that automatic behavior may not be unconscious (but for a different account, see also Williams, 2009); rather, awareness is optional in automatic behavior, because automatic behavior is "so adapted that it can unfold without the need for conscious monitoring." (2011, p. 6).

Skill acquisition theory (Anderson, 1982, 1993) draws on the power law of practice as a key property of human learning mechanisms, which is ubiquitously applicable to a wide array of cognitive (e.g., reading) and psychomotor skills (e.g., telegraphy). When humans practice a certain skill over time, the skill will show "development from initial representation of knowledge through initial changes in behavior to eventual fluent, spontaneous, largely effortless, and highly skilled behavior" (DeKeyser, 2015, p. 94), across so many linguistic and non-linguistic domains. In 1981, Newell and Rosenbloom presented a seminal study on automaticity, in which they maintained the power function to be the central law of human learning. Because practice resulted in a specific shape of learning curve, namely the power function (i.e.,*speci*), across diverse tasks, this mathematical function became referred to as the power law of practice. In the mathematical formula, *Y* represents response times (RTs) or error rate; *X* represents the amount of practice; the

exponent *n* represents the learning rate at which performance improves with practice. Importantly, the power learning curve of RT and error rate involves an initial quick and drastic decrease, which can be interpreted as a shift from "knowledge that" (*declarative knowledge*) to "knowledge how" (*procedural knowledge*). This learning period is then followed by a slow and gradual plateau, which is an index of the automatization of procedural knowledge (Anderson, 1982, 1993). According to this view, for various human skills, practice and the power law of practice entail different stages that lead a qualitative change over time. The speed-up observed in the power learning curve can be considered as the adaptation to the statistical structure of the input (Anderson & Schooler, 1991).

Building on Anderson's skill acquisition theory (1982, 1993), DeKeyser (1997) conducted the first study to empirically investigate whether L2 skill development followed the power law of practice and whether a skill can be transferred to other skills, from comprehension to production and vice versa. In this longitudinal study, the participants explicitly learned grammar rules and vocabulary items of an artificial language before they engaged in computer-administered comprehension (sentence picture matching) and production tasks (picture description) over 8 weeks. The results confirmed the prediction of skill acquisition theory, by showing that both longitudinal RT data and error rates followed the power function. The learning curves for both L2 comprehension and production revealed fast proceduralization followed by slow automatization. This finding, therefore, provided evidence showing that L2 skills, like other cognitive and psychomotor skills, can be automatized with practice, undergoing two important stages—proceduralization and automatization.

The review of these theoretical explanations of domain-general principles of learning provides important lessons to understand the development of L2 knowledge. First, repeated

exposure and extensive experience bring about a qualitative change in mental representations. Second, the accumulation of implicit cognition may be necessary for further knowledge development. Third, the involvement of awareness is likely to aid knowledge development, by providing control over behavior. Lastly but most importantly, knowledge development can be represented in a speedup in behavior, which can be highly informative of how language users can automatically and fluently access their mental representations.

As Bialystok (1994) rightly pointed out, essentially, L2 acquisition can be different from other developmental processes, including L1 acquisition. This is because adult L2 learners need to build a new representation on a fully elaborated L1 linguistic system. Of note, adults already possess "a more highly analyzed conception of language and more well-developed procedure for directing attention" (Bialystok, 1994, p. 163) than children learning their L1 do. Therefore, it would be an interesting exploration to see how prior L1 knowledge, which may or may not be domain-general, affects L2 development, which is hypothesized to draw on domain-general learning mechanisms, such as the power learning curve. In light of this, in the present study I aim to explore empirically how L1 grammatical knowledge affects the change in real-time processing of syntactic structures in incidental learning conditions.

3.1.2 Incidental exposure and changes in reading times

Recently, eye-tracking methodology emerged as a powerful tool to explore cognitive processes that underlie L2 development in diverse learning contexts (e.g., vocabulary: Godfroid, Ahn, et al., 2018; Elgort et al., 2018; grammar: Indrarathne, Ratajczak, & Kormos, 2018). Eye-movement recording can offer detailed information about how the learning process unfolds over time (Godfroid, 2020; Godfroid & Winke, 2015; Leow, Grey, Marijuan, & Moorman, 2014) because it gives a spatial and temporal reflection of attentional processes in real-time. That is, as

an online processing measure, eye movement records can provide a window into peripheral, selective, and focal attention (Godfroid, 2019). One of the benefits of eye-movement data is that it offers research participants a fairly natural reading experience (but see Spinner, Gass, & Behney, 2013), as an eye-tracking experiment usually does not require a secondary task (e.g., pressing a button) that may interfere with normal reading processes (Dussias, 2010; Godfroid & Spino, 2016). Thus, for reading long texts, such as novels, eye-tracking may well be the preferred methodology to observe moment-by-moment cognitive processing.

Combining natural reading materials with an incidental learning experiment, Godfroid, Ahn, et al. (2018) used eye-tracking methodology to investigate changes in reading times for unfamiliar vocabulary that occurred repeatedly in the novel and how the reading times related to vocabulary learning. The participants read five chapters of an authentic English novel, in which target Farsi-Dari words were naturally embedded, while their eye-movements were recorded. Using growth-curve modelling, the authors found that with repeated exposure, processing time (i.e., total reading time) on the target words decreased in a non-linear fashion, generating an S-shaped pattern. The authors proposed that the speed-up in lexical processing over time could reflect "implicit learning processes or the gradual build-up and specification of a new word representation that can be accessed increasingly fluently" (p. 574).

Using an eye-tracking methodology, Indarathne et al. (2018) focused on the cognitive processing during implicit and explicit learning of a syntactic construction. Similar to Godfroid, Ahn, et al. (2018), the authors also found an S-shaped curve in their eye-movement data. The initial decrease in total fixation duration in the early stage of learning was considered a reflection of habituation; that is, the reader's increased familiarity with the new grammar construction. On the other hand, the sharp decrease in total fixation duration during the last stage of learning,

which was shown only in the explicit instruction conditions, was taken as evidence of increased processing proficiency because it correlated with learners' grammar learning gains on an unannounced post-test.

Taken together, a decrease in eye fixation times (i.e., speed-up) can be used as evidence of a change in mental representation, such as proceduralization and automatization. Eye-tracking data in these studies provided a window into cognitive changes during learning processes. To obtain a multi-faceted understanding of real-time L2 processing and learning, researchers could further triangulate eye-movements with other measures of knowledge and awareness (Godfroid & Winke, 2015; Rebuschat et al., 2013).

3.1.3. Incidental exposure and sensitivity to L2 violations

Recently, many SLA researchers have begun to pay special attention to online sensitivity to grammatical violations as a reflection of integrated L2 knowledge (Godfroid, 2016; Granena, 2013, Jiang, 2007, 2012; Leung & Williams, 2011, 2012, 2014; Sanz & Grey, 2015; Spinner & Foote & Upor, 2017; Spinner & Jung, 2018; Suzuki, 2017; Suzuki & DeKeyser, 2015, 2017; Vafaee et al., 2016). Integration of knowledge into the linguistic system is likely to enable automatic competence, which involves spontaneous language use in both receptive and productive tasks (Jiang, 2007). If L2 knowledge is integrated, learners can engage in automatic and fluent processing without attending to grammatical accuracy. A typical example of integrated knowledge would be the L1, in the sense that L1 can be automatically retrieved and fluently produced without attending to grammatical accuracy. Researchers and practitioners ultimately aim for learners to achieve this type of integrated knowledge because it is automatic in its activation and functioning and underlies fluent language use (Doughty, 2003).

Grammatical sensitivity can be indexed by an increased processing time in response to

L2 violations (e.g., Godfroid, 2020). which can be assessed by psycholinguistic techniques, such as RTs (e.g., Leung & Williams, 2011), eye movements (e.g., Keating, 2009), and event-related brain potentials (ERP) (e.g., Tokowicz & MacWhinney, 2005). Researchers can compute sensitivity by subtracting processing times for grammatical sentences from those for matched, ungrammatical sentences; the difference reflects the degree of sensitivity for each participant. The core assumption of this methodology is that someone with integrated knowledge will unconsciously slow down to a grammatical violation in the sentence while processing. However, if a participant has incomplete or little implicit knowledge, then his or her processing time for violations and non-violation sentences would not be significantly different.

Adopting a RT methodology, Leung and Williams (2011) bolstered Williams's (2005) findings of implicit learning of form-meaning mappings by introducing a real-time component. They found online sensitivity—increased RTs in the violation block—among participants who remained verbally unaware. Using a pretest-treatment-posttest design, Godfroid (2016) also found unaware participants' sensitivity to violation, not only during the treatment (a picture matching task) but also at the posttest (a word monitoring task) that followed the intervention. In these studies, an unconscious slowdown to grammatical errors was interpreted as the automatic activation of implicit knowledge (Godfroid, 2016; Leung & Williams, 2011).

Interestingly, online sensitivity appears to be influenced by learners' prior L1 knowledge (Foucart & Frenck-Mestre, 2011; Jiang, 2007, 2011; Keating, 2009; Lim & Christianson, 2015; Sagarra & Ellis, 2013; Tokowicz & MacWhinney, 2005; Tolentino & Tokowicz, 2011, 2014). To illustrate, Jiang (2007) used RTs from a self-paced reading task to reveal how online sensitivity, which was hypothesized to reflect integrated L2 knowledge, can be selectively activated depending on the target structure. Native speakers of English and Chinese English-as-

a-second-language (ESL) speakers read grammatical and ungrammatical English sentences containing two target structures: verb categorization and plural -s. Through a comparison of reading times between English and Chinese speakers, he found out that L2 knowledge integration was dependent on the type of structure. While the English speakers showed sensitivity to both target forms, the Chinses speakers were sensitive only to verb categorization; they did not show a delay when reading sentences containing errors (omissions of) plural -s. One possible explanation for this phenomenon was the transfer of L1 knowledge because in Chinese, plural marking is rare while the verb categorization system is similar.

An ERP study by Tokowicz and MacWhinney (2005) documented the effect of L1-L2 similarity on online grammatical sensitivity. English learners of Spanish performed a GJT while their brain responses were recorded. Importantly, there were three types of target forms: a) tense marking (L1-L2 similar form), b) determiner number agreement (L1-L2 dissimilar form), c) determiner gender agreement (L2 unique form). Although GJT accuracy was only near chance-level, the researchers found online sensitivity, evidenced by a difference in brain responses to ungrammatical and grammatical sentences, for tense marking and determiner-gender agreement, but not for determiner-number agreement. In other words, participants showed differential sensitivity to syntactic anomalies depending on the L1-L2 similarity. Specifically, they showed a P600 effect[3] when the L2 form was similar to the L1 and when it was unique to the L2, but not when it was dissimilar. Based on these findings, the authors emphasized the importance of cross-language similarity in implicit processing. However, it remains unknown how awareness played a role during online processing because no awareness measure was employed in their study. An open question, therefore, is whether participants were aware of the L1-L2 relationships to which

[3] P 600 is a brain wave form that takes place to syntactic anomalies.

they showed sensitivity.

3.1.4 Research questions

In this chapter, I will present the answers to the following research questions:

RQ1: Learning curve

1-1. (prior L1 knowledge and learning curve)

Do the English and the Korean speakers' residualized sentence reading times for Koreanish sentences change according to the power law of practice?

1-2. (awareness and learning curve)

Do the aware and the unaware groups differ in terms of changes in residualized sentence reading times over time?

RQ2: Sensitivity

2-1. (prior L1 knowledge and sensitivity)

Do the English and Korean speakers show online sensitivity by slowing down their reading times for sentences with word order violations?

2-2. (awareness and sensitivity)

Do the aware and unaware groups differ in terms of online sensitivity to word order violations?

3.2 Methods

3.2.1 Participants

The same group of English speakers and Korean speakers who participated in the study reported in Chapter 2 also participated in this study. English speakers (n = 40, M_{age} = 23.71, SD = 3.90) and Korean speakers (n = 40, M_{age} = 25.25, SD = 4.42) were recruited from a large, Midwestern university in the United States and two universities in South Korea. Those invited to

participate held at least a bachelor's degree or were enrolled in a degree-granting university program at the time of the experiment.

Each group was subdivided into two: an experimental and a control group. This resulted in four subgroups: (a) an English experimental group (EE, n = 25), (b) an English control group (EC, n = 15), (c) a Korean experimental group (KE, n = 25), and (d) a Korean control group (KC, n = 15). The experimental groups, EE and KE, read the novel *The Mysterious Affair at Styles* (Christie, 1920) in a semi-artificial language that consisted of English vocabulary and Korean word order. The control groups, EC and KC, read the English version of the novel; that is, the same novel, but with English vocabulary and English word order.

The English speakers did not have any knowledge of Korean or any other head-final languages, such as Japanese or Turkish. Further, the experimental groups had no experience in German, the language used to create ungrammatical sentences in the violation blocks during the testing phase (see section 3-2-3). The Korean speakers were Korean-English bilinguals who were highly proficient in English, such that they had no difficulty in reading L2 English novels. At the time of the research, 19 Korean speakers resided in the United States, whereas the remaining 31 Korean participants resided in South Korea. Their mean iBT TOEFL score was 111.67 (SD = 4.65) and their mean length of residence in an English-speaking country was 7.89 years (SD = 3.27). Table 3.1 summarizes the Korean speakers' English learning backgrounds and proficiency levels. The KE and KC groups showed no difference in their age of exposure (U = 180.00, p = .847, r = 0.01), length of residence (U = 174.50, p = .720 r = 0.05), TOEFL score (U = 127.00, p = .474 r = 0.16), or self-rated proficiency (U = 170.00, p = .639, r = 0.10). The recruitment process involved flyers, web postings, and word of mouth, and compensation took the form of a $30 reward at the end of the experiment.

Table 3.1

Korean speakers' English learning backgrounds and English proficiency

		KE (n = 25)		KC (n = 15)	
		M	*SD*	*M*	*SD*
Age at testing		25.32	4.61	25.13	4.24
Age of exposure		5.80	1.76	5.93	1.67
Length of residence (years)		7.74	3.26	8.12	3.38
TOEFL score		112.22	4.99	110.69	3.97
Self-evaluation of proficiency	Total	34.96	1.79	34.53	2.50
	Listening	8.96	0.68	8.80	1.01
	Reading	8.88	0.78	8.60	0.78
	Speaking	8.56	0.82	8.47	1.13
	Writing	8.56	0.87	8.67	0.98

3.2.2 Targeted semi-artificial language: Koreanish

For the present study, I adopted a semi-artificial language paradigm, in which vocabulary from the participants' native or second language was rearranged to follow the target language's patterns. Specifically, the semi-artificial language involved English vocabulary and Korean syntax. To create this hybrid language, hereafter named "Koreanish", the English words from the English version of the Agatha Christie novel, *The Mysterious Affair at Styles*, were rearranged according to Korean word order. The words were thus familiar to the participants, but the syntax formed an unknown system that created difficulty in comprehension.

Word order in Korean and English seem to mirror each other, because each language has a different system of ordering heads and complements (Onnis & Thiessen, 2013). Syntactically,

English is a head-initial language in which heads of phrases precede their complements, thus forming a right-branching structure; Korean, on the other hand, is a head-final language, wherein heads follow complements to form a left-branching structure. To illustrate, the English sentence "John ate an apple" would be structured as "John apple ate" in Korean, because the verb (head) comes at the end of the phrase, and its complements must thus occur to the left. Likewise, "at church" is glossed as "church at", because the preposition (head) does not occur until the end of the phrase; that is, the Korean "at" here is postpositional. The following sentences are excerpts from the novel used in the experiment and demonstrate how an English sentence – (1) and (3) – can be rearranged according to Koreanish rules, as shown in (2) and (4) respectively.

(1) English: A vague suspicion of everyone and everything filled my mind.

(2) Korean: Everyone and everything of a vague suspicion my mind filled.

First, the PP, "of everyone and everything" is transformed into "everyone and everything of " because "of," the apposition, is the head of the phrase and is moved to the final position. Second, the NP, "a vague suspicion everyone and everything of" is reordered into "everything and everyone of a vague suspicion," as the NP, "a vague suspicion," serves as the head of the phrase. Third, the verb "filled" in the VP "filled my mind" is shifted to the end, resulting in "my mind filled." The full sentence thus reads "Everyone and everything of a vague suspicion my mind filled," and follows SOV word order.

(3) English: "I have a cousin who is a nurse," I remarked.

(4) Koreanish: "I a nurse is who a cousin have," I remarked.

First, the verb, or head, in the VP "is a nurse," "is", shifts to the end, forming "a nurse is". The CP "who is a nurse" is then changed to "a nurse is who," because "who," the relative complementizer, is the head of the CP, which should at the end of the phrase. Third, "a nurse is who" should modify the head, "a cousin," by preceding rather than following it. This results in the relative clause–noun sequence, "a nurse is who a cousin." The complex noun phrase "a nurse is who a cousin" is next merged with the verb "have," thus forming the verb phrase "a nurse is who a cousin have," wherein the verb "have," the head, migrates to the end of the phrase. Hence, the full sentence reads, "I a nurse is who a cousin have." In this manner, Koreanish consistently features the head, the main element of the phrase, in the final position.

3.2.3 Materials

The exposure task: Novel reading. The exposure text for training, *The Mysterious Affair at Styles* by Agatha Christie, provided incidental exposure to the semi-artificial word order system, Koreanish, to participants. The participants read the first two chapters, Chapters 1 and 2 (approximately 6,500 words), of *The Mysterious Affair at Styles*. The same novel has been used in a series of eye-tracking studies that examine various aspects of English monolinguals' and Dutch-English bilinguals' reading processes (Cop, Dirix, Drieghe, & Duyck, 2017; Cop, Dirix, Van Assche, Drieghe, & Duyck, 2017; Cop, Drieghe, & Duyck, 2015; Cop, Keuleers, Drieghe, & Duyck, 2015). Cop and her colleagues noted that this novel was selected based on its appropriate difficulty level for college level L2 English speakers (Flesch Reading Ease = 81.3, SMOG grade = 7.4),[4] as well as its similarity in word frequency distribution to natural language based on the Subtlex database (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014). In other

[4] Flesch Reading Ease gives a scale between 0 (difficult to read) and 100 (easy to read); higher scores reflect greater readability. The SMOG grade is an estimation of how many years of educations are required to understand the text.

words, this novel represents suitable reading material for an L2 experiment because it is easy to read and reflects real-life, natural language use.

The participants were told that they were going to read the first two chapters of a detective novel (92 screens), and that—for some of the participants—the sentences would be presented with the words scrambled. Before reading, the participants were given a brief description of the novel's main characters to help them understand the story. The experimental groups read the Koreanish version of the novel, consisting of English sentences with Korean word order, while the control groups read the English version. The participants were informed that there would be a practice session at the beginning and a short break approximately 15 screens later (total 6 breaks). At every other break, the participants were asked to answer comprehension check questions. The purpose of the comprehension test was to keep participants engaged while reading the scrambled sentences for about an hour. These comprehension tests, which contained eight simple true or false statements regarding plot-specific information, were administered three times throughout the reading. The total number of the questions was 24. One point was assigned for correct answers, and the total score of overall comprehension was 24. The overall reliability coefficients (Cronbach's alpha) were 0.78 for the EE group and 0.71 for the KE group.

It is important to note that the novel was employed not only to introduce Koreanish patterns, but also to concurrently measure the participants' implicit knowledge of Koreanish during training. Unbeknownst to participants, the exposure text comprised three phases: (a) a practice phase, (b) a training phase, and (c) a testing phase. In the practice phase, all sentences were presented in English. I selected a practice passage from the English version of the same novel and presented it to the participants in order to familiarize them with reading on the screen

and moving from one screen to the next. After that, an introductory part of the novel was presented, which provided participants with the context of the story and helped familiarize them with the plots. Next, the training phase involved an extensive amount of Koreanish sentences (*n* = 479) that flooded participants with the target syntactic input. Lastly, the testing phase was used to measure participants' sensitivity to violations. It was subdivided into three blocks of 20 sentences each: (a) a control_pre block, (b) a violation block, and (c) a control_post block. The control_pre and control_post blocks followed Koreanish word order and were used as baselines to calculate the increase in reading times in the violation block. Sentences in the violation block were ungrammatical, in that they did not follow the Koreanish word order rules. Rather, they followed the German word order (Rebuschat & Williams, 2012). Therefore, in the violation block, English sentences were rearranged in accordance with German syntactic rules, in which the verb is placed at the first, second, or final position depending on the clause type and sequence. For example, the original English sentence "In truth, he looked pathetic, totally covered in mud" would be rearranged in German as "In truth, looked he pathetic, totally in mud covered." This violation sentence contrasted with the Koreanish version of the same sentence, "Truth in, mud in totally covered, he pathetic looked". It is important that the sentences in the violation block were ungrammatical not only by Koreanish rules but also English ones; however, sometimes the violation sentences, which were rearranged in German word order, also followed English word order and were identical to the original English sentences. Those sentences were excluded from the sensitivity analysis because participants' prior English knowledge could undesirably affect the reading times for these items, which may obscure any sensitivity effects.

Three counterbalanced versions of the exposure text were created to control for lexical and topical confounds (see Table 3.2). The sentences in the testing phase rotated between the

control_pre, violation, and control_post blocks, as shown in Table 3.2. For example, 20 sentences, from Sentence 524 to Sentence 543, formed the violation block in Version B, the control-post block in Version A, and the control pre-block in Version C. Due to this rotation, the starting point of the training phase, in which Koreanish sentences were presented, varied depending on the text version. However, regardless of which version participants saw, they received the same amount of training, since they read the same amount of Koreanish sentences. To even out the number of presented sentences across all participants, participants who read Version A and Version B read an extra 40 (Sentence 544–Sentence 583) and 20 sentences (Sentence 564–Sentence 583) respectively, which were taken from the testing phase in Version C. Those extra sentences were presented in regular English syntax as a practice passage during the practice session, which enabled me to control the content familiarity with the exposure text.[5]

[5] For Version C, participants read an additional passage selected from chapter 4 (25 sentences) before they started reading the main text.

Table 3.2

The counterbalanced version of the exposure task

Phase	Syntax	*N* of sentences	Sentence ID		
			Version A	Version B	Version C
Practice	English	44	544 – 583, 1 – 4	564 – 583, 1 – 24	1 – 44
Training	Koreanish	479	5 – 483	25 – 503	45 – 523
Testing: Control_pre	Koreanish	20	484 – 503	504 – 523	524 – 543
Testing: Violation	German	20	504 – 523	524 – 543	544 – 563
Testing: Control_post	Koreanish	20	524 – 543	544 – 563	564 – 583

Online measures of learning and knowledge: Eye-movements. From the readers' eye movements during the exposure task, I calculated two online measures of learning and knowledge: changes in sentence reading times as a learning measure (DeKeyser, 1997; Godfroid, Ahn, et al., 2018; Indarathne et al., 2018) and increased reading times of (i.e. sensitivity to) violations as a knowledge measure (Godfroid, 2016; Jiang, 2007; Leung & Williams, 2011; Suzuki & DeKeyser, 2015, Suzuki, 2017; VanPatten & Smith, 2015). These eye-movement measures were advantageous, in that they enabled me to employ a within-subject design (revealing how individuals' eye movements change over time) in conjunction with a between-subjects design (showing how much extra reading time the experimental groups spend compared to the control groups). From the eye-tracking data in the exposure task, I derived changes in reading times over time and from the testing phase, I calculated changes (increases) in reading times in response to violations.

Changes in sentence reading times: learning curve. Changes in *residualized sentence reading times* were used as evidence of learning. To calculate the residualized sentence reading time, I subtracted the control group's sentence reading time for the English sentence from that of the experimental group' reading time for the corresponding Koreanish sentence. Thus, the residualized sentence reading time represents the extra processing time participants spend when reading Koreanish word order, compared to English word order. This measure yielded one difference value per sentence and enabled me to conduct a sentence-level analysis (Cop et al., 2015).

With the residualized sentence reading times, I focused on how and in what form the processing of Koreanish sentences changed over time. Specifically, I investigated whether the power law of learning can be applied to changes in residualized sentence reading times. A

learning curve described by a power function involves two critical components: a rapid decrease in the early stage of learning, followed by a gradual plateau in the later stage. If the changes in residualized sentence reading times of Koreanish are governed by the power function (i.e. $Y = X^{-n}$), where Y is the residualized sentence reading time and X is Time, this suggests that participants started processing Koreanish sentences more automatically and fluently as a result of repeated exposure (Anderson, 1982, 1993; DeKeyser, 1997).

Sensitivity to violations: integrated knowledge. Sensitivity to word order violations was employed as the concurrent measure of integrated knowledge (Jiang, 2007). Integrated knowledge brings about automatic competence in processing (Jiang, 2007). If participants succeeded in integrating Koreanish word order in their mental representation, it was expected that they would automatically slow their reading speed when processing ungrammatical sentences in the violation block, even though they were focusing on comprehension rather than form.

To compute sensitivity, I subtracted the average residualized reading times for the control blocks from those for the violation block; the difference in reading times reflects the degree of sensitivity. To better understand the nature of this measure and its underlying construct, I triangulated the sensitivity data with one of the awareness measures, verbal reports. Immediately after completing the exposure task, participants were asked if they noticed anything odd during reading. If they said no and showed an unawareness of the target rules, this was taken as evidence supporting the development of unconscious, integrated knowledge of Koreanish word order during learning.

Debriefing interview. The purpose of the debriefing interview was to probe participants' levels of awareness of the targeted Koreanish syntactic patterns (see Appendix D). The interview

was carried out over two sessions. At the end of Session 1, after the post-test (see Chapter 2), I conducted the first part of the debriefing interview to examine participants' awareness of the violation block. The goal was to determine whether participants had noticed the change in word order in the violation block. I asked the participants whether they had noticed anything odd while reading the novel. In addition, I asked them to report whether they had noticed any particular rule or regularity, to specify when they might have noticed it (i.e., during reading or on the test), and to describe what they believed they had noticed.

In Session 2, the second part of the debriefing interview was conducted online. I began by asking whether participants ever indicated recollection or rule knowledge as a source of their grammaticality judgement. If so, they were asked to explain why and state what they were thinking. In addition, I asked participants whether they had tried to search for a pattern in the scrambled text while reading. Moreover, in an attempt to tap into participants' lower awareness levels, I asked them to reflect specifically on the placement of words within the sentences and to recall any specific rule or regularity. Finally, I asked them questions about their usual reading experiences.

Apparatus. The participants' eye movements during reading were recorded with an Eyelink Portable Duo (SR Research, Canada) laptop-mounted eye-tracking camera with a sample rate of 2,000 Hz. In this study, reading was binocular, but only the right eye was recorded. A chin-rest was used to minimize head movement. The text of the novel was presented in the form of paragraphs in 16-point Consolas font, triple-spaced, on a light gray background. At a 63-seating distance, one letter subtended 0.37 degrees of visual angle. A maximum of ten lines were presented on the screen, and a drift check was performed between screens. The eye-tracker was recalibrated three times per participant, and more frequently if I deemed it necessary.

3.2.4 Procedure

The experiment was conducted in two sessions, Session 1 (90 minutes) and Session 2 (30 minutes). Session 1, which was carried out in a quiet, dimly-illuminated study room, included the exposure phase, the immediate testing phase, and the first part of the debriefing interview (short version). Session 2 comprised the delayed testing phase, the second part of the debriefing interview (long version), and the background questionnaires.

In Session 1, the participants first signed the consent form and then filled out a questionnaire regarding their language background. I explained that they were going to read two chapters from a detective novel (Appendix A) and provided a brief description of the main characters. The participants were informed that there would be comprehension questions after reading each section (Appendix B). Then the practice session started, which allowed them to become accustomed to reading on a laptop with an eye-tracker recording their eye movements. The eye-tracker was recalibrated three times per participant (at every other break) and more frequently if I deemed it necessary. They silently read the first two chapters of *The Mysterious Affair at Styles* on a laptop screen while their eye movements were recorded by the Eyelink Portable Duo eye-tracking system. The experimental groups read the novel in Koreanish, and the control groups read the English version of the novel. After they finished reading the novel, a surprise GJT with source attributions was administered for immediate testing (see Chapter 2). This test was followed by the first part of the debriefing interview regarding participants' awareness of the violation block during reading and Koreanish word order. Session 2 took place two weeks after Session 1 and involved delayed testing, along with the second part of the debriefing interview. To each participant, the researcher sent an e-mail containing the links to the test and the debriefing interview questions. The participants were asked to complete the test

within three days.

3.2.5 Analysis

In this study, I took a chronometric approach to mental processing by analyzing changes in residualized sentence reading times over repeated exposure to nonnative syntax. Before the analysis, the eye-movement recordings were screened for data quality. At first, I removed short fixations under 80 milliseconds using a built-in program in Data Viewer (SR research, Canada) to reduce noise in the data. Following Godfroid's (2020) suggestion, I inspected individual participant recordings on a trial-by-trial basis. After a careful inspection, I manually corrected fixations for drift, which is a discrepancy between the recorded eye gaze location and a participant's actual eye gaze location. Drift is generally due to human errors (e.g., participant movement) or technical errors (e.g., track loss). If it was too difficult to judge the true locations of drift fixations, I deleted the corresponding sentences to ensure the recording accuracy, which amounted to 1.3% of the entire dataset. I also decided to exclude the recordings from one Korean participant and one English participant because they skipped large parts of the reading. After cleaning the eye-movement data, I computed the residualized sentence reading times by subtracting the control group's sentence reading time for the English text from that of the experimental group for the Koreanish text. I did this for the English speakers (EE – EC) and the Korean speakers (KE – KC) separately. This new measure, which reflected sentence-level extra processing time for Koreanish sentences, was used as a dependent variable, sometimes with log-transformation if necessary. All the data from the exposure task (novel reading) were entered into the Statistical Package for the Social Sciences, version 25. The significance level alpha was set at 0.05.

Research question 1. To investigate the real-time learning effect reflected in the changes

in sentence reading time during the training phase (Blocks 1–24), I explored whether the eye-movement data followed the power law of practice (DeKeyser, 1997). Combining visual and statistical evidence, I specifically examined the effects of L1 and awareness on the learning curve. The mathematical function that describes the learning curve is a power function: $Y = X^{-n}$. The most popular approach for testing the power learning curve is to perform a logarithmic transformation. The logarithmic transformation is applied to the independent and the dependent variables. This will render the relationship between log X and log Y (i.e., the learning curve) linear if it follows a power law. Thus, a straight line in a log-log plot can be taken as evidence of the power function. Furthermore, such linearity can be statistically tested by means of a linear regression model.

Consequently, to answer research question 1-1, which investigated the effect of L1 on the learning curve, I first created log-log plots for each L1 group. On the x-axis, log Block was displayed as the time-course variable. On the y-axis, log Residualized sentence reading time was shown. Then, I used a linear mixed-effect model to test whether the two L1 groups differed in their linear relationship on the log-log scale, for instance, whether they had a different slope. The dependent variable was log Residualized sentence reading time. The independent variables were log Block (within-subject, time-course variable) and L1 (between-subject, group-difference variable). To answer research question 1-2, which explored the effect of awareness on the learning curve, I ran separate linear mixed effects models for the English and the Korean speakers because the two groups had different levels for the Awareness variable. The dependent variable was the log Residualized sentence reading time. Log Block and Awareness were entered as the independent variables. For all the analyses, the main effects and the interactions with log Block were of interest. I fitted the models with random by-subject intercepts and random slopes

for log Block to account for individual variations in how experimental participants processed the Koreanish sentences initially and over time.

Research question 2. To investigate the effect of L1 and awareness on online sensitivity during the testing phase (Blocks 25–27), I performed a series of mixed -design ANOVAs. Online sensitivity was assessed by means of the differences between the violation block (Block 26) and the control-pre block (Block 25) and the control-post block (Block 27). The main effects and the interactions of Block were of interest. To answer research question 2-1, which focused on the effect of L1 on sensitivity, I carried out a 2 x 3 mixed-design ANOVA on residualized sentence reading times with L1 (English, Korean) as a between-subject variable and Block (control_pre block, violation, control_post block) as a within-subject variable. Additionally, I performed one-sample t-tests on each group's sensitivity score using zero as a test value. A positive sensitivity score indicates that the participant is sensitive to grammatical violations because of their integrated, implicit knowledge of the target syntax; a negative sensitivity score could signal reverse sensitivity, which may be indicative of the activation of non-targetlike grammar rules. In order to answer research question 2-2, which examined whether subgroups with different awareness levels differed in sensitivity, I performed a mixed-design ANOVA for each L1 group separately. The analyses were conducted on Residualized sentence reading times with Awareness (English: unaware, VP-aware; Korean: unaware, VP-aware, L1-aware) as a between-subject variable and Block (control pre_block, violation, control post_block) as a within-subject variable.

3.3 Results

3.3.1 Comprehension test

The overall comprehension scores of the EE and KE groups were 18.17 (SD = 2.51) and 18.13 (SD = 3.17), respectively, which indicated that their comprehension level of the scrambled text was acceptable. The difference between the two L1 groups' comprehension scores was not significant [$t(46) = 0.31$, $p = .759$, $d = 0.09$]. The overall comprehension scores for the control groups, the EC and KC groups, were 23.13 (*SD* = 0.96) and 22.80 (*SD* = 1.05), respectively.

3.3.2 Analyses of retrospective verbal reports

The retrospective verbal report data from the debriefing interviews from Sessions 1 and 2 were analyzed to gauge participants' level of awareness. To enhance the inter-rater reliability in the coding, two independent raters (the researcher and a trained rater) coded the verbal report data, which was gathered from two debriefing interview sessions. The raters identified 1) what type of awareness the participants developed (unaware, VP-aware, L1-aware), 2) when they became aware of the rules (exposure and testing), and 3) when they reported awareness (Session 1 and Session 2). The reliability coefficient (kappa) for the identification of awareness type was $\kappa = 0.94$, for the identification of the timing of awareness was $\kappa = 0.99$, and for the identification of the timing of reporting was $\kappa = 1.00$.

This analysis enabled me to identify subgroups of participants based on their reported awareness of the target rules. Note that only the experimental groups, who read the Koreanish version of the novel, participated in the debriefing interview.

From the verbal report data, I attempted to confirm the internal validity of the incidental learning condition (cf. Leow, 1997; Leow & Hama, 2013); that is, whether participants engaged in primarily meaning-focused processing as intended. To this end, I asked the participants

whether they tried to search for patterns and focused on word order while they were reading. This inspection led me to exclude two participants—one from each L1 group—who engaged in intentional learning. Their answers, presented in Example (6), clearly showed that they intentionally searched for a pattern while reading, which meant that their attention was primarily oriented toward forms.

Example (6)

P119 (English participant)

"At first, I did not think there was a pattern. But after the first break, I started to notice a pattern, so I would try to figure it out and unscramble the sentences. I realized that the verb at the end of a dependent clause was the action of the first noun in the sentence."

P304 (Korean participant)

"I have noticed that the order of noun, verb, and etc. were messed up regularly. I did try to find the order of the words in the sentence but could not find a clear pattern."

The verbal data of the remaining 48 participants (English: n = 24; Korean: n = 24) did not show evidence of such rule-search behavior. An analysis of the English speakers' verbal report data indicated that half of the participants—12 of 24 participants—became aware of the VP rule. This was a partial awareness of Koreanish word order rules, since participants did not mention any other rules regarding head finality. Based on their awareness of the VP rule, I categorized them as VP-aware. Among the 12 participants, 10 participants reported awareness of the VP rule at Session 1, during the short version of the interview, and two participants reported awareness at Session 2 (two weeks after Session 1), during the long version of the interview. All participants

said they became aware of the VP rule during the GJT. The other 12 participants did not report awareness of any target rules and were categorized as unaware.

Unlike the analysis of the English speakers' data, an analysis of the Korean speakers' verbal report data exhibited three types of awareness: L1 aware (n = 9), VP-aware (n =10), and unaware (n = 5). First, I found an interesting subgroup, the L1-aware group, who reported their awareness of the relationship between the target language and their L1 (Korean). These nine participants reported that they became aware that the language used in the novel followed Korean word order while they were taking the GJT. It is important that they were unaware of any rules (n = 7) or only discovered the VP rule (n = 2) during the exposure phase; it was in the testing phase that they figured out that the Koreanish sentences in the novel matched their L1. In terms of the timing of reported awareness, among 9 L1-aware participants, 2 participants reported their L1 awareness after taking the immediate GJT at Session 1; the rest, 7 participants, reported their awareness after taking the delayed GJT at Session 2. Second, 10 participants became VP-aware. They reported noticing the VP rule but did not notice other rules regarding head finality in the Koreanish text. All VP-aware participants reported their VP-awareness at Session 1 after taking the immediate GJT. Lastly, five participants remained verbally unaware and were categorized as unaware.

Additionally, I also aimed to ensure the nature of online sensitivity (i.e., conscious or unconscious) from the retrospective verbal reports. Given that online sensitivity can be evidenced by an automatic slowdown in processing of violations without awareness (Godfroid, 2016; Jiang, 2007), which may take place outside awareness, I asked the participants whether they noticed anything odd or changes in the pattern while they were reading the two chapters of the novel. This information enabled me to exclude four participants (two from each L1 group),

who noticed the presence of the violation block during the testing phase (see Example 7).

Example (7)

P 117 (English participant)

"Not quite at the end, but close to the end, the sentences appeared to go back to a normal sentence structure. Then they went back to being in a strange structure."

P 326 (Korean participant)

"I think the grammar structure at the end may have changed."

As shown in Example 4, some participants noticed that the word order changed at the violation block, and I decided to exclude these violation finders from the main sensitivity analyses. These four violation finders showed a lower sensitivity score (M = -55, SD = 346) than the rest of the participants (see Table 3.6, Englsih: M = -12, SD = 483; Korean: M = 158, SD =292). This indicated that after noticing the change in the word order, they read the violation sentences, which followed German word order, faster than the Koreanish sentences.

3.3.3 Changes in residualized sentence reading time: Learning curve

The visualizations of the log-log plots for the English speakers' and the Korean speakers' residualized sentence reading times are presented in Figures 3.1 and 3.2, respectively. In these graphs, log Block is on the x-axis, and log Residualized sentence reading time is on the y-axis.

The two log-log plots revealed distinctive patterns. While the English speakers' data fit a straight line, such was not the case for the Korean speakers' data. To test for differences in the linear relationship, I carried out a linear mixed-effect model. The dependent variable was log Residualized sentence reading time. The independent variables were the log Block as a time-

course variable and L1 as a grouping variable. I tested the following series of models to determine whether the addition of fixed effects and interactions improved the model fit:

- M0, an intercept-only model;
- M1, a model adding a fixed effect for log Block;
- M2, a model adding a fixed effect for L1;
- M3, a model adding an interaction term between log Block and L1.

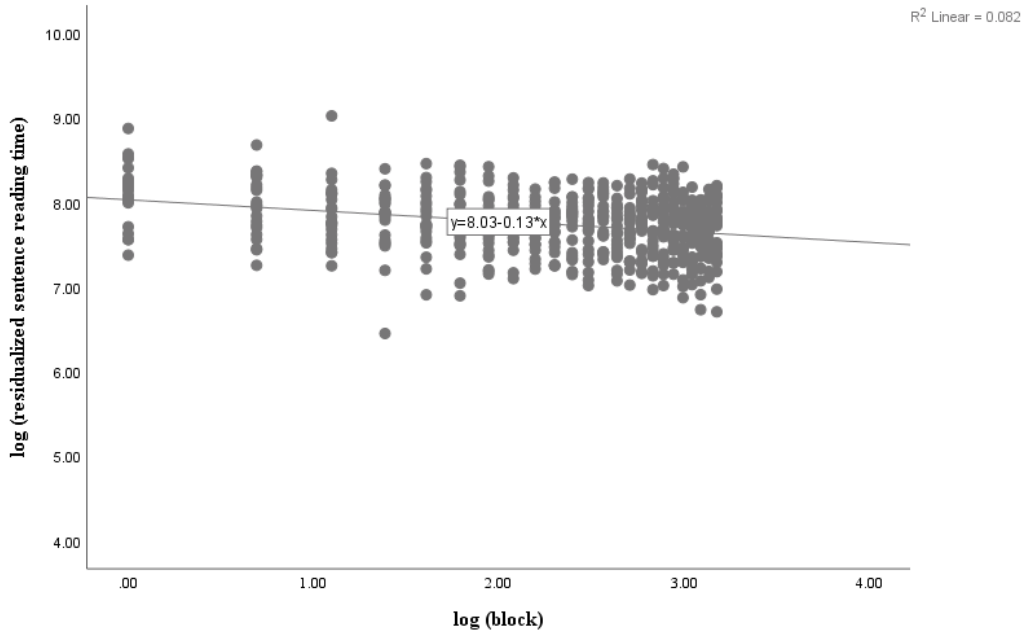

Figure 3.1

The English speakers' log-log plot

All models, except for M0, included random slopes for log Block and the random by-subject intercepts. M0 included only the random by-subject intercept. The model fit was assessed by evaluating the change in chi-square statistic through a comparison of the successive models' goodness of fit (Baayen, 2008). Each consecutive model yielded a significantly better fit than the

previous one [M1: $\chi 2$ (3) = 114.26, $p < .01$; M2: $\chi 2$ (1) = 4.78, $p < .05$; M3: $\chi 2$ (1) = 10.43, $p < .01$].

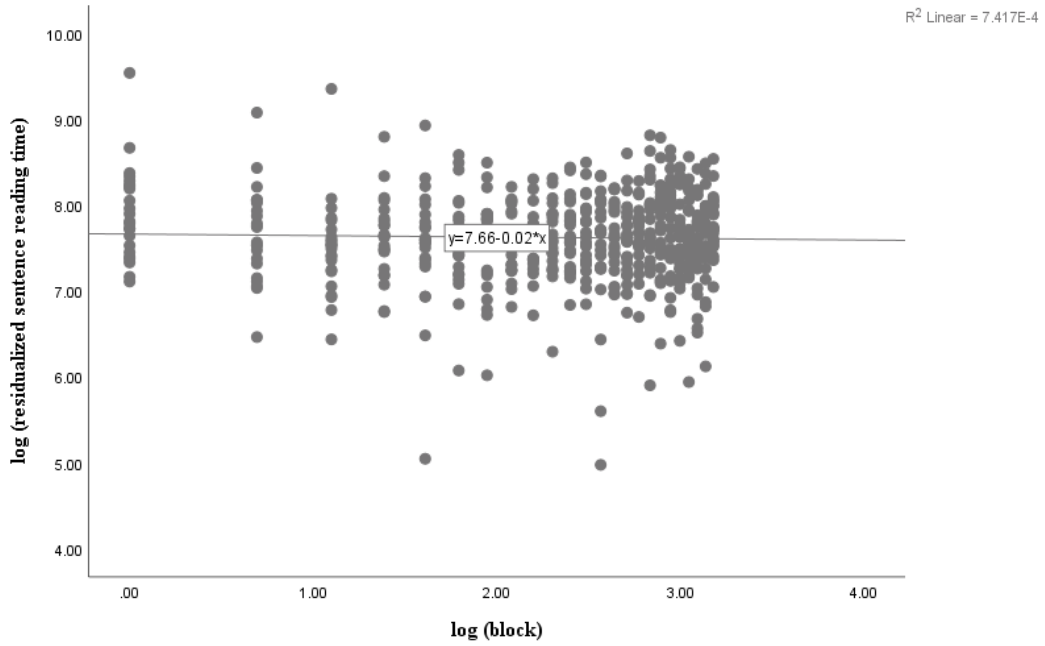

Figure 3.2

The Korean speakers' log-log plot

The final model (M3) (Table 3.3) revealed that L1 [$F(1, 46) = 6.56$, $p = .014$], log Block [$F(1, 46) = 15.87$, $p < .001$], and the interaction between L1 and log Block [$F(1, 46) = 9.10$, $p = .004$] significantly predicted log Residualized sentence reading time. The relationship between log Block and log Residualized sentence reading time showed significant variance in the intercepts [var(u0j) = 0.21, $\chi 2$ (1) = 968.97, $p < .01$] and the slopes [var(u1j) = 0.01, $\chi 2$ (1) = 46.71, $p < .01$] across the participants. Crucially, the two-way interaction indicated that the two L1 groups followed different learning trajectories.

Next, I fitted separate multilevel models to the English and the Korean speakers to

understand the nature of the interaction. In the separate models, I retained log Block as a predictor; additionally, to answer research question 1-2, I included the main effect of Awareness in the models. For the English speakers, Awareness was a two-level factor (unaware and VP-aware); for the Korean speakers, it was a three-level factor (unaware, VP-aware, and L1-aware). The visualizations of the log-log plots for the English speakers' and the Korean speakers' residualized sentence reading times by awareness are presented in Figures 3.4 and 3.5, respectively.

Table 3.3

Regression output of the final model (M3)

	b	*SE*	*p*
Fixed effects			
Intercept	7.66	0.10	< .001
Log block	-0.17	0.03	.497
L1: English	0.37	0.14	.014
Log block x L1	-0.11	0.04	.004
Random effects by subject			
Within-person variance	0.06	2.68E-3	< .001
Between-person intercept	0.21	0.05	< .001
Between-person slope: log block	0.01	3.10E-3	< .001

For the English speakers, the log block significantly predicted log Residualized sentence reading time [b = -0.13, SE = 0.03, t(23) = -4.03, p = .001]. However, the main effect of Awareness [F(1, 23) = 0.02, p = .089] and the interaction between Awareness and log Block

[$F(1, 23) = 0.01$, $p = .097$] were not significant (Table 3.4). Thus, there was a significant linear relationship between log Residualized sentence reading time and log Block, which supported the notion that the English speakers' residualized sentence reading times decreased according to a power function. However, for the Korean speakers, the main effects of log Block [$F(1, 23) = 0.58$, $p = .455$], Awareness [$F(2, 23) = 0.36$, $p = .702$], and the interaction between Awareness and log Block [$F(2, 23) = 0.22$, $p = .803$] were all not significant (Table 3.5), which indicated that their learning curve on the log-log scale was essentially flat.

Table 3.4

Regression output for the English speakers

	b	*SE*	*p*
Fixed effects			
Intercept	8.02	0.11	< .001
Log block	-0.13	0.03	.001
Awareness: VP-aware	0.02	0.16	.888
Log block x awareness	-1.47E-3	0.45	.974
Random effects by subject			
Within-person variance	0.03	1.93E-3	< .001
Between-person intercept	0.13	0.04	.002
Between-person slope: log block	9.03E-3	3.49E-3	.005

Table 3.5

Regression output for the Korean speakers

	b	*SE*	*p*
Fixed effects			
Intercept	7.74	0.25	< .001
Log block	-0.05	0.06	.434
Awareness: L1-aware	5.17E-3	0.31	.987
Awareness: VP-aware	-0.20	0.31	.530
Log block x L1-aware	0.05	0.07	.515
Log block x VP-aware	0.03	0.07	.716
Random effects by subject			
Within-person variance	0.09	5.67E-3	< .001
Between-person intercept	0.28	0.09	.002
Between-person slope: log block	0.01	5.18E-3	.022

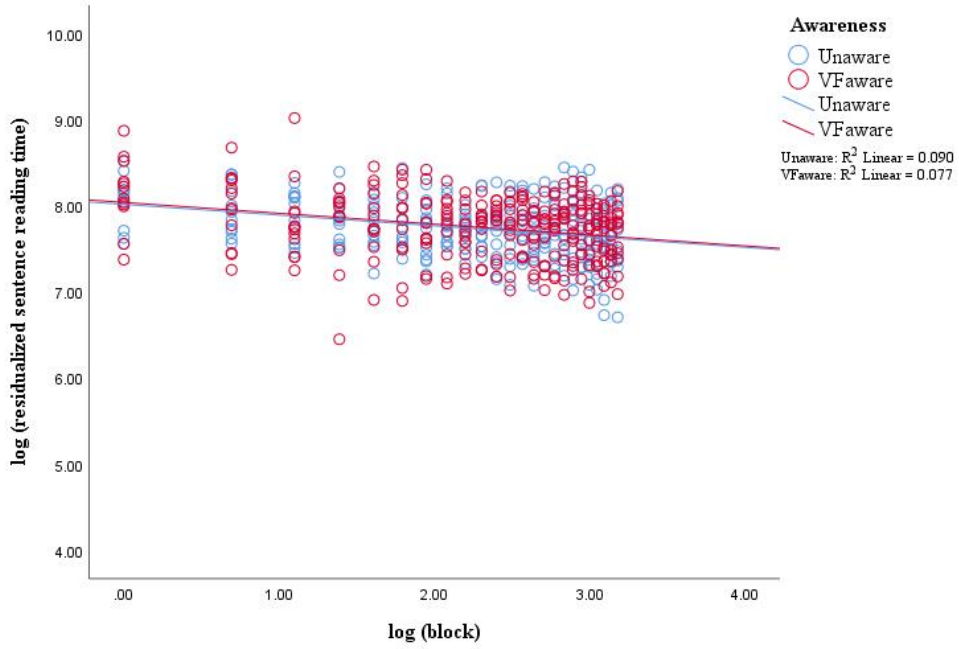

Figure 3.3

The log-log plot for unaware and VP-aware English speakers

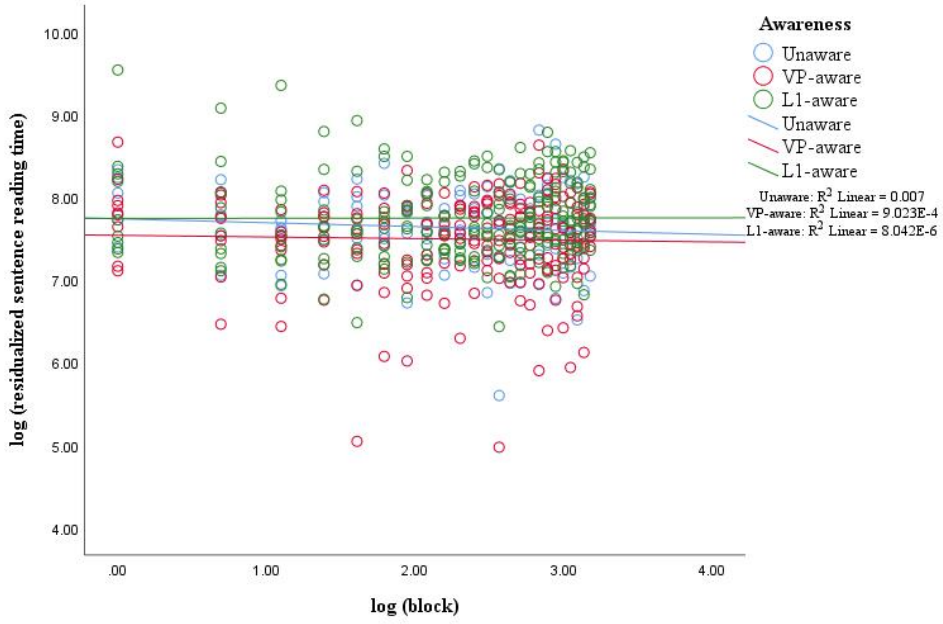

Figure 3.4

The log-log plot for unaware, VP-aware, and L1-aware Korean speakers

To sum up, the significant two-way interaction between log Block and L1 and the follow-up analyses showed that the two L1 groups followed a different learning curve. On the one hand, the English speakers' residualized sentence reading time data followed the power learning curve; on the other hand, the Korean speakers' residualized sentence reading times did not follow a clearly discernible pattern. For both L1 groups, there was no significant relationship between awareness of the target syntax and residualized sentence reading times. Therefore, this diverging pattern suggested that prior L1 knowledge affected the learning curve under incidental exposure (RQ 1-1), regardless of whether participants later developed conscious knowledge about the target syntax (RQ 1-2).

3.3.4 Sensitivity to violations: Integrated knowledge

The descriptive statistics of the English speakers' and the Korean speakers' residualized sentence reading times on the testing block are presented in Table 3.6 and visualized in Figure 3.5. Figure 3.5 showed that the two L1 groups responded very differently to the violations; the Korean speakers slowed down, whereas the English speakers sped up.

Table 3.6

The English speakers' and the Korean speakers' sensitivity to violations

	L1: English (*n* = 21)			L1: Korean (*n* = 21)		
	M	*SD*	*95% CI*	*M*	*SD*	*95% CI*
Control_pre	497	533	[254, 739]	713	718	[386, 1040]
Violation	387	537	[143, 632]	825	847	[440, 1211]
Control_post	302	657	[2, 601]	625	750	[283, 967]
Sensitivity score	-12	483	[-231, 208]	157	292	[24, 290]

Note. The sensitivity score was computed as RSRTviolation – ((RSRTcontrol-pre + RSRTcontrol-post)/2)

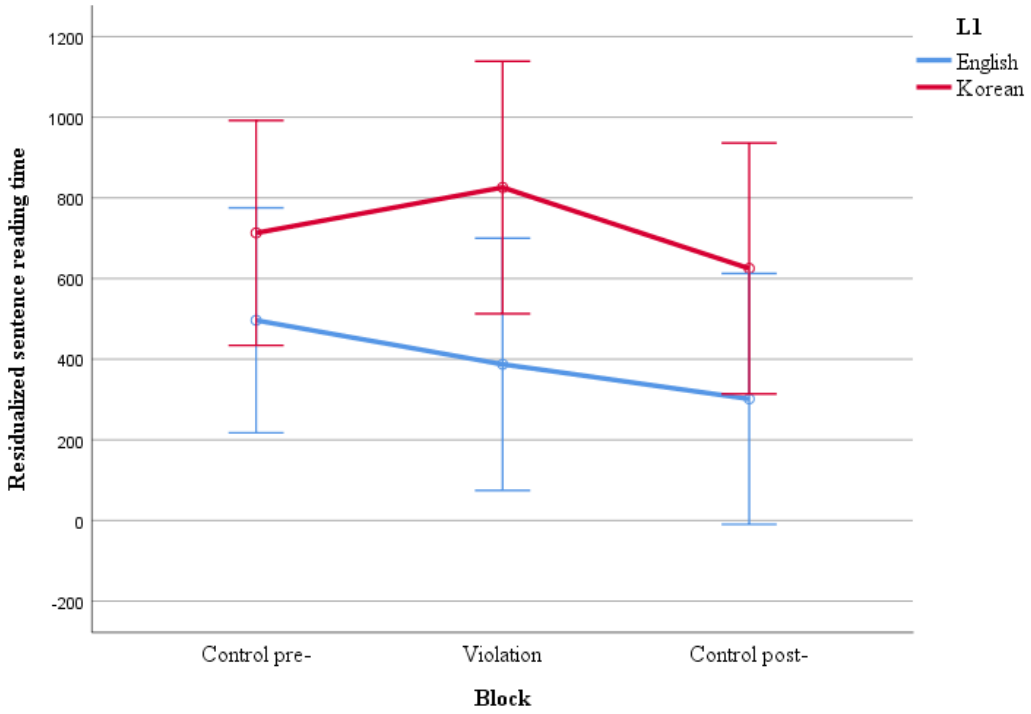

Figure 3.5

The English and the Korean speakers' sensitivity during the testing phase

To investigate whether prior L1 knowledge would affect the development of sensitivity to violations under incidental exposure, a 2 (L1) x 3 (Block) mixed-design ANOVA was conducted. The Shapiro–Wilk test showed that the distributions of the English and the Korean speakers' residualized sentence reading times by block were all normal ($p > .05$). This analysis showed the significant main effect of Block [$F(2, 80) = 3.23$, $p = .045$, $\eta_p^2 = 0.08$], but the main effect of L1 [$F(1, 40) = 2.74$, $p = .106$, $\eta_p^2 = 0.06$] and the interaction between L1 and Block [$F(2, 80) = 1.47$, $p = .235$, $\eta_p^2 = 0.04$] were not significant. The post-hoc paired t-tests showed that the significant difference stemmed from the difference between the control pre-block and the control post-block [$t(41) = 2.47$, $p = .018$, $d = 0.38$]. The difference between the violation and the

control post-block was marginally significant [$t(41) = 1.98$, $p = .055$, $d = 0.35$], while the difference between the violation and the control-pre block was not significant [$t(41) = -0.03$, $p = .978$, $d = 0.00$]. Therefore, the ANOVA results did not provide evidence of online sensitivity because it suggested that the participants did not show a robust effect of a slowdown in reading times during the violation block.

Additionally, I ran one-sample t-tests on each L1 group's sensitivity score using zero as a test value (English: $M = -12$; Korean: $M = 157$). The English speakers' sensitivity score was not significantly different from zero [$t(20) = -0.11$, $p = .913$, $d = 0.02$], whereas the Korean speakers' sensitivity score was positive [$t(20) = 2.46$, $p = .023$, $d = 0.54$]. However, the latter result should be interpreted with caution because as seen in the ANOVA results, this positive score was primarily due to the decrease in reading time during the control-post block rather than the increase during the violation block. Thus, it could be concluded that there was some evidence for the Korean speakers' sensitivity, while the English speakers did not show any evidence of sensitivity.

The descriptive statistics of the English and the Korean speakers' sensitivity to violations by awareness are presented in Table 3.7. To answer research question 2-2, which investigated the relationship between awareness and sensitivity, I ran separate analyses on each L1 group. The reason was that each L1 group had different levels of Awareness (English: 2 levels; Korean: 3 levels). For the English experimental group, a 2 (Awareness) x 3 (Block) mixed-design ANOVA on the group's residualized sentence reading time was carried out. According to the Shapiro–Wilk test, the distributions of the residualized sentence reading times for each block were all normal ($p > .05$). The analysis showed that the main effect of Block [$F(2, 38) = 1.86$, $p = .170$, $\eta_p^2 = 0.09$], awareness [$F(1, 19) = 0.14$, $p = .712$, $\eta_p^2 = 0.01$], and the block by awareness

interaction [$F(1, 19) = 0.01$, $p = .918$, $\eta_p^2 = 0.00$] were all not significant, which meant that there was no significant relationship between the English speakers' awareness of the target syntax and their sensitivity to Koreanish word order violations.

Table 3.7

The English speakers' and the Korean speakers' sensitivity to violations by awareness

		Control_pre		Violation		Control_post		Sensitivity score	
	n	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
L1: English									
Unaware	12	514	563	470	552	312	642	57	503
VP-aware	9	473	523	276	528	287	715	-104	469
L1: Korean									
Unaware	4	630	644	296	196	413	362	-225	229
VP-aware	8	472	643	566	714	275	665	192	264
L1-aware	9	964	871	1292	939	1030	802	295	191

For the Korean experimental group, the Shapiro–Wilk test indicated that the distributions of the residualized sentence reading times for each block were all normal ($p > .05$), except for the unaware group's residualized sentence reading time for the control-post block. A 3 (Awareness) x 3 (Block) mixed-design ANOVA showed a significant main effect of Block [$F(2, 36) = 3.58$, $p = .038$, $\eta_p^2 = 0.17$], which was qualified by an Awareness by Block interaction [$F(4, 36) = 2.94$, $p = .033$, $\eta_p^2 = 0.25$] (Figure 3.6). The main effect of Awareness was not significant [$F(2, 18) = 2.25$, $p = .134$, $\eta_p^2 = 0.02$].

To follow up on the two-way interaction, I analyzed the simple effect of Block for each

Awareness group separately. The follow-up tests revealed a significant effect of Block for the L1-aware group [$F(2, 16) = 4.87$, $p = .022$, $\eta_p^2 = 0.38$] but not for the VP-aware group [$F(2, 14) = 2.75$, $p = .098$, $\eta_p^2 = 0.28$] and the unaware group [$F(2, 6) = 1.96$, $p = .221$, $\eta_p^2 = 0.40$]. Although there was no significant effect of Block for the unaware group, the effect size was large ($\eta_p^2 = 0.40$), which may indicate a power issue due to a small sample size ($n = 4$). In fact, Table 3.7 and Figure 3.6 suggested that the Korean unaware group's residualized sentence reading times showed the opposite tendency as the VP-aware and the L1-aware groups during the testing phase. The unaware group's reading time for the violation block was faster (296 ms) than their reading time for the control_pre (630 ms) and the control_post (413 ms) blocks, whereas the L1-aware (1292 ms) and the VP-aware (566 ms) groups' reading times for the violation block were slower than their reading times for the control_pre (L1-aware: 964 ms; VP-aware: 472 ms) and the control_post blocks (L1-aware: 1030 ms; VP-aware: 275 ms). For the L1-aware group, the mean sentence reading time differences between the violation block and the control pre-block (328 ms) as well as between the violation block and the control-post block (-262 ms) were statistically significant (control-pre block: $t(8) = 3.22$, $p = .012$, $d = 1.08$; control-post block: $t(8) = -3.05$, $p = .016$, $d = 1.01$). These significant differences indicated that the L1-aware group increased their reading times for ungrammatical Koreanish sentences and then decreased their reading times for grammatical Koreanish sentences. Although the L1-aware group did not become aware of the relevance of the L1 until the grammaticality judgment posttest and, hence, was unaware of the relationship between Koreanish and their L1 (Korean) during reading, they unconsciously exhibited a large amount of sensitivity during the critical violation block. Moreover, the L1-aware participants were the slower, more careful readers at a descriptive level (see Table 3.7 and Figure 3.6).

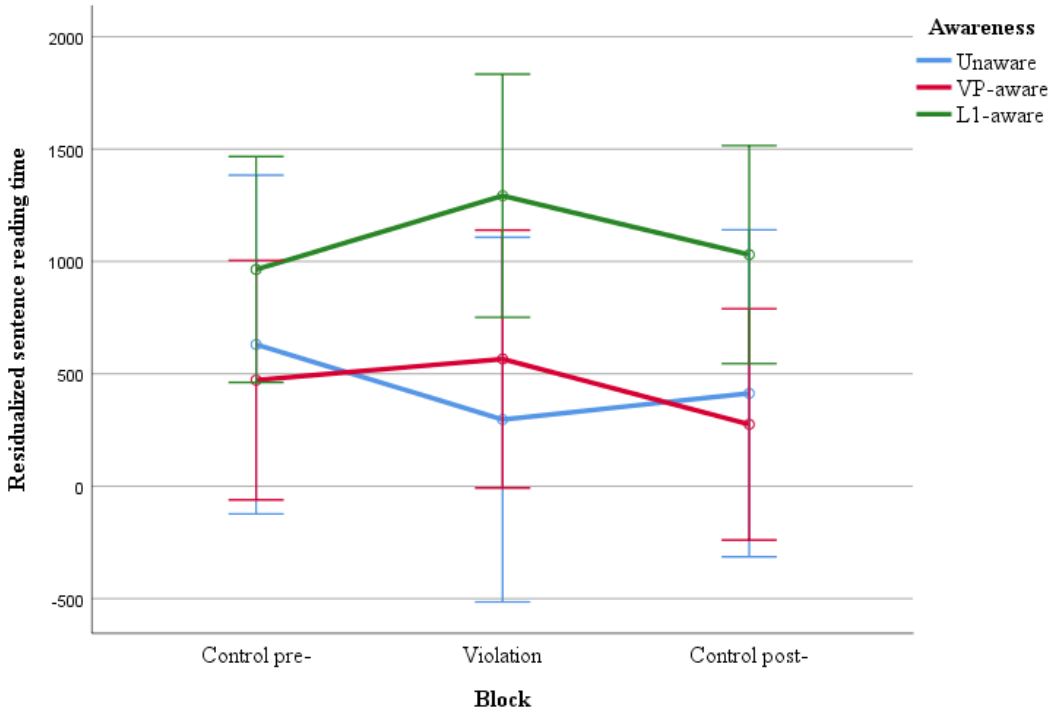

Figure 3.6

Two-way interaction between Block and Awareness for the Korean speakers

3.3.5 Summary of results

- The learning curve seen in the English speakers' reading time data followed the power law of practice; however, the Korean speakers' reading time data did not follow a clearly discernible pattern.
- The English speakers did not show statistically reliable sensitivity to word order violations.
- There was a suggestion of the Korean speakers' sensitivity to violations.
- The Korean L1-aware group members, who later became aware of the cross-linguistic similarity between the target language and their L1, showed robust sensitivity to syntactically ungrammatical sentences in the input.
- The Korean VP-aware group showed a response pattern that was consistent with the L1-

aware group.

- The Korean unaware group showed a qualitatively different response.

3.4 Discussion

Using an eye-tracking methodology in this study, I demonstrated the effects of prior L1 knowledge on online learning (i.e., a decrease in sentence reading times) and knowledge (i.e., sensitivity to violations) of nonnative syntax. The eye-movement data revealed an intriguing interaction between L1 knowledge and online L2 development. On the one hand, changes in the English speakers' residualized sentence reading times followed the power learning curve during the training phase; however, these participants failed to show sensitivity to violations during the testing phase. On the other hand, changes in the Korean speakers' residualized sentence reading times did not fit the power learning curve, but the Korean L1-aware group showed sensitivity to word order violations and the Korean VP-aware group (but not the unaware group) showed a response pattern that was consistent with this. In this way, the two L1 groups' different patterns on the two online measures captured multifaceted aspects of L1 influence on L2 development, uncovering the complex nature of linguistic transfer, variability in L2 development, and the moderating role of awareness in meaning-focused, incidental learning conditions.

Residualized sentence reading times reflect the amount of extra attention needed for processing Koreanish input, and changes in this measure could reveal online learning processes. Under incidental exposure, the English speakers' residualized sentence reading times followed the power learning curve (DeKeyser, 1997), which is known to be a hallmark of domain-general learning (Anderson, 1982, 1993; Newell & Rosenbloom, 1981). Given that the power function has been commonly found in intentional learning contexts (DeKeyser, 2015), I extended

DeKeyser (1997) by observing the the power function even under incidental learning conditions. The nonlinear decrease in the English speakers' residualized sentence reading times (Godfroid, Ahn, et al., 2018; Indrarathne et al., 2018) likely reflects a gradual withdrawal of attentional processing. Regardless of whether they obtained conscious rule knowledge (i.e., awareness), the English speakers showed a lower need for controlled attentional processing over time. This gradual removal of attention can be taken as an indication of increased processing efficiency.

In fact, the participants' self-reported reading patterns enabled me to infer how they controlled their attentional processing. In the debriefing interview, I asked the participants what they were thinking when reading the assigned novel and how they read the sentences with the words scrambled. The analyses of their answers led me to conclude that twice the number of English speakers (n = 19) than Korean speakers (n = 9) tended to use self-developed reading strategies for comprehension. Their verbal reports revealed two broad types of reading strategies—meaning-oriented and form-oriented ones. First, four English speakers used a meaning-oriented strategy by focusing on keywords in the sentences. They reported that they "extracted keywords at a glance" and "found skimming to be more effective than trying to read each sentence syntactically." To answer the comprehension check questions, they chose to pay attention to important information to understand the content of the story. Second, more than half of the English speakers (15 of 23) used a form-focused strategy by searching for the subject and the verb in each sentence to rearrange the structure. After realizing that the verb was placed at the end of the sentence, they deliberately "read the verb after reading the subject of the sentence" to "put the words in an English sentence structure mentally to try and understand." Although this form-focused strategy could have contributed to the acceleration of their reading speed (as they became faster at locating the verb and the subject), it also suggested that they were still

processing the Koreanish sentences based on their L1 (English) representation. A similar L1-based parsing tendency was also shown in Onnis and Thiessen's (2013) study. Further, it also revealed that the incidental, meaning-focused condition elicited some conscious, deliberate processing strategies as well. This might be the reason that the power law of practice was applicable even in the incidental context.

Crucially, the English speakers failed to show sensitivity to word order violations, suggesting that they had not integrated their new syntactic knowledge yet. Incidental exposure from reading a novel was probably insufficient for English speakers to develop a complete set of integrated, implicit knowledge (Williams & Kuribara, 2008), primarily due to the dissimilarity between English and Korean word order. Because the Koreanish word order had exactly the reverse pattern of their L1 English, they would perhaps need more explicit instruction, such as rule-search instruction or metalinguistic explanations, to master the L2 rules that differed from their L1 rules (Indrarathne & Kormos, 2017). For example, a recent study (McManus & Mardsen, 2019) found that L2 practice, combined with explicit instruction about the L1–L2 differences in form-meaning mapping, significantly facilitated L2 speed and accuracy, which pointed to the pedagogical value of instruction on the L1–L2 differences for L2 grammatical development. In this sense, L2 teaching would benefit greatly from an in-depth understanding of the L1–L2 relationship, especially when a great deal of differences between learners' L1 and L2 exist. In other words, in an L2 classroom, it should be kept in mind that the starting point of L2 acquisition is the L1.

In contrast to the English speakers' reading data, the Korean speakers' residualized sentence reading times did not fit the power function. What then do the changes in their sentence reading times indicate? A visual representation of the two L1 groups' residualized sentence

reading times by block (Figure 3.7) suggests that the Korean speakers showed a sharper initial decrease during the early stage of learning compared with the English speakers.

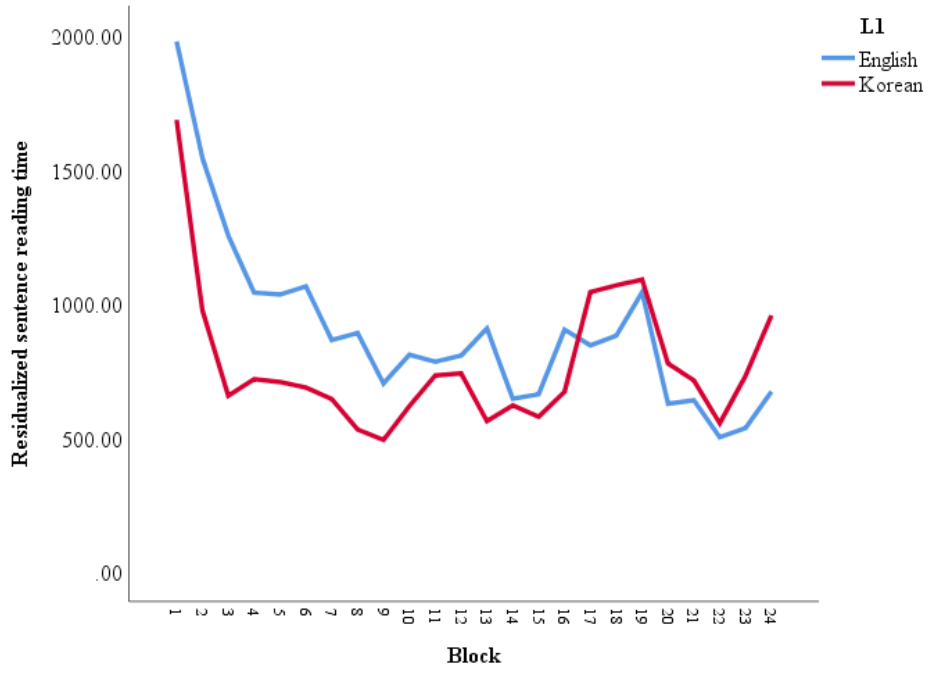

Figure 3.7

The English and the Korean speakers' residualized sentence reading times across blocks

To test the group difference shown in the early stage of reading, I conducted an additional 2 (L1) x 24 (Block) mixed ANOVA on log-transformed residualized sentence reading times.[6] The

[6] Log-transformation was used to meet the normality assumption. According to the Shapiro–Wilk test, the distributions of both the English and the Korean speakers' raw data were not normal ($p < 0.01$). However, with the log-transformation, the distribution of the English speakers' log Residualized sentence reading time was normal ($p = 0.245$); the distribution of the Korean speakers' data was still not normal ($p < 0.01$).

results[7] indicated that the Korean speakers read faster than the English speakers during the very early stage of reading the novel (from Sentence 20 to Sentence 60), but after that, reading times for the two groups could not be distinguished statistically. Thus, above all, the Korean speakers showed a steeper initial decrease than the English speakers did.

The Korean speakers' faster initial speedup could be attributable to their prior L1 knowledge of the Korean word order. Given that the Korean speakers in this study were Korean-English bilinguals, they possibly could have spontaneously accessed their L1 (Korean) while reading the vocabulary in L2 English because among bilinguals, both their languages are known to be automatically co-activated even in a task performed in one language (Hatzidaki, Branigan, & Pickering, 2011; Sanoudaki & Thierry, 2015). Perhaps the cross-language activation helped the Korean speakers become familiar with the Koreanish syntax more rapidly than the English speakers, accelerating their speed of reading Koreanish sentences. However, despite the sharp drop-off in their reading times, they did not show a consistent and gradual decrease for the rest of their reading as the English speakers did. Rather, they showed subsequent peaks and valleys in their residualized sentence reading times. One possibility is that the Korean speakers' unstable processing patterns could be accounted for by the representational redescription model (Karmiloff-Smith, 1992). According to this developmental framework, when learners build their abstract rule knowledge from the procedural representation (i.e., Explicit Level 1), they tend to make more errors as the knowledge is redescribed into a more efficient format. During this

[7] The analyses showed a significant main effect of Block [$F(9.38, 412.53) = 5.84$, $p = .020$, $\eta_p^2 = 0.12$], which was qualified by a significant L1 by Block interaction [$F(9.38, 412.53) = 2.63$, $p = .005$, $\eta_p^2 = 0.06$]. The simple effect test of L1 for each Block showed that this effect was significant only for Block 2 [$F(1, 44) = 4.52$, $p = .039$, $\eta_p^2 = 0.09$] and Block 3 [$F(1, 44) = 6.20$, $p = .017$, $\eta_p^2 = 0.12$].

period, their performance may present a U-shaped learning pattern. In this regard, it could be assumed that the Korean speakers' fluctuating reading patterns were reflections of potential error-correcting behaviors, which led them to formulate unconscious structural knowledge.

Another notable finding of this study is that the Korean speakers, especially the L1-aware group, exhibited integrated knowledge of Koreanish as reflected in their sensitivity to word order violations. The L1-aware group members, who later recognized the cross-language similarity after finishing the exposure task, already showed an automatic slowdown in their reading times for violations prior to that. Such a sensitivity effect indicated that they developed integrated knowledge of the Koreanish word order under incidental exposure conditions. It is notable they did not report any awareness of the violations in their verbal reports, which suggest that their slowdown in processing of violations might have taken place outside awareness; therefore, it can be argued that their integrated knowledge was also implicit and unconscious. The Korean L1-aware group's sensitivity probably stemmed from their unconscious access to their well-established L1 (Korean), which could be taken as evidence of unconscious L1 transfer. They in turn developed metacognition of the cross-language similarity, which I referred to as L1 awareness in this study. In this way, the L1-aware participants' performance showed an interesting connection between unconscious and conscious access to prior L1 knowledge. This finding appears to be in line with the radical plasticity theory proposed by Cleeremans (2007, 2011). According to Cleeremans (2011), unconscious knowledge, which can give rise to sensitivity effects, will shape higher-order consciousness as a result of learning. When the brain unconsciously processes information and learns about its own unconsciously accrued representations, one will eventually go through conscious experience of the unconscious knowledge and become aware of what we know. This theory could explain the L1-aware group's

performance; the ones who showed evidence of unconscious access to Koreanish (i.e., integrated knowledge) subsequently gained conscious access to the metacognitive knowledge of the Korean–Koreanish relationship. In this way, the L1-aware group's L2 development showed an interesting progression from unconsciousness to consciousness.

The methodological approach in this study resulted in a more solid understanding of the nature of grammar development. The use of control groups was advantageous, because they provided baseline eye-tracking data, which enabled the researcher to more accurately assess extra processing times. This methodological feature enabled me to employ a within-subject design (revealing how individuals' eye movements change over time) in conjunction with a between-subjects design (showing how much extra reading time the experimental groups needed when compared to the control groups). In addition, triangulating online measures with awareness measures was highly informative, particularly when making inferences about cognitive processes in incidental learning conditions (Rebuschat, 2013; Rebuschat et al., 2015). This triangulation allowed me to probe the complex, multi-faceted effects of prior L1 knowledge during processing, which could not have been possible with traditional accuracy measures.

CHAPTER 4

GENERAL DISCUSSION AND CONCLUSION

4.1 Summary of the Findings

4.1.1 English speakers

Online measures. During the exposure task, the English experimental group's change in residualized sentence reading times followed the power function. This learning curve, which was governed by the power law of practice, was indicative of learning processes that involved proceduralization and automatization (Anderson, 1982, 1993; DeKeyser, 1997, 2015). Nonetheless, when faced with word order violations, the English experimental group' online grammatical sensitivity was not statistically significant (no slowdown). This suggested that, in spite of their reading time gains, the English speakers did not show evidence of integrated knowledge of head-final word order. There were no significant differences in reading patterns or grammatical sensitivity for the awareness subgroups.

Offline measures. The analysis of the English group's GJT performance suggested that the English experimental group was able to acquire the target grammar after incidental exposure through a novel. The English experimental group exhibited strong evidence of explicit knowledge; however, evidence of their implicit knowledge was weak and limited. The effect of awareness was not significant for syntactic knowledge development.

4.1.2 Korean speakers

Online measures. During the exposure task, the Korean experimental group showed a strong initial decrease in residualized sentence reading time; however, their changing reading times did not follow a clearly discernible pattern. Crucially, there was a suggestion of the Korean speakers' sensitivity to violations, which indicated that they possessed integrated, implicit

knowledge of the target syntax. The Korean experimental group tended to unconsciously slow down their reading during the violation block. Interestingly, the Korean experimental groups' grammatical sensitivity differed depending on the levels of awareness. The L1-aware participants, who later became aware of the cross-linguistic similarity between the target language and their L1, showed robust sensitivity to syntactically ungrammatical sentences in the input. The Korean VP-aware group showed a response pattern that was consistent with the L1-aware group. The Korean unaware group showed a qualitatively different response.

Offline measures. On the GJT, the Korean experimental group showed robust learning of the target grammar after incidental exposure through a novel. Furthermore, after two weeks, the Korean experimental group's knowledge was significantly enhanced, which prompted them to outperform the English experimental group on the delayed GJT. This enhanced knowledge consolidation was led by the Korean L1-aware group, whose members noticed during the immediate GJT (not during reading) that the target grammar matched their L1 grammar, Korean. Two weeks later, at delayed testing, these participants judged sentences with 87.22% accuracy, compared to 58.91% accuracy for the English speakers.

Notably, unlike the English experimental group, the Korean experimental group acquired not only explicit but also implicit knowledge. The Korean experimental group clearly demonstrated their advantage over the English experimental group in developing both implicit and explicit knowledge. This became apparent in the triangulation of GJT scores with the two offline awareness measures. First, a triangulation of GJT performance with retrospective verbal reports showed that the Korean unaware group developed more implicit knowledge than the English unaware group. Second, the analysis of GJT performance as a function of source attributions revealed that the Korean experimental group developed more explicit knowledge

than the English experimental group. Table 4.1 summarizes the findings from both the online and offline measures for the two participant groups.

Table 4.1

Summary of the findings

L1	Decrease in reading times	Power law of practice	Sensitivity	Grammaticality judgments	
				Implicit	Explicit
English	✓	✓	x	Δ	✓
Korean	✓	x	✓(L1-aware)	✓	✓

4.2 General Discussion

In the present dissertation, I observed stark overall differences between the English and the Korean experimental groups' developmental trajectories in an incidental language learning task. The effect of prior L1 knowledge on the acquisition of L2 syntax was pervasive throughout the experiment. During the exposure task (reading a novel), the Korean group exhibited a faster initial decrease in sentence-reading times than did the English group. Furthermore, only the Korean group showed grammatical sensitivity while reading as evidenced by increased sentence-reading times in response to syntactic violations. Subsequently, on the testing task (GJT), the Korean experimental group demonstrated reliable implicit learning and knowledge across time while the English experimental group did not. Moreover, the Korean experimental group acquired more explicit knowledge than did the English experimental group, largely through their awareness of the connection between their L1 and the target language. L1 awareness, the conscious perception of cross-similarity between the L1 (Korean) and the target Koreanish

syntax, contributed to memory consolidation over time. In this sense, the Korean experimental group consistently had advantages over the English experimental group.

In particular, the triangulation of diverse measures uncovered intriguing patterns of convergence as well as divergence in terms of how speakers with different L1 backgrounds interacted with the task. An interesting pattern of convergence concerns the online and offline knowledge measures. The Korean speakers had advantages over the English speakers on both the online violation block and the offline GJT. This pattern supported the idea that cross-language similarity boosts L2 learning (Ringbom, 2007). On the online knowledge measure, the Korean experimental group, especially those who later noticed the cross-language similarity between their L1 and the target language (L1-aware participants), showed evidence of grammatical sensitivity to violations, whereas the English experimental group did not. This may reflect the Korean experimental group's development of integrated, implicit knowledge of the target syntax. Even before awareness of the similarity between L1 and the target language emerged, the L1-aware group reliably showed evidence of implicit knowledge of the target word order. On the offline knowledge measure, the Korean experimental group outperformed the English experimental group, and it exhibited strong evidence of implicit knowledge while the English experimental group did not (as seen in the comparison of the English and Korean unaware learners' GJT performances as well as in the Korean group's above-chance performance on grammaticality judgments attributed to guess and intuition across testing sessions). Moreover, the L1-aware group excelled over all others and showed excellent performance on the GJT based primarily on their conscious knowledge of L1-L2 cross-linguistic similarities. In this way, both knowledge measures (sensitivity and GJTs) showed that the Korean L1-aware group led the way in the acquisition of the head-final target syntax. These converging findings from both the online

and offline measures provide compelling evidence of the significant L1 influence on the development of implicit syntactic knowledge under incidental exposure. The Korean speakers certainly had advantages in acquiring the target word order, which followed their L1 Korean, and the conscious registration of L1-awareness gave them an edge in knowledge development.

However, unlike the knowledge measures, the learning measure—the learning curve—revealed a different picture. That is, the learning and knowledge measures diverged in terms of L1 influence. Although both the Korean and English experimental groups sped up over time during reading the exposure text (i.e., evidence of learning), their learning curves had a different form. The English experimental group's reading data followed the power function (Anderson, 1993; Newell & Rosenbloom, 1981), while the Korean experimental group's reading data did not. This was quite an unexpected but very interesting diverging pattern. This divergence could be somewhat resolved if we posit that the English experimental group's power learning curve was a reflection of the English speakers' automatization of explicit knowledge of the target word order (DeKeyser, 2015), such as the VP rule. Although the English speakers' processing times followed the power function, it is plausible that incidental exposure through a novel was insufficient for them to integrate the target syntactic knowledge into their mental representation. Therefore, on the violation block at the end of the reading task, they were not sensitive to the word order violations. Similarly, they were unable to show evidence of implicit knowledge on the GJT; however, they reliably showed a fair amount of explicit knowledge on the GJT. Such evidence of explicit knowledge on the GJT may support the idea that the English speakers were automatizing declarative, explicit knowledge of Korean word order during learning. It may be that without the guidance of prior L1 knowledge, they needed more exposure and practice to develop unconscious grammatical reflexes to word order violations. Furthermore, one possible

explanation for the Korean speakers' reading data did not show a power function is that there was more heterogeneity in this group. However, this account must remain speculative for now. The subgroup analyses for the three awareness groups did not show marked differences between the subgroups, but there may be other sources of variation that account for the atypical learning curves. Future researchers could explore this possibility in greater depth.

Changes in the Korean group's sentence reading times showed a faster learning rate (steeper decrease) at the initial stage of reading. Moreover, their L2 development showed a progression from unconscious, implicit knowledge to conscious, explicit knowledge (Cleeremans, 2007, 2011; Karmiloff-Smith, 1992), which was seen in its full form in the L1-aware subgroup. Conversely, the English speakers, who did not have relevant prior knowledge, started off with explicit knowledge and fell short of achieving implicit knowledge (Jiang, 2007, 2011; Keating, 2009; Tokowicz & MacWhinney, 2005). Thus, the two L1 groups evidenced different learning trajectories, owing most likely to the prior L1 knowledge that they brought to the task.

The findings of this dissertation demonstrate the significant and durable effects of prior L1 knowledge on the incidental acquisition of L2 syntax. First, L1 syntactic transfer occurred both consciously and unconsciously under incidental learning conditions; indeed, 15 of the 24 Korean experimental participants remained unaware of the cross-language similarity in the experiment. Second, L1 syntactic transfer occurred at both online and the offline knowledge measures. Third, L1 awareness, which is the availability of conscious, metacognitive knowledge of the L1/L2 relationship, functioned as a catalyst for the development of implicit and explicit L2 syntactic knowledge. Thus, a reasonable conclusion of the present dissertation is that L2 learners with different L1 backgrounds perform on an unequal footing due to their prior L1 experience,

especially in naturalistic learning contexts. The two L1 groups, in fact, presented opposing directions of knowledge development. The Korean participants—who had relevant prior L1 knowledge—progressed from implicit to explicit knowledge under incidental exposure while the English participants—who lacked relevant prior L1 knowledge—began with explicit knowledge and failed to reach implicit knowledge.

4.3 Limitation

There were some limitations in the present study. First, the sample size for the awareness subgroups was quite small. For example, I could not have anticipated that there would be five participants in the Korean unaware subgroup. This could not be controlled for because I was not able to predict how many participants would become aware of the rules prior to the actual experiment. Therefore, future research with a larger number of participants would be needed to confirm the findings regarding the relationship between awareness and syntactic knowledge development. Second, the two L1 groups had different levels of familiarity with the vocabulary that made up the Koreanish text. English speakers were reading the Koreanish text in their L1, English, whereas the Korean-English bilinguals, though highly fluent in English, were reading the words in the text in their L2. This meant their speed of lexical processing was likely slower and some English vocabulary may have been unfamiliar to them. Any difficulties in comprehending the text, however, would have been relatively small, considering the two experimental groups performed on a par on the reading comprehension tests. Third, we cannot exclude the possibility of the materials effects for the reading task. Unlike sentence-processing experiments that allow tight control over the input, the order of the sentence presentation for the reading material in the study—the unmodified, authentic novel—was not random and the

possibilities for counterbalancing the text were limited. In other words, every participant read the sentences in a predetermined, fixed order during the exposure phase. Such fixed order and limited counterbalancing of the sentences could have introduced noise in the data that may have obscured some of the learning effects. Hence, the nonsignificant effect of power law and the flat slope for the log-log plot for the Korean speakers (see Figure 3.4) should be interpreted with caution. Given that increased noise levels may be a trade-off when using authentic materials, future research should attempt to seek an optimal balance between ecological validity and experimental control depending on its research goals.

4.4 Conclusion and Outlook

I introduced notable methodological advances in the present dissertation. First, I used an authentic novel to incidentally expose participants to the target semi-artificial language, creating a more ecologically valid research context. A coherent, longer text from the novel was likely to ensure a stronger participant focus on meaning, and it stimulated more natural reading processes than isolated sentences or unconnected short texts. Second, I employed both online and offline measures to track two L1 groups' learning trajectories at multiple stages: eye movements as an online measure of real-time processing and knowledge as well as grammaticality judgments as an offline knowledge measure. This allowed me to triangulate online measures derived from eye movements (residualized sentence-reading times and grammatical sensitivity) with offline knowledge measures (grammaticality judgments) and with awareness measures (source attributions, retrospective verbal reports). Third, I used delayed testing to track changes in knowledge development over time. These methodological components contributed to a comprehensive exploration of the pervasive and durable transfer effects of L1 in L2 syntax

acquisition. Overall, the L1 Korean (head-final) speakers had crucial advantages over the L1 English (head-initial) speakers when learning a new head-final language under meaning-focused, incidental learning conditions. L1 syntactic transfer occurred at both the conscious and unconscious levels, and L1 syntactic transfer occurred both online and in the offline knowledge measures. L1-awareness, defined as the availability of conscious, metacognitive knowledge of the L1-L2 relationship, aided the development of implicit and explicit L2 syntactic knowledge.

Despite the extensive amount of input, the English speakers, who lacked the guidance of cross-language similarity, did not show strong learning effects. In light of this, future researchers could incorporate different treatment conditions, such as enhancement and explicit instruction, to examine whether such treatments can diminish the L1 influence on the development of L2 syntactic knowledge (for example, see McManus & Marsden, 2019). The interaction between treatment conditions and L1 influence certainly merits further investigation. Moreover, in follow-studies, it would be interesting to examine individual differences in cognitive style and personality to see who becomes aware and who does not. Further, given that the present dissertation employed only receptive tasks as a knowledge measure, future research would benefit from using both receptive and productive tasks to expand the scope of observation concerning incidental learning. Another The in-depth investigation of the effect of prior L1 knowledge on implicit and explicit L2 learning will help researchers and educators both to better understand the variations in L2 acquisitional processes and to design tailored instructions for diverse learner populations.

APPENDICES

Appendix A: Sample of exposure task (novel reading)

Task instruction

Welcome to the study!

Thank you for your participation. In this experiment you will read a detective novel. The goal is to read this text as natural as possible.

You are going to read the first two chapters of the detective novel. For some of you, the sentences will be presented with the words scrambled. There will be practice screens at the beginning, which will be followed by a short break, and a break approximately every 15 screens afterwards. You will have six breaks throughout the experiment. You will be asked to answer simple comprehension questions at every other break. There will be 92 screens in total. To move from one screen to the next, press any button on the controller in front of you.

Try not to move your head during the experiment please.

Before each slide, you will be asked to look at a dot in the upper left-hand corner of your screen to continue. If you need to stop at any time during the experiment, please stop only when the dot is in the upper left-hand corner of the screen and notify the researcher.

Sample of the practice phase

There used to be intense public interest in what was known at the time as “The Styles Case,” but people have now begun to forget about it. Nevertheless, because of that previous interest, both by my friend Poirot and the family themselves asked me to write an account of the whole story. This, we all hope, will silence the sensational rumors that still persist. I will therefore briefly write of the circumstances which led to my being involved in the affair.

Sample of the training phase

I injury due to the battlefield from home had been sent. And, rather a depressing rehab facility in some months spending after, a month's sick leave was given. I no friends or family had, and I John Cavendish ran across when what do to my mind make up to was trying. I the past few years in him much had not seen. Indeed, I him particularly well never had known. He his forty-five years than younger looked though, he me than a good fifteen years was older. A boy as I Essex in his mother's house Styles at often had stayed.

Sample of the violation block

I was anxious John to get a hold of, but he was nowhere to be seen. Evidently had something very momentous that afternoon occurred. I tried the few words to forget I had overheard. But, no matter how much I tried, I could them not altogether from my mind dismiss. What was Mary Cavendish's concern in the matter? Mr. Inglethorp was in the parlor, when I to supper down came. His face was impassive as ever, and the strange unreality of the man struck me again. Mrs. Inglethorp came last down. She looked still agitated, and during the meal was there a somewhat awkward silence. As usual, surrounded he his wife with little attentions, placing her a cushion at back, and playing altogether the role of the devoted husband.

Appendix B: Sample of comprehension check questions

1. The narrator (Hastings) is recovering from an injury. (True False)
2. John Cavendish's family home is called Styles. (True False)
3. John Cavendish has a younger sister named Jessica. (True False)
4. John Cavendish's stepmother, Emily Inglethorp, remains single after his father's death. (True False)
5. Evie Howard is a delicate and pale lady. (True False)
6. John Cavendish's wife, Mary Cavendish, is a beautiful woman. (True False)
7. The group eats outside. (True False)
8. Alfred Inglethorp is a welcoming and charming host. (True False)

Appendix C: Sample of the grammaticality judgment test

The scrambling of the sentences in the novel (Task 1) was **not arbitrary but followed a complex system.** This was actually based on the grammar of a real language.

In the second part of the experiment, you will read 80 new sentences.

40 of these sentences were generated by means of the same language grammar as in the previous part of the experiment. The sentences follow the system presented in the novel. These sequences are called GRAMMATICAL.

The other 40 sentences were generated randomly, i.e., they do NOT confirm to the same language grammar presented in the novel. The sentences do not follow the system presented in the novel. These sequences are called UNGRAMMATICAL.

Your task is to decide which ones are grammatical and ungrammatical. Try to respond as quickly and accurately as possible. Simply pick whatever comes to your mind first.

In addition to deciding on whether each new sentence follows the complex system of the previous sentences, we will also ask you what the basis of your decision was.

* 1: GUESS
Your decision was based on a true guess, i.e., you might as well have flipped a coin.

* 2: INTUITION
Your decision was based on intuition, i.e., you feel that your decision is correct but you have no idea why.

* 3: RECOLLECTION
Your decision was based on the recollection of specific sentences (or parts of sentences) that you have read before that seemed similar; or you failed to recollect a specific sentence that was similar.

* 4: RULE KNOWLEDGE
Your decision was based on rule knowledge, i.e., you followed a rule when making the decision and you are able to describe the rule at the end of the experiment.

From now on, you will see two practice items with the researcher first. If you have any questions, please let the researcher know.

[Practice]

1. My student Spanish learned.

( ) grammatical ( ) ungrammatical

What is the basis of your decision? Please write down the number to give your response (1: Guess, 2: Intuition, 3: Recollection, 4: Rule knowledge)

_____

2. The teacher who asked many questions the student liked.

( ) grammatical ( ) ungrammatical

What is the basis of your decision? Please write down the number to give your response (1: Guess, 2: Intuition, 3: Recollection, 4: Rule knowledge)

_____

Now you are ready to start the main part of Task 2.

Your task is:

1) to decide whether it is grammatical (follows the word order system presented in Task 1) or ungrammatical (does NOT follow the word order system presented in Task 1),

2) and to report what the basis of your judgment was. Don't forget to leave your answer in the comments.

If you have any questions, please let the researcher know. If not, please proceed.

1. Susan the table touched.

( ) grammatical ( ) ungrammatical

What is the basis of your decision? Please write down the number to give your response (1: Guess, 2: Intuition, 3: Recollection, 4: Rule knowledge)

_____

2. Until Megan her driver's license renewed, she a bicycle rode.

( ) grammatical ( ) ungrammatical

What is the basis of your decision? Please write down the number to give your response (1: Guess, 2: Intuition, 3: Recollection, 4: Rule knowledge)

_____

3. Julie Mr. Lee the prize won that believed.

( ) grammatical ( ) ungrammatical

What is the basis of your decision? Please write down the number to give your response (1: Guess, 2: Intuition, 3: Recollection, 4: Rule knowledge)

_____

4. Betsy the coffeeshop at waited.

( ) grammatical ( ) ungrammatical

What is the basis of your decision? Please write down the number to give your response (1: Guess, 2: Intuition, 3: Recollection, 4: Rule knowledge)

_____

5. My father his driveway shoveled who the neighbor thanked.

( ) grammatical ( ) ungrammatical

What is the basis of your decision? Please write down the number to give your response (1: Guess, 2: Intuition, 3: Recollection, 4: Rule knowledge)

_____

Appendix D: Sample of the debriefing interview questions

[Session 1: short version of the debriefing interview]

1. While reading the novel, did you notice any rules or regularity? If yes, please indicate what you have noticed and during when you noticed it.

2. Please describe what you were thinking during reading.

3. Did you notice anything odd at the end of the novel? If yes, please indicate what you believe you have noticed.

[Session 2: long version of the debriefing interview]

1. Have you ever indicated rule knowledge as a source as a basis for your decisions? If so, please describe what you had been thinking. Why did you select rule knowledge category? If not, please share any other ways in which you made your choices.

2. As mentioned in the experiment, the scrambling of the sentences was not arbitrary. Instead, the word order in the sentences was based on a complex system. Reflecting now specifically on the placement of words within the sentences, can you recall any specific rule or regularity?

3. In Session 1, how did you read the sentences with the words scrambled? Please describe your own reading processes.

4. In Session 1, did you try to find a pattern from the scrambled word order during reading? Did you focus on the word order during reading? If so, please describe how you tried to find the pattern.

5. How much do you like reading a detective novel?

6. How many novels do you read a year?

7. How many Agatha Christie's books have you read?

8. Have you ever read the novel "The Mysterious Affair at Styles" before you participated the study?

REFERENCES

REFERENCES

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review*, *89*(4), 369.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological science*, *2*(6), 396-408.

Andringa, S. (in press, 2020). The emergence of awareness in uninstructed L2 learning: A visual world eye tracking study. *Second Language Research*.

Andringa, S. J., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning. *Studies in Second Language Acquisition*, 37, 185–196.

Batterink, L. J., Oudiette, D., Reber, P. J., & Paller, K. A. (2014). Sleep facilitates learning a new linguistic rule. *Neuropsychologia*, *65*, 169-179.

Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in second language acquisition*, *16*(2), 157-168.

Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D. J., & Pylkkänen, L. (2012). Syntactic structure building in the anterior temporal lobe during natural story listening. *Brain and language*, *120*(2), 163-173.

Brooks, P. J., & Kempe, V. (2013). Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Memory & cognition*, *41*(2), 281-296.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Christie, A. (1920). *The mysterious affair at Styles*. N.p.: John Lane. Retrieved on 11 October, 2017, from Project Gutenberg, www.gutenberg.org

Cintrón-Valentín, M., & Ellis, N. C. (2015). Exploring the interface: Explicit focus-on-form instruction and learned attentional biases in L2 Latin. *Studies in Second Language Acquisition*, *37*(2), 197-235.

Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied psycholinguistics*, *27*(1), 3-42.

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract and against stimulus-specific represntations. *Psychological Science*, *17*, 905–912.

Cop, U., Dirix, N., Drieghe, D., & Duyck, W. (2017). Presenting GECO: An eye-tracking corpus

of monolingual and bilingual sentence reading. *Behavioral Research, 49*, 602–615.

Cop, U., Dirix, N., Van Assche, E., Drieghe, D., & Duyck, W. (2017). Reading a book in one or two languages? An eye movement study of cognate facilitation in L1 and L2 reading. *Bilingualism: Language and Cognition*, *20*(4), 747–769.

Cop, U., Drieghe, D., & Duyck, W. (2015). Eye movement patterns in natural reading: A comparison of monolingual and bilingual reading of a novel. *PLoS ONE*, *10*, e0134008. doi:10.1371/journal.pone.0134008

Cop, U., Keuleers, E., Drieghe, D., & Duyck, W. (2015). Frequency effects in monolingual and bilingual natural reading. *Psychonomic Bulletin & Review*, *22*, 1216–1234.

Cleeremans, A. (2007). Consciousness: the radical plasticity thesis. *Progress in brain research*, *168*, 19-33.

Cleeremans, A. (2011). The radical plasticity thesis: how the brain learns to be conscious. *Frontiers in psychology*, *2*, 86.

Curcic, M., Andringa, S., & Kuiken, F. (2019). The role of awareness and cognitive aptitudes in L2 predictive language processing. *Language Learning*, *69*, 42-71.

Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 256-310). Oxford, UK: Blackwell Publishing.

DeKeyser, R. M. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in second language acquisition*, *19*(2), 195-221.

DeKeyser, R. (2007). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp. 97–112). Mahwah, NJ: Erlbaum.

Denhovska, N., Serratrice, L., & Payne, J. (2016). Acquisition of second language grammar under incidental learning conditions: The role of frequency and working memory. *Language Learning*, *66*(1), 159–190.

Denhovska, N., & Serratrice, L. (2017). Incidental learning of gender agreement. *Journal of Psycholinguistic Research*, *46*, 1187–1211.

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, *93*(5), 880-896. doi:10.1037/0022-351493.5.880

Dienes, Z. (2004). Assumptions of subjective meausres of unconscious mental states: Higher order thoughts and bias. *Journal of Consciousness Studies*, *11*, 25–45.

Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Progress in Brain Research*,

168, 49–64.

Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research, 69*, 338–351.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and brain sciences*, 22(5), 735-808.

Dussias, P. E. (2010). Uses of eye-tracking data in second language sentence processing research. *Annual Review of Applied Linguistics, 30,* 149-166.

Elgort, I., Brysbaert, M., Stevens, M., & Van Assche, E. (2018). Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition*, *40*(2), 341-366.

Elgort, I., & Warren, P. (2014). L2 vocabulary learning from reading: Explicit and tacit lexical knowledge and the role of learner and item variables. *Language Learning*, *64*(2), 365-414.

Ellis, N. C. (1996). Sequencing in SLA: Phonological Memory, Chunking and Points of Order. *Studies in Second Language Acquisition*, *18*, 91-126.

Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language learning*, *48*(4), 631-664.

Ellis, N. C. (2006). SLA: The Associative Cognitive CREED. In B. VanPatten & J. Williams; A. F. Williams (Eds.). *Theories in second language acquisition: An introduction*. Mahwah NJ: Erlbaum.

Ellis, N. C., & Schmidt, R. (1997). Morphology and longer distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition*, *19*(2), 145-171.

Ellis, N. C., & 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., & Sagarra, N. (2011). Blocking and learned attention in language acquisition: A replication and generalization study. *Studies in Second Language Acquisition*, *33*, 589–624.

Ellis, N. C., Hafeez, K., Martin, K. I., Chen, L., Boland, J., & Sagarra, N. (2014). An eye-tracking study of learned attention in second language acquisition. *Applied Psycholinguistics*, *35*(3), 547-579.

Fischer, S., Drosopoulos, S., Tsen, J., & Born, J. (2006). Implicit learning–explicit knowing: a role for sleep in memory system interaction. *Journal of Cognitive Neuroscience*, *18*(3), 311-319.

Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2: Electrophysiological evidence of the effect of L1–L2 syntactic similarity. *Bilingualism: Language and Cognition, 14*(3), 379-399.

Gass, S. (1979). Language transfer and universal grammatical relations. *Language learning, 29*(2), 327-344.

Gass, S. (1984). A review of interlanguage syntax: Language transfer and language universals. *Language learning*, *34*(2), 115-132.

Godfroid, A. (2016). The effects of implicit instruction on implicit and explicit knowledge development. *Studies in Second Language Acquisition, 38*(2), 177–215.

Godfroid, A. (2019). Investigating instructed second language acquisition using L2 learners' eye-tracking data (pp. 44-57). In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning*. New York: Routledge.

Godfroid, A. (in press, 2020). *Eye tracking in second language acquisition and bilingualism: A research synthesis and methodological guide*. New York: Routledge.

Godfroid, A., Ahn, J., Choi, I., Ballard, L., Cui, Y, Johnston, S., Lee, S., Sarkar, A., & Yoon, H. (2018). Incidental vocabulary learning in a natural reading context: an eye-tracking study. *Bilingualism: Language and Cognition, 21*(3), 563-584.

Godfroid, A., & Spino, L. A. (2015). Reconceptualizing reactivity of think-alouds and eye tracking: Absence of evidence is not evidence of absence. *Language Learning*, *65*(4), 896-928.

Godfroid, A., & Winke, P. (2015). Investigating implicit and explicit processing using L2 learners' eye-movement data. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 325-348). Philadelphia, PA: John Benjamins.

Grey, S., Williams, J. N., & Rebuschat, P. (2014). Incidental exposure and L2 learning of morphosyntax. *Studies in Second Language Acquisition, 36*, 1–35.

Hama, M., & Leow, R. P. (2010). Learning without awareness revisited: Extending Williams (2005). *Studies in Second Language Acquisition, 32*, 465–491.

Hamrick, P. (2013). Development of conscious knowledge during early incidental learning of L2 syntax. *Georgetown University, Washington, DC.*

Hamrick, P., & Sachs, R. (2017). Establishing evidence of learning in experiments employing artificial linguistic system. *Studies in Second Language Acquisition,* 1–17.

Hatzidaki, A., Branigan, H. P., & Pickering, M. J. (2011). Co-activation of syntax in bilingual language production. *Cognitive psychology*, *62*(2), 123-150.

Indrarathne, B., & Kormos, J. (2017). Attentional processing of input in explicit and implicit conditions: An eye-tracking study. *Studies in Second Language Studies, 39,* 401–430.

Indrarathne, B., Ratajczak, M., & Kormos, J. (2018). Modelling changes in the cognitive processing of grammar in implicit and explicit learning conditions: Insights from an eye-tracking study. *Language Learning*, *68*(3), 669-708.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge.

Jiang, N. (2007). Selective integration of linguistics knowledge in adult second language learning. *Language Learning*, *25*, 603-634.

Jiang, N., Novokshanova, E., Masuda, K., & Wang, X. (2011). Morphological congruency and the acquisition of L2 morphemes. *Language Learning*, *61*(3), 940-967.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Keating, G. D. (2009). Sensitivity to violations of gender agreement in native and nonnative Spanish: An eye-movement investigation. *Language Learning, 59*(3), 503-535.

Kerz, E., Wiechmann, D., & Riedel, F. B. (2017). Implicit learning in the crowd: Investigating the role of awareness in the acquisition of L2 knowledge. *Studies in Second Language Acquisition*, *39*(4), 711-734.

Kim, M., & Godfroid, A. (in press, 2019). Should we listen or read? Modality effects in implicit and explicit L2 knowledge of grammar. *The Modern Language Journal*, *103*(3).

Kuhn, G., & Dienes, Z. (2008). Learning non-local dependencies. *Cognition*, *106*(1), 184-206.

Leow, R. P., Grey, S., Marijuan, S., & Moorman, C. (2014). Concurrent data elicitation procedures, processes, and the early stages of L2 learning: A critical overview. *Second Language Research*, *30*(2), 111-127.

Leow, R. P. (2015a). *Explicit learning in the L2 classroom: A student-centered approach*. New York: Routledge.

Leow, R. P. (2015b). The roles of attention and (un) awareness in SLA: Conceptual replication of N. C. Ellis & Sagarra (2010a) and Leung & Williams (2012). *Language Teaching*, *48*(1), 117-129.

Leung, J. H. C., & Williams, J. N. (2011). The implicit learning of mappings between forms and contextually-derived meanings. *Studies in Second Language Acquisition*, *33*, 33–55.

Leung, J. H. C., & Williams, J. N. (2012). Constraints on implicit learning of grammatical form-

meaning connections. *Language Learning, 62*, 634–662.

Leung, J. H. C., & Williams, J. N. (2014). Crosslinguistic differences in implicit language learning. *Studies in Second Language Acquisition*, 36, 634–662.

Lim, J. H., & Christianson, K. (2015). Second language sensitivity to agreement errors: Evidence from eye movements during comprehension and translation. *Applied Psycholinguistics*, *36*(6), 1283-1315.

MacWhinney, B. (1997). Second language acquisition and the competition model. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 113–144). Mahwah, NJ: Erlbaum

MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. de Groot (eds.). *Handbook of bilingualism: Psycholinguistic approaches*. Oxford: Oxford University Press.

MacWhinney, B. (2008). A unified model. In N. Ellis & P. Robinson (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341–371). London, UK: Routledge.

McManus, K. & Marsden E. (2019). Signatures of automaticity during practice. Explicit instruction about L1 processing routines can improve L2 grammatical processing. *Applied Psycholinguistics*, *40*(1), 205-234.

Miller, Z., & Godfroid, A. (in press, 2019). Emotions in incidental language learning: An individual differences approach. *Studies in Second Language Acquisition.*

Mohamed, A. A. (2018). Exposure frequency in L2 reading: An eye-movement perspective of incidental vocabulary learning. *Studies in Second Language Acquisition, 40*(2), 269-293.

Monaghan, P., Schoetensack, C., & Rebuschat, P. (2019). A single paradigm for implicit and statistical learning. *Topics in Cognitive Science*.

Morgan-Short, K., Deng, Z., Brill-Schuetz, K. A., Faretta-Stutenberg, M., Wong, P. C., & Wong, F. C. (2015). A view of the neural representation of second language syntax through artificial language learning under implicit contexts of exposure. *Studies in Second Language Acquisition*, *37*(2), 383–419.

Morgan-Short, K., Finger, I., Grey, S., & Ullman, M. T. (2012). Second language processing shows increased native-like neural responses after months of no exposure. *PLoS One, 7*(3), e32974.

Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T. (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-

related potential study. *Language Learning*, *60*, 154–193.

Nijhof, A. D., & Willems, R. M. (2015). Simulating fiction: individual differences in literature comprehension revealed with fMRI. *PLoS One*, *10*(2), e0116492.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language learning*, *50*(3), 417-528.

Onnis, L., & Thiessen, E. (2013). Language experience changes subsequent learning. *Cognition*, *126*, 268–284.

Pacini, R., & Epstein, S. (1999). The relation of rational and experimental information processing styles to personality, basic beliefs, and the ratio-bias phenomenon. *Journal of Personality and Social Psychology*, *76*(6), 972-987.

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental Vocabulary Acquisition from an Authentic Novel: Do" Things Fall Apart"?. *Reading in a Foreign Language*, *22*(1), 31-55.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238.

Plihal, W., & Born, J. (1997). Effects of early and late nocturnal sleep on declarative and procedural memory. *Journal of cognitive neuroscience*, *9*(4), 534-547.

Rayner, K. (1998). Eve movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422. DOI: 10.1037/0033-2909.124.3.372

Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, 63, 595–626.

Rebuschat, P., Hamrick, p., Riestenber, K., Sacks, R., & Ziegler, N. (2015). Triangulating measures of awareness: A contribution to the debate on learning without awareness. *Studies in Second Language Acquisition*, 37, 299–334.

Rebuschat, P., & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics*, *33*, 829–856.

Rebuschat, P., & Williams, J. N. (2012). Introduction: Statistical learning and language acquisition. In P. Rebuschat, & J. N. Williams (Eds.), *Statistical learning and language acquisition.* (pp. 1-12). Berlin: Mouton de Gruyter.

Ringbom, H. (2007). *Cross-linguistic similarity in foreign language learning*. Clevedon: Multilingual Matter.

Robinson, P. (2002). Learning conditions, aptitude complexes, and SLA. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 113-133). Philadelphia, PA: Benjamins.

Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 leaning: 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.

Rogers, J., Reves, A., & Rebuschat, P. (2016). Implicit and explicit knowledge of inflectional morphology. *Applied psycholinguistics*, 37, 781–812.

Sanoudaki, E., & Thierry, G. (2015). Language non-selective syntactic activation in early bilinguals: the role of verbal fluency. *International Journal of Bilingual Education and Bilingualism*, *18*(5), 548-560.

Sanz, C., & Grey, S. (2015). Effects of conditions on L2 development: Moving beyond accuracy. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 301-324). Amsterdam: John Benjamins.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language learning*, *60*(2), 263-308.

Spinner, P., Foote, R., & Upor, R. A. (2017). Gender and number processing in second language Swahili. *Linguistic Approaches to Bilingualism*, *8*(4), 446-476.

Spinner, P., & Gass, S. M. (2019). *Using judgments in second language acquisition research*. Routledge: New York.

Spinner, P., Gass, S. M., & Behney, J. (2013). Ecological validity in eye-tracking: An empirical study. *Studies in Second Language Acquisition*, *35*(2), 389-415.

Spinner, P., & Jung, S. (2018). Production and Comprehension in Processability Theory: A self-Paced Reading Study. *Studies in Second Language Acquisition*, *40*(2), 295-318.

Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics*, *38*(5), 1229-1261.

Suzuki, Y., & DeKeyser, R. M. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, *65*(4), 860-895.

Suzuki, Y., & DeKeyser, R. M. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, *67*(4), 747-490.

Speer, N. K., Reynolds, J. R., Swallow, K. M., & Zacks, J. M. (2009). Reading stories activates neural representations of visual and motor experiences. *Psychological science*, *20*(8), 989-999.

Tagarelli, K., Ruiz, S., Vega, J., & Rebuschat, P. (2016). Variability in second language learning: The roles of individual differences, learning conditions, and linguistic complexity. *Studies in Second Language Acquisition*, *38*(2), 293–316.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in second language acquisition*, *27*(2), 173-204.

Tolentino, L. C., & Tokowicz, N. (2011). Across language, space, and time. *Studies in Second Language Acquisition*, *33*, 91-125.

Tolentino, L. C., & Tokowicz, N. (2014). Cross-language similarity modulates effectiveness of second language grammar instruction. *Language Learning*, *64*(2), 279-309.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 141–178). Washington, DC: Georgetown University Press.

Vafaee, P., Suzuki, Y., & Kachisnke, I. (2017). Validating grammatical judgment tests. *Studies in Second Language Studies*, 39(1), 59-95.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176-1190.

VanPatten, B., & Smith, M. (2015). Aptitude as grammatical sensitivity and the initial stages of learning Japanese as a L2. *Studies in Second Language Acquisition*, *37*, 135-165.

VanPatten, B., & Rothman, J. (2015). What does current generative theory have to say about the explicit-implicit debate? In P. Rebuschat (Ed.), *Explicit and implicit learning of languages* (pp. 89-116). Amsterdam: John Benjamins.

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature*, *427*(6972), 352.

White, L. (2000). Second language acquisition: from initial to final state. In J. Archibald (ed.), *Second language acquisition and linguistic theory* (pp. 130–155). Blackwell.

Whiteside, S. P., & Lynam, D. R. (2001). The five factor model and impulsivity: Using a structural model of personality to understand impulsivity. *Personality and Individual Differences*, *30*(4), 669-689. doi:10.1016/S0191-8869(00)00064-7.

Willems, R. M., Frank, S. L., Nijhof, A. D., Hagoort, P., & Van den Bosch, A. (2015). Prediction during natural language comprehension. *Cerebral Cortex*, *26*(6), 2506-2516.

Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition*, *27*, 269–304.

Williams, J. N. (2009). Implicit learning in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 319–353). Bingley, UK: Emerald Press.

Williams, J. N., & Lovatt, P. (2003). Phonological memory and rule learning. *Language learning*, *53*(1), 67-121.