# CROSS-LINGUISTIC ACTIVATION IN KOREAN L2 LEARNERS' PROCESSING OF REMENTION BIAS IN ENGLISH

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF
THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

SECOND LANGUAGE STUDIES

DECEMBER 2018

By

Hyunwoo Kim

Dissertation Committee:

Theres Grüter, Chairperson
Kamil Ud Deen
William O'Grady
Bonnie D. Schwartz
Shinichiro Fukuda

Keywords: remention bias, cross-language activation, causative marking, reference processing

# ACKNOWLEDGMENTS

Among all those people who have contributed to this dissertation, I would like to first give my deepest thanks to my advisor, Theres Grüter. Her guidance, valuable advice, and encouragements made possible much of what I could do during my doctoral program, including this dissertation. Whenever I felt lost or discouraged, she always provided me the strength to move forward and supported me in the best way she could. Without her help, I would never have completed the research. It was a great blessing to have her as my advisor.

I also would like to express my deepest gratitude to my committee members, Kamil Ud Deen, Shinichiro Fukuda, William O'Grady, and Bonnie D. Schwartz for their support and helpful comments on the dissertation. Kamil provided insightful feedback on my research and was willing to help me by joining my committee when I needed him. Fukuda sensei shared his expertise of Japanese linguistics, which greatly inspired me when I developed ideas regarding causative constructions in Korean. William not only encouraged me with his father-like caring, but also taught me a broad range of information on Korean syntax. It was Bonnie who constantly motivated me with her keen insight and vast knowledge. Special thanks also go to Amy Schafer for encouraging me to develop the remention bias study that was initially a final project for her class and for teaching me various aspects of psycholinguistics.

I owe much to other faculty members at Seoul National University. Professor Hyun-Kwon Yang, who supervised my master's thesis in Korea, greatly influenced me in terms of both my academic career and view of life. I also thank Professor Oryang Kwon, Professor Moonsu Shin, Professor Jin-Wan Kim, Professor Hyunkee Ahn, Professor Byungmin Lee, Professor Sun-Young Oh, and Professor Byungcheon Min for their support and helpful encouragement.

I also want to thank many people I met in Hawai'i and in Seoul for their support: Hyunah Ahn, Amber B. Camp, Ivan Bondoc, Yunchuan Chen, Minyoung Cho, Keiko Hata, Haerim Hwang, Kum-Jeong Joo, Hanbyul Jung, Hye Young Jung, Hong Ri Kim, Jonny Kim, Kitaek Kim, Rakhun Kim, Sangki Kim, Yeonju Kim, Heejung Lee, Josephine Lee, Wenyi Ling, Mari Miyao, Eunseok Ro, Gyu-Ho Shin, Minchang Sung, Aya Takeda, Nozomi Tanaka, Zhijun Wen, Sejung Yang, Fred Zenker, and Jing Crystal Zhong. In addition, I sincerely thank Professor Hye-ryeong Hahn, Professor HyunSook Ko, Professor Nayoung Kwon, Professor Jin-Hwa Lee,

Professor Sun Hee Park, Professor Yangon Rah, Professor Sang-Keun Shin, and Dr. Jai-Ho Yoo for their insights and warm-hearted support. I also thank Pastor Junsang Kang and members at Logos Presbyterian Church for their support both materially and spiritually. I am also grateful to Bonnie Fox and Ivana Matson for English annotations, Jihyun Kim and Gyu-ho Shin for Korean annotations, and Laura Ahn, Ann Im, Jihyun Kim, and Heejung Seo for translation.

Finally, I would like to thank National Science Foundation for providing me with a doctoral dissertation research grant (BCS-1749240, PI: Theres Grüter) and the Fulbright Commission of Korea and the United States for supporting my doctoral program (IIE Grantee ID: 15130996).

My wholehearted thanks and love go to my wife Jihyun Kim for supporting me in all ways and putting up with hardship during our stay in the U.S. I also want to thank my parents and parents-in-law for their support.

## ABSTRACT

Effects of cross-linguistic activation in L2 learners have been demonstrated abundantly at the word level (e.g., Dijkstra, 2005; Prior, Degani, Awawdy, Yassin & Korem, 2017; Van Assche, Duyck & Brysbaert, 2013), yet less is known about the consequences of cross-language activation at the word level on processing at the sentence and discourse levels. This dissertation investigates if and how Korean learners of English are affected by the strength of referential biases associated with certain interpersonal predicates in Korean in their reference choices and processing in English. The study also tests potentially modulating roles of translation priming, L2 proficiency, and L2 learning experience, whose effects on cross-language activation remain an issue of ongoing investigation.

Remention bias is a well-known phenomenon whereby certain verbs appear to create biases to remention either its subject or its object in a causal dependent clause (Garvey & Caramazza, 1974; Hartshorne, 2014). Importantly, some English remention bias verbs have no lexical translation equivalents in Korean and can only be translated with a periphrastic construction involving explicit marking of causality (e.g., 'surprise'→*nolla-key ha*, be surprised-RESULT do). Experiment 1 tested whether such predicates in Korean, which contain explicit causality marking, lead to stronger remention biases than predicates with no causality marking. Results from written sentence-completion tasks in Korean and English showed stronger subject bias with predicates with causality marking than predicates with no causality marking among native Korean speakers, as well as similar biases for the English translation equivalents of these predicates among native English speakers. Experiment 2 further explored whether the stronger bias with predicates encoding explicit vs. implicit causality in Korean affects Korean-speaking L2 learners' sentence completions in English. The study also probed for potential effects of translation priming by having L2 learners complete a translation task either preceding or following the sentence-completion task. Results indicated that the strength of a verb's referential bias in Korean affected learners' reference choices in English. This effect emerged independent of the presence or absence of translation priming. Experiment 3 tested whether the results from Experiment 2 could be replicated with a different set of remention bias verbs with more uniform argument structures. The results of Experiment 3 not only replicated the effects of cross-linguistic activation in L2

referential choices but also showed that these effects emerged regardless of learners' proficiency or learning experience (immersed vs. instructed). Based on the effect of cross-linguistic activation in L2 learners' (offline) referential choices observed in Experiments 2 and 3, Experiment 4 used the visual world eye-tracking paradigm to investigate whether the effect extends to online processing. Results showed that while L2 learners used remention bias information during real-time listening, their use of the information was delayed compared to that of native speakers. Yet no robust evidence was found that either proficiency or cross-linguistic activation interacted with L2 learners' use of remention bias.

Overall, the results from this study indicate that the effect of cross-linguistic activation goes beyond the word or construction level and influences Korean-speaking L2 learners' referential choices at a discourse level. These effects were robust and replicable in two offline tasks, and emerged irrespective of the presence of translation priming, L2 proficiency, and L2 learning experience. These effects are assumed to arise through the mental models created under the influence of cross-linguistic activation at the word and construction level during L2 learners' production of written discourse continuations. In the visual world eye-tracking task, by contrast, no clear effects of cross-linguistic activation emerged, potentially due to L2 listeners' delayed use of remention bias in real-time processing.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

The following abbreviations are used to label the linguistic terms employed in this dissertation.

| | |
|---|---|
| ACC | Accusative case particle |
| COMP | Complementizer |
| DAT | Dative case particle |
| DECL | Declarative sentence-type suffix |
| NOM | Nominative case particle |
| PAST | Past tense suffix |
| PL | Plural suffix |

# CHAPTER I
# INTRODUCTION

Language users rely on various linguistic and non-linguistic cues to interpret a referring expression in a discourse context. While some of these cues come from explicit linguistic devices such as number (e.g., *it, they*), case marking (e.g., *he*, *him*), and gender marking (e.g., *he, she*), others are less explicit. Consider (1), for example.

(1) Tom hated Bill because he…

Several implicit cues may guide a reader in the interpretation of the ambiguous pronoun *he* in (1). First, a reader may resolve the pronoun *he* to the subject *Tom*, affected by a preference to disambiguate a pronoun to the subject of the preceding clause, a highly accessible antecedent in the discourse (Arnold, 2010; Hobbs, 1979). Also, the reader may disambiguate the reference of *he* as referring to *Tom* based on a first-mention strategy, a preference to associate a pronoun with a first-mentioned referent in the previous clause (e.g., Crawley, Stevenson, & Kleinman, 1990; Gernsbacher, Hargreaves, & Beeman, 1989). Another cue that the reader may deploy during the interpretation of a pronoun is a grammatical role preference or parallel processing strategy (Sheldon, 1974), whereby an antecedent that shares the same grammatical role with a given pronoun is preferred. In this example, the reader may resolve the pronoun to *Tom* since both arguments, *he* and *Tom*, bear the subject role in the subordinate and matrix clause, respectively. Alternatively, the reader may choose the object *Bill* as the antecedent of the pronoun because it is the most recently processed referent and thus remains highly accessible in the reader's memory (Arnold, 1998; Gernsbacher et al., 1989). Last but not least, the reader may have a preference for *Bill* as the antecedent of *he* because the verb *hate* evokes an event that calls upon a particular explanation as to what aspects or action of *Bill* caused *Tom* to hate him. Like the verb *hate*, some interpersonal verbs have been shown to induce biases to rementation either its subject or object in a causal dependent clause when followed by an explanation. In the psychology and psycholinguistics literature, this phenomenon has been called *IMPLICIT CAUSALITY* (Brown &

Fish 1983; Garvey & Caramazza 1974; Au, 1986) or *REMENTION BIAS* (Hartshorne, 2014).[1] This dissertation focuses on remention bias as a tool to investigate potential effects of cross-linguistic activation on referential interpretation in second language (L2) sentence and discourse processing..

Remention biases are indicated by a comprehender's preference to remention one of the verb's arguments as the likely cause of the event when asked to continue the sequence [Argument–Verb–Argument *because*], as in (2).

(2a) John frightened Peter because …

(2b) Mary feared Nancy because ...

The sentence fragments in (2) provide no explicit information as to which protagonist in the main clause is more likely to cause the event. Nevertheless, people tend to have a bias toward certain interpretations depending on the verb semantics, which allow them to derive probabilistic inferences about who is more responsible for the event. In case of the *frighten* event in (2a), *John* – the subject of the main clause – appears more likely to be the cause of the event than *Peter*, which gives rise to greater bias to remention the subject referent in the ensuing causal dependent clause as a causer. In the *fear* event in (2b), on the other hand, *Nancy* – the object of the main clause – seems the more likely cause of the event, consequently inducing a stronger preference to be mentioned in the following clause. Verbs with a bias toward the subject as the underlying cause of the event, such as *frighten*, are referred to as 'subject-biased verbs' or 'NP1-biased verbs', and verbs with a bias toward object, such as *fear*, are referred to as 'object-biased verbs' or 'NP2-biased verbs'.

Numerous studies have provided robust evidence that monolingual or first language (L1) speakers have consistent remention biases when processing sentences such as (2) (e.g., Cozijn, Commandeur, Vonk, & Noordman, 2011; Ferstl, Garnham & Manouilidou, 2011; Hartshorne &

[1] The term *implicit causality bias* focuses on pronoun resolution, whereas the term *remention bias* is more comprehensive, covering a wider range of referential possibilities including interpretations of overt pronouns, null subjects and repeated names (Hartshorne, 2014). Since this study investigates various types of referential expressions, not limited to pronouns, *remention bias* rather than *implicit causality bias* will be used throughout this dissertation.

Snedeker, 2013; Hartshorne, Sudo & Uruwashi, 2013; Itzhak & Baum, 2015; Pyykkönen & Järvikivi, 2010; Stewart, Pickering & Sanford, 2000). Relative to the extensive literature on L1 speakers' use of remention bias in reference interpretation and processing, much less is known about how this information is utilized by L2 speakers. To the best of my knowledge, only four studies have so far examined this issue (Cheng & Almor, 2017, 2018; Contemori & Dussias, 2018; Liu & Nicol, 2010), with findings pointing to the conclusion that L2 learners are generally able to use verb-related bias information, yet not as much as L1 speakers do. However, none of these studies has considered the potential role of cross-linguistic influence, L2 proficiency, or language learning experience – some important areas that are known to affect L2 sentence and discourse processing (Kaan, 2014) – on L2 speakers' processing of remention biases.

Investigating the L2 processing of remention biases and the potential influence of cross-linguistic activation, L2 proficiency, and language learning experience allows for testing how L2 learners differ from native speakers in their use of remention bias information during incremental processing. In the L1 processing literature, it is well-attested that native speakers actively recruit remention bias information for formulating expectations about who will be mentioned in the upcoming discourse (e.g., Cozijn et al., 2010; Itzhak & Baum, 2015; Pyykkönen & Järvikivi, 2010). In the L2 processing literature, however, L2 use of remention bias information during incremental processing has received little attention. To the best of my knowledge, only one study has tested L2 processing of remention bias during real-time listening (Contemori & Dussias, 2018), focusing on highly advanced early bilinguals. Thus, a number of unresolved issues remain regarding potential factors influencing L2 processing of remention bias. For example, Kaan (2014) pointed out that L2 learners may be restricted in their ability to use linguistic information for predictive processing compared to L1 speakers due to several factors such as less stable lexical representations, cross-language co-activation, L2 proficiency, and/or language learning experience (i.e., immersive vs. non-immersive experience). This dissertation attempts to address these gaps by exploring the influence of these factors on the processing of remention biases in L2.

Examining L2 processing of remention bias information also contributes to the investigation of potential differences between L1 and L2 processing at the discourse level, where comprehenders draw on remention bias in the construal of causal relations across clauses. Whereas many studies have compared L1 and L2 processing at the sentence level (e.g., Clahsen

& Felser, 2006; Dussias, 2003; Dussias & Piñar, 2010; Dussias & Sagarra, 2007; Frenck-Mestre, 2002; Hopp, 2016; Omaki & Schulz, 2011; Pliatsikas & Marinis, 2013; Witzel, Witzel & Nicol, 2012), fewer studies have probed discourse-level processing in L2, with findings remaining somewhat inconclusive (e.g., Foucart, Romero-Rivas, Gort & Costa, 2016; Grüter, Rohde & Schafer, 2017; Pan, Schimke, & Felser, 2015; Roberts, Gullberg & Indefrey, 2008; Trenkic, Mirkovic & Altmann, 2014). While some studies showed that L2 learners strongly rely on discourse cues to make inferences about inter-clausal relations (e.g., Foucart et al., 2016) or resolve reference and syntactic ambiguities (e.g., Trenkic et al., 2014), others suggest that L2 learners have a reduced ability to generate discourse-level expectations (e.g., Grüter et al., 2017) or to use discourse information for reference resolution (e.g., Roberts et al., 2008). In this regard, exploring L2 processing of remention bias information adds to the debate regarding whether and how L2 processing differs from L1 processing at the discourse level.

Motivated by the debate about differences between L1 and L2 sentence and discourse processing and the potential influence of cross-language activation, proficiency, and language learning experience, this dissertation investigates L2 learners' use of remention bias information. This study focuses particularly on how each of the potentially modulating factors – cross-language activation, L2 proficiency, and L2 learning experience – plays a role in L2 learners' processing of remention biases. For the role of cross-language activation, I investigate if and how Korean-speaking learners of English are affected by parallel access to English verbs and their Korean translation counterparts when they construe causality in the comprehension and production of English discourse. Specifically, I focus on the effect of cross-linguistic activation at word and construction levels on referential choices and processing in causal dependent clauses where the dependent clause provides an explanation for the event in the matrix clause (e.g., *Eliza surprised Natalie because she was hiding and then popped out*). I will follow the assumption that biases to remention an event participant (e.g., *Eliza* or *Natalie*) from the matrix clause in a causal dependent clause are closely associated with the matrix verb's semantic structure (Hartshorne & Snedeker, 2013). As discussed in more detail in Chapter 2, some English predicates have different syntactic/semantic structures from their Korean translation equivalents. Specifically, some NP1-biased verbs in English can only be translated into Korean predicates that contain a construction including explicit causative marking. In the context of constructions in Japanese

that also contain explicit marking of causation, Hartshorne et al. (2013) hypothesized that explicit marking of causality may give rise to stronger remention biases than implicit causality in the lexical verbs typically examined in research on remention biases. In an offline sentence-completion study (Experiment 1), I will test Hartshorne et al.'s hypothesis by investigating whether explicit causative marking in Korean predicates gives rise to cross-linguistic differences between Korean (Experiment 1a) and English (Experiment 1b) in the strength of biases for rementioning one of the event participants in a causal dependent clause. I will then test whether these cross-linguistic differences give rise to cross-language activation that affects Korean-speaking learners of English in their offline referential choices in English (Experiment 2). In subsequent experiments, I will test effects of cross-language activation in offline (Experiment 3) and online (Experiment 4) referential processing in English, with remention bias verbs selected from consistent VerbNet classes (Kipper, Korhonen, Ryant & Palmer, 2008). In addition to the effect of cross-linguistic influence, Experiment 3 will investigate to what extent L2 proficiency and learning experience modulate the effect of cross-linguistic activation. In Experiment 4, I will explore whether L2 learners can make online use of remention bias information, whether cross-linguistic activation influences L2 learners' online referential processing, and how L2 proficiency modulates their use of remention biases and the effect of cross-linguistic activation during real-time listening.

The rest of the dissertation is organized as follows. Chapter 2 provides a review of the L1 and L2 literatures on remention bias, including discussion of cross-linguistic differences between Korean and English remention bias predicates. This chapter also reviews relevant previous literature on cross-linguistic activation and the roles of proficiency and learning experience in L2 processing, and it concludes with the research questions and predictions for the experiments presented in this dissertation. Chapter 3 reports two parallel sentence-completion experiments with L1 speakers of Korean and L1 speakers of English. Chapter 4 reports an English sentence-completion experiment with L1 speakers of English and L1 Korean learners of L2 English. Chapter 5 reports the results from an English sentence-completion experiment, with remention bias verbs selected from consistent semantic classes and additional factors of L2 proficiency and L2 learning experience. Chapter 6 discusses a visual-world eye-tracking experiment, which tested L2 learners' use of remention bias information and effects of cross-linguistic influence.

Lastly, Chapter 7 concludes the dissertation with a summary and discussion of the findings and suggestions for future research.

# CHAPTER II
# LITERATURE REVIEW

## 2.1 Remention biases in causal dependent clauses

A verb-mediated remention bias is a well-attested factor in L1 reference resolution and processing. Despite extensive evidence that native speakers use this information consistently during the comprehension of sentences such as (2) (repeated in (3) below) (e.g., Au, 1986; Bott & Solstad, 2014; Brown & Fish, 1983; Caramazza, Grober, Garvey & Yates, 1977; Cozijn et al., 2011; Ferstl et al., 2011; Hartshorne & Snedeker, 2013; Itzhak & Baum, 2015; Pyykkönen & Järvikivi, 2010; Stewart et al., 2000), the precise mechanisms underlying this phenomenon still remain only partially understood (Hartshorne & Snedeker, 2013).

(3a) John frightened Peter because …

(3b) Mary feared Nancy because ...

While there is a general consensus that remention biases are associated with properties of verbs, researchers have long debated what aspects of a verb contribute to the biases.[2] So far, three main accounts have been proposed regarding this issue: the arbitrary semantic tag account, the world knowledge account, and the semantic structure account.

The arbitrary semantic tag account claims that each verb carries a unique feature associated with a remention bias (e.g., Caramazza et al., 1977; Garvey & Caramazza, 1974). According to this account, whether a verb has a subject- or object-oriented bias is arbitrary, just as grammatical gender is an arbitrary feature encoded on nouns in languages like Spanish. On this

[2] In addition to the verb semantics, the type of connector that introduces a subordinate clause also plays an important role in inducing remention biases. Numerous studies have shown that the strength and direction of remention biases drastically change when "because" is replaced with other connectors, such as "and so" and "but," indicating a close association between the activation of remention bias and the coherence relation associated with a connector. In particular, the type of bias induced by the connector "and so" is called implicit consequentiality (Crinean & Garnham, 2006; Stewart et al., 1998), as this connector preferentially elicits the description of consequences, not causes, of the event in the main clause. Since this dissertation focuses exclusively on implicit causality and sentences containing the connector *because*, effects of different connectors in contributing to remention biases will not be discussed further here.

account, the strength and direction of a verb's remention bias cannot be predicted by any properties of the verb, including its syntactic and semantic features.

The world knowledge account attributes a verb's remention bias to the language user's knowledge about the situation in which the event takes place (e.g., Brown & Fish, 1983; Corrigan, 2001; Pickering & Majid, 2007). This account focuses on nonlinguistic knowledge such as the social status of event participants in a sentence, rather than verbal meaning itself, as a main source of remention biases. Like the arbitrary semantic tag account, this perspective does not assume any predictive role of a verb's syntactic and semantic structures in accounting for remention biases.

Recent studies have provided evidence against the arbitrary tag and world knowledge accounts, demonstrating that a verb's remention bias may be explained, at least in part, by the verb semantics. This so-called semantic structure account proposes that remention biases are systematically associated with a verb's semantic structure, namely the lexical content of the verb and its arguments' thematic roles (Bott & Solstad, 2014; Hartshorne, O'Donnell, & Tenenbaum, 2015; Hartshorne & Snedeker, 2013). This proposal is best illustrated by Hartshorne and Snedeker (2013), who classified 328 monosemic verbs into classes according to VerbNet (Kipper et al., 2008), an extended version of Levin's (1993) taxonomy of verb argument structure, and tested whether verbs in the same class (i.e., verbs that involve the same thematic role structures) would consistently exhibit the same direction of remention biases. The basic assumption of this semantics-driven account is that since semantic constructs such as thematic roles constitute major components of the meaning of an event denoted by the verb, certain classes of verbs with the same thematic properties will induce the same direction of the remention bias. In support of this hypothesis, Hartshorne and Snedeker's (2013) analyses revealed that five classes – class 31.1 (Stimulus-Experiencer verbs, e.g., *frighten, surprise*), class 31.2 (Experiencer-Stimulus verbs, e.g., *admire, love*), class 33 (Judgment verbs, e.g., *blame, thank*), class 45.4 (Other Alternating Verbs of Change of State, e.g., *improve, revive*), class 59 (Force verbs, e.g., *dare, fool*) – predicted the bias direction for each verb significantly above chance. These results suggest that remention biases may be driven, at least partially, by a verb's argument structure and thematic roles.

More recently, Bott and Solstad (2014; see also Solstad & Bott, 2017) provided a more fine-grained semantic account of remention biases by proposing that a verb's remention bias emerges as a function of underspecified causal content in the verb's semantic structure. According to their account, lexical properties of certain interpersonal verbs (verbs that involve two human arguments) lack specific information regarding a potential cause of the event denoted by the verb and thus may elicit a preference to refer to one of the arguments when explicit explanations of the event are required. For example, the verb *annoy* denotes a causal relation between a stimulus-role-bearing subject and an experiencer-role-bearing object, such that an object becomes annoyed either as a result of an action initiated by a subject or due to some properties of a subject. However, the verb underspecifies information about what exact property or action of the subject caused the object to be in the state of being annoyed. Bott and Solstad proposed that such underspecified information can trigger explanations regarding an underlying cause of the event, allowing for continuations specifying properties or actions of the subject that may have resulted in the state of the object. They further hypothesize that verbs with different semantic structures require different types of missing information and thus entail different types of explanations. They proposed that Stimulus-Experiencer (SE) verbs (e.g., *annoy*) and Experiencer-Stimulus (ES) verbs (e.g., *love*) encode a *propositional* relationship between the arguments and thus entail *simple causes* (i.e., causes of events, states or attitudinal states that do not involve any volition or agentivity, e.g., *John disturbed Mary because* ***he was making lots of noise***) as explanations of the event, while Agent-patient (AP) verbs (e.g., *punish*) imply a *presupposition* that leads to the introduction of the causal dependent clause as *external causes* (i.e., causes external to the agent's mind, e.g., *John punished Mary* ***because she stole the money***). Their predictions about different types of explanations depending on verb classes were supported by a sentence completion study with German and Norwegian speakers in which each language group was asked to provide written continuations for sentence fragments in their own language (e.g., NP1 *verbed* NP2 *because*…). Bott and Solstad (2014) found that despite some variability in each group's responses, both language groups showed a similar pattern in general: They provided more simple causes than other types as explanations following ES verbs (87.9%) and SE verbs (74.7%), but more external reasons (76.6%) than other types as explanations following presuppositional AP

verbs. These results suggest that the fine-grained semantics of a verb, including missing information about causality, contribute significantly to the phenomenon of remention bias.

The issue of what exactly triggers remention biases remains a matter of continued debate among semanticists, and this dissertation is not intended to arbitrate between these accounts. The primary goal of this dissertation is to identify cross-linguistic differences between Korean and English in terms of remention bias strength, and to investigate effects of cross-linguistic activation in L2 learners' processing of remention biases in English that may arise from these differences. These accounts will thus not be further discussed in this dissertation.

### *2.1.1 Cross-linguistic differences between English and Korean remention predicates*

A majority of the English predicates that have been examined in the previous literature with regard to implicit causality or remention biases are interpersonal transitive verbs (Ferstl et al., 2011), in which information about causal relations between arguments is delivered in a single lexical item, as in (3). Unlike in English, many interpersonal predicates in Korean include light verb constructions (e.g., Chae, 1997) that are composed of a noun of Chinese origin and the light verb *ha* ('do'), as shown in (4a) and (4c). There is also a small number of verbs of Koreanic origin, which consist of a single morpheme, as in (4d), which I will henceforth call 'lexical verbs.' In addition to these types, Korean also has a (subject-biased) syntactic causative (SC) construction, best translated as 'cause X to be Y' (e.g., Lee, H-S, 2017; Lee, K., 1996), as illustrated in (4b).[3]

---

[3] In addition to the causative meaning, the "key ha" construction can also be associated with a permissive interpretation, as in the following example.

Nay-ka haksayng-tul-eykey i kyosil-eyse tampay-lul phiwu-key hay-ss-ta.
I-NOM student-PL-DAT this classroom-in cigarette-ACC smoke-COMP do-PAST-DECL
"I let the students smoke cigarettes in this classroom."

(O'Grady, 1991, p. 172)

Given that the *key h*a construction can have either a causative or a permissive interpretation, the term "syntactic causative" does not properly capture its full range of meaning. Although the referent of the subject is in control of the event denoted by the lexical verb, it is not necessarily the causer. Nevertheless, I will continue to use the term "syntactic causative" construction, since the main focus of this dissertation is the pattern that presents the referent of the matrix subject as the causer of the event.

(4a) Tom-i John-ul hyeppak-ha-yess-ta.
Tom-NOM John-ACC threat-do-PAST-DECL
'Tom threatened John.' (subject-biased light verb construction)

(4b) Tom-i John-ul nolla-key ha-yess-ta.
Tom-NOM John-ACC be surprised-COMP do-PAST-DECL
'Tom caused John to be surprised.' (subject-biased SC construction)

(4c) Tom-i John-ul pinan-ha-yess-ta.
Tom-NOM John-ACC criticism-do-PAST-DECL
'Tom criticized John.' (object-biased light verb construction)

(4d) Tom-i John-ul mit-ess-ta.
Tom-NOM John-ACC believe-PAST-DECL
'Tom believed John.' (object-biased verb)

A Korean SC construction (e.g., 4b) is characterized by the presence of the resultative suffix *-key* on the adjectival predicate (e.g., *nolla-* 'be surprised'), which is in turn followed by the causative verb *ha-* (literally 'do'; see O'Grady, 1991; Park, 1994; Sohn, 2001). Thus, the embedded predicate in a Korean SC construction describes a caused event, and the matrix verb *ha-* denotes the action that caused the event. For instance, the sentence (4b) is interpreted as John's becoming surprised as a result of Tom's action or behavior.

Note that like Korean, English also has an SC construction, such as *Tom caused John to be surprised,* where the cause and effect events are expressed by the matrix and embedded predicates (e.g., *caused* and *be surprised*, respectively). However, a noticeable cross-linguistic difference is found between experiencer-object verbs (i.e., verbs that require an experiencer as their internal argument) in English and their Korean translation counterparts. Several experiencer-object lexical verbs in English, such as *amuse, anger, please,* and *surpris*e can only be translated into Korean as an SC construction. For example, the English verbs *anger, bore,* and *surprise*, are translated into the Korean SC predicates *hwana-key ha-*, *cilwuha-key ha-*, and *nolla-key ha-*. By contrast, other English remention bias verbs have Korean translation equivalents which are simple lexical verbs (e.g., *pwulu-* 'call') or light verb constructions (e.g., *apwu-ha-* 'flatter', *hyeppak-ha-* 'threaten', and *sakwa-ha-* 'apologize to').

The cross-linguistic differences between Korean SC predicates and their English counterparts may affect their remention bias strength as they involve different structures. In particular, it has been hypothesized that predicates containing explicit marking of causality may give rise to stronger remention biases than lexical verbs. This hypothesis is put forward by Hartshorne and colleagues (Hartshorne et al., 2013), who note that some experiencer-object verbs in Japanese include a causative morpheme –(*s*)*ase*, generally translated as *cause*. Although causative markings in Japanese and Korean work in different ways in several respects (see Shibatani & Chung, 2001, for detailed comparisons of the two constructions), they both pick out the referent of the matrix subject as the causer of an event by means of an explicit marker. Hartshorne and colleagues hypothesize that the predicates containing the Japanese causative morpheme denote a more explicit causal relation between event participants, potentially leading to clearer remention biases for these verbs than for verbs in other languages that do not involve this marking. Given that, like the Japanese *–(s)ase* construction, the Korean SC (*-key ha*) construction denotes an explicit causal relation between the main and embedded clauses, the first experiment in this dissertation (Experiment 1, Chapter 3) was designed to test Hartshorne et al.'s conjecture by investigating the effect of causative marking on remention bias strength in the case of Korean predicates. Testing Hartshorne et al.'s hypothesis in the context of Korean allows for investigating whether cross-linguistic differences exist between the Korean and English in terms of the strength of remention biases, a crucial prerequisite for exploring whether such cross-linguistic differences may affect Korean-speaking learners of English in their referential processing in English, which constitutes the major motivation for this dissertation.

### *2.1.2 Previous studies on remention biases in L1 speakers*

There has been a substantial body of previous research on adult L1 speakers' use of remention biases in offline and online tasks (e.g., Au, 1986; Bott & Solstad, 2014; Brown & Fish, 1983; Caramazza et al., 1977; Cozijn et al., 2011; Featherstone & Sturt, 2010; Ferstl et al., 2011; Garvey & Caramazza, 1974; Garvey, Caramazza & Yates, 1974; Greene & McKoon, 1995; Hartshorne & Snedeker, 2013; Hartshorne et al., 2013; Itzhak & Baum, 2015; Koornneef & VanBerkum, 2006; Stewart et al., 2000). These studies include sentences comprised of a remention bias verb with two arguments in the first clause, followed by a second clause

providing an explanation for the first clause and investigate whether participants provide continuations consistent with the verb's remention bias when asked to complete the second clause in a sentence completion task, or whether they interpret an ambiguous pronoun in the second clause as consistent with the verb's remention bias during online processing.[4] These studies provide consistent evidence that comprehenders use remention bias information to construct event representations associated with causal relationships between preceding and ensuing clauses.

Based on the seminal work of Garvey and Caramazza (1974), who first identified the phenomenon of remention bias, Garvey and colleagues (Garvey et al., 1974) provided the earliest empirical evidence that remention biases influence the way that a comprehender resolves an ambiguous pronoun in a causal dependent clause. In a sentence-completion task, they asked 28 adult English speakers to provide written continuations for sentence fragments that contained remention bias verbs and two human arguments in the main clause, followed by *because* and an ambiguous pronoun, as in (5).

(5) The father scolded his son because he...

Results from participants' responses demonstrated that there were certain verbs creating strong biases to resolve the pronoun to either the previous subject or the previous object. Among the 16 verbs tested in Garvey et al.'s study, verbs such as *call* and *join* exclusively led to NP1 responses, and verbs such as *blame* and *kill* to NP2 responses. Other verbs that did not exclusively elicit NP1 or NP2 responses still induced consistent NP1 response patterns (e.g., *confess to, confide in, sell*) or NP2 response patterns (*criticize, distrust, fear, rush to, scold*) more than 70% of the time. From these results, Garvey et al. (1974) concluded that remention biases created by these verbs provide an important source of information that helps guide a comprehender to the specific interpretation of a sentence. Since Garvey et al. (1974), the strength and direction of remention biases from a wide variety of verbs have been tested across different languages in numerous

[4] In offline tasks, some studies included an overt pronoun after *because* (e.g., Commandeur, 2010; Cozijn et al., 2011; Garnham et al., 1996; Koornneef & Van Berkum, 2006), while others presented a sentence fragment without an overt referential expression in the embedded subject position (e.g., Bott & Solstad, 2014; Ferstl et al., 2011).

sentence-completion studies (e.g., Bott & Solstad, 2014; Commandeur, 2010; Cozijn et al., 2011; Dery & Bittner, 2016; Ferstl et al., 2011; Fukumura & Van Gompel, 2010; Goikoetxea, Pascual & Acha, 2008; Hartshorne & Snedeker, 2013; Hartshorne et al., 2013; Koornneef, Dotlačil, van den Broek & Sanders, 2016; Koornneef & Van Berkum, 2006; Long & De Ley, 2000; Mannetti & de Grada, 1991; Park, 2009; Stewart et al., 2000), which confirmed a robust effect of remention bias in reference resolution among L1 speakers.

In addition to offline work, a number of studies have used online methodologies to investigate the L1 processing of remention bias, with a particular focus on when the bias information becomes available to comprehenders in the course of sentence processing. Two hypotheses make distinct predictions on this timing issue: the integration account and the focusing account. The integration account claims that the effect of remention bias is only available when disambiguating information is obtained, usually at the end of a causal dependent clause (Garnham, Traxler, Oakhill, & Gernsabacher, 1996; Stewart et al., 2000). This perspective predicts that in "NP1 verb NP2 because he/she..." comprehenders show sensitivity to remention biases at regions where the causality information is explicitly delivered and thus the embedded subject pronoun is fully disambiguated, for example, at/near the end of the sentence or when disambiguating information is provided. Supporting evidence of the integration account comes from Garnham et al. (1996), who conducted a probe recognition task with English speakers to test when the effect of remention bias is observed during sentence processing. In a series of experiments, sentences like (6) were visually presented word by word in the center of a screen with an inter-word interval of 150ms, and a probe name (e.g., David or Brian) appeared at the top of the screen in capital letters at one of three time points: before the pronoun (150ms after the offset of *because*), after the pronoun (150ms after the offset of *he*), and at the end of the sentence. Participants were asked to respond by pressing one of the two buttons designated for the probe names as soon as they recognized the name. The probe name disappeared after participants responded or after 2.5 second following its appearance.

(6a) David approached Brian after school because he wanted some advice.

(6b) David approached Brian after school because he looked friendly.

(Garnham et al., 1996, p. 521)

Garnham et al. (1996) investigated the timing of a congruity effect, indicated by the response time gap between bias-consistent and bias-inconsistent sentences. For example, the pronoun in the bias-consistent sentence (6a) is preferentially interpreted as referring to *David*, congruent with the bias direction of the NP1-biased verb *approach*, whereas the pronoun in (6b) is more likely disambiguated into the bias-inconsistent referent *Brian*. Garnham et al. predicted that if participants show sensitivity to this congruity between the referent and the verb's bias, they will take longer to respond to the probe in the bias-inconsistent than bias-consistent condition. Their results showed a significant congruity effect when the probe name appeared at the end of the sentence, but no evidence of such an effect was found when the probe name was presented before or after the pronoun region. Garnham and colleagues interpreted these results as support for the integration account. However, some methodological limitations in their study should be noted. As Stewart et al. (2000) pointed out, their probe recognition task may have placed substantial processing demands on participants as they were asked to read a target sentence and respond to a probe name at the same time, which could have made it difficult for them to detect the probe names in earlier regions. Moreover, the secondary task during reading – probe recognition – is far from natural processing and may induce participants' engagement of strategic parsing. In addition, participants' sensitivity to remention bias was investigated at only three regions – immediately preceding/following the pronoun and at the end of the sentence, thus overlooking the possibility that any congruity effect may have spilled over to regions not captured in the task.

The same concern holds for the results from Stewart et al. (2000), another study advocating the integration account. In a self-paced reading task, they presented a target sentence in only two fragments by splitting it immediately after the pronoun in the *because*-clause (e.g., [*Daniel apologised to Arnold profusely because he*] / [*had been behaving selfishly*]). While they found a congruity effect only in the second fragment, thus interpreting their results as evidence of the integration account, their results do not unequivocally support this account, since one cannot rule out the possibility that a congruity effect from the pronoun region in the first segment might have spilt over to the second fragment (see Pyykkönen & Järvikivi, 2010).

Recently, more ecologically valid methods, including eye-tracking, have afforded more fine-grained investigations of the timing issue. In particular, the visual-world paradigm allows one to measure participants' temporal sensitivity to remention bias in spoken language comprehension without hampering natural sentence processing. Remention bias studies using this paradigm assess how often participants look at a picture of either the NP1 referent or the NP2 referent in a visual display while they listen to linguistic stimuli that include NP1 and NP2 protagonists in the subject and object positions along with a remention bias verb in the main clause, followed by a *because*-clause that provides an explanation of the event described in the main clause.

Several studies using this paradigm suggest that remention bias information may become available from much earlier points than what is predicted by the integration account during sentence processing. One proposal aligned with an earlier detection of remention bias information is the focusing account, which assumes that remention biases are available, potentially as early as the remention bias verb in the main clause, but at least before encountering any disambiguating information (Greene & McKoon, 1995; Koornneef & Van Berkum, 2006; Long & De Ley, 2000). Depending on the exact point of activation of remention biases, researchers have used different terms for the focusing account. Some proposed the "anticipation account" (McDonald & MacWhinney, 1995, p. 553) or the "truly 'proactive' anticipation account" (Koornneef & van Berkum, 2006, p. 460), which predict that remention biases are activated upon processing the remention bias verb and its object. While the anticipation account holds that a remention bias verb serves to increase the activation level of one of its argument as the likely causer, the truly proactive account views the activation of remention bias as a result of the processing of the verb in terms of a predictive, forward-looking mechanism.[5] Other proponents of the focusing account predict that remention biases are activated as soon as the connector and the subordinate pronoun have been processed (Cozijn et al., 2011). This proposal assumes that the verb's remention bias brings one of the NP referents in the main clause into the focus of attention, and the biased protagonist is rapidly activated when a pronoun is encountered. Since all these accounts claim that remention biases can be activated prior to disambiguating

[5] The anticipation account and the truly proactive account mentioned in McDonald and MacWhinney (1995) and Koornneef and van Berkum (2006), respectively, are discussed only in the context where the second clause following the first clause provides an explanation. Neither study mentions other types of coherence relations (e.g., consequence).

information, I will follow previous studies in grouping these accounts together under the term *focusing account* (e.g., Commandeur, 2010; Cozijn et al., 2011), in contrast to the integration account.

Several studies have provided evidence consistent with the focusing account by demonstrating that a remention bias effect emerges at or soon after the point when an embedded subject pronoun is encountered, immediately after the offset of the pronoun in a causal dependent clause (e.g., Cozijn et al., 2011; Featherstone & Sturt, 2010; Itzhak & Baum, 2015). For example, in a Dutch visual-world eye-tracking study, Cozijn and colleagues (Cozijn et al., 2011) found that speakers fixated on the bias-consistent referent significantly more often than the bias-inconsistent referent both for NP1- and NP2-based verbs soon after the offset of the pronoun. Their findings are taken to support the focusing account since the remention bias effect manifested before participants encountered any disambiguating information in the subordinate *because*-clause. Similar results were reported by Itzhak and Baum (2015), who investigated English speakers' sensitivity to remention bias in a visual-world eye-tracking task. While the main focus of their study was on the effect of prosody on speakers' use of remention bias information, they also report a separate analysis for processing of remention bias without prosodic information (their 'No-Accent' condition). In their results of the 'No-Accent' condition, there was a reliable effect of remention bias (indicated by significantly more fixations on NP1 than NP2 target with NP1-biased verbs and more fixations on NP2 than NP1 target with NP2-biased verbs) from the onset of the pronoun to 400ms after pronoun onset, consistent with the focusing account.

Even earlier activation of remention bias was reported by Pyykkönen and Järvikivi (2010), who argued that remention bias information can be assessed and activated as early as a remention bias verb and its arguments have been processed. In a Finnish visual-world eye-tracking experiment, they demonstrated that participants began to look at the bias-consistent antecedent significantly more often than the bias-inconsistent antecedent starting 900ms after the onset of the verb in the main clause, even before the connector *koska* ('because') or a pronoun was encountered in the *because*-clause. Pyykkönen and Järvikivi (2010) attributed these results to a predictive mechanism that enabled the comprehenders to recruit remention bias information to generate proactive expectations about one of the event participants as an underlying cause as

soon as they processed the remention bias verb. However, one cannot dismiss the possibility that their results may be an artifact of participants' strategic processing, as Cozijn et al. (2011) pointed out. As the experimental sentences in Pyykkönen and Järvikivi (2010) only included bias-consistent endings, it is possible that their participants developed a strategic focus on the remention bias information. Indeed, other studies that tried to fix this problem by including both bias-consistent and bias-inconsistent sentences found that remention biases are activated only after the connector has been processed (e.g., Cozijn et al., 2011; Itzhak & Baum, 2015).

While the previous literature on the timing issue of remention bias is still somewhat inconclusive, the more recent evidence points to the general conclusion that native speakers can use remention bias information quite early during sentence processing, even before they encounter information that disambiguates the subject pronoun in the subordinate *because*-clause.

### *2.1.3 Previous studies on remention biases in L2 learners*

Despite extensive evidence for native speakers' use of remention bias in reference interpretation and processing, less is known about how this information is utilized by L2 speakers. To the best of my knowledge, only four studies have examined this issue (Cheng & Almor, 2017, 2018; Contemori & Dussias, 2018; Liu & Nicol, 2010).

Liu and Nicol (2010) used a self-paced reading task to investigate whether advanced Chinese learners of English show online sensitivity to the mismatch between a verb's remention bias and the gender of the subject pronoun in a subordinate causal clause, as in (7).

(7a) The mother amused the father because he/she told funny jokes at dinner. (NP1-biased)
(7b) The boy admired the girl because he/she was so intelligent. (NP2-biased)

(Liu & Nicol, 2010, p. 154)

Liu and Nicol (2010) observed a congruity effect in terms of significant reading slowdowns in the dependent clause when the pronoun was inconsistent with verb's bias as compared to when the pronoun was consistent with verb-bias. The congruity effect was present in both their L1 and their L2 groups, albeit in slightly different regions. Notably, the congruity effect occurred earlier for NP1- than for NP2-biased verbs in the L2 group, but this was not the case in the L1 group,

indicating that L2 speakers may have depended more heavily than native speakers on other cues for generating an expectation for remention of the subject referent, such as the well-known subject bias (Hobbs, 1979), first-mention bias (Gernsbacher, 1990), and/or parallel structure processing strategies (Sheldon, 1974). Overall, Liu and Nicol's (2010) findings provide evidence that L2 learners, at least those at an advanced level, can use remention bias information in their referential processing in an L2, soon after an embedded subject pronoun is encountered.

Similarly, Cheng and Almor (2017) found that advanced Chinese-speaking L2 learners can make use of remention bias information during referential choices in a written English sentence-completion task with items such as *Ben embarrassed James because he....* Clear preferences for bias-consistent continuations were found for both L2 learners and native speakers, yet some group-differences were also observed depending on the verb type. In the case of NP1-biased verbs, both groups showed similar results: significantly more remention of the subject than object referent. For NP2-biased verbs, however, the bias effect was significantly weaker for the L2 group than the L1 group. Cheng and Almor interpreted the weaker effect with NP2-biased verbs in the L2 group as a reduced ability to integrate multiple sources of information, referring to Grüter et al.'s (2017) RAGE hypothesis, which holds that non-native speakers have '**R**educed **A**bility to **G**enerate **E**xpectations' about upcoming referents during discourse processing. Cheng and Almor also maintain that the L2 speakers in their study may have been strongly affected by other cues such as a subject- and/or first-mention bias, presumably driven by the presence of an overt pronoun in the causal dependent clause (e.g., Ariel, 1990; Gordon, Grosz & Gilliom, 1993). In other words, the presence of a pronoun in their sentence-completion task may have induced stronger reliance on form-related constraints associated with pronominal subjects.

In a follow-up study, Cheng and Almor (2018) addressed the role of referential form by presenting another group Chinese-speaking L2 learners with sentence fragments with and without an overt pronoun in the causal dependent clause in a written English sentence

completion task (e.g., *Mary called Sara because she... / Mary called Sara because ...*).[6] Their results showed that when the pronoun was not provided, both groups were biased to rementioning the subject referent following NP1-biased verbs and the object referent following NP2-biased verbs to the same extent. These results led Cheng and Almor (2018) to the conclusion that L2 learners can use remention bias information as efficiently as native speakers while making referential choices when no overt pronoun is provided as a prompt. When an overt pronoun was provided, however, L2 participants had a lower probability of rementioning the subject referent than native speakers with NP1-biased verbs, while producing significantly more continuations with the subject referent with NP2-biased verbs than native speakers did. Although Cheng and Almor speculated that form-related constraints associated with the overt pronoun might have contributed to these results, they admitted that the L2 learners' pattern in the pronoun condition was difficult to interpret.

The findings of Liu and Nicol (2010) and Cheng and Almor (2017, 2018) indicate that L2 learners can use verb-related biases. At the same time, L2 learners appear to be affected more strongly than L1 speakers by form-related constraints associated with pronouns. Weaker sensitivity to verb biases and stronger reliance on form-related cues among L2 learners is compatible with Grüter et al.'s (2017) findings from a story-completion study in which Japanese- and Korean-speaking learners of English showed target-like sensitivity to referential biases associated with referential form (pronoun vs. repeated name), but were less able than native speakers to exploit verb-related cues such as grammatical aspect (perfective vs. imperfective). While the effect of relative weighting of cues on L2 reference resolution and processing at various levels of linguistic representation is an interesting and relevant phenomenon, the present study is more specifically focused on L2 speakers' use of verb-related remention biases and on potential effects of cross-language influence in this regard. Therefore, in an attempt to diminish the potential influence of form-related constraints, the sentence-completion tasks in this

[6] It is difficult to compare proficiency levels of the L2 participants between this and their 2017 studies, since each study employed different tasks for measuring L2 proficiency: Oxford Quick Placement Test in the 2017 study and Test for English Majors (TEM) Band 4 and C-test in the 2018 study. Participants' proficiency in each study can only be estimated by the authors' statement that the participants in the 2017 study were classified as intermediate-advanced to advanced English learners, and the participants in the 2018 study were classified as advanced English learners.

dissertation do not provide an overt referring expression in the subject position of the subordinate clause as a prompt, thereby permitting participants to choose a reference form of their preference.

Unlike previous studies, which focused either on L2 learners' online sensitivity to remention biases (Liu & Nicol, 2010) or on their ability to use the information in offline referential choices (Cheng & Almor, 2017, 2018), Contemori and Dussias (2018) used both visual-world eye-tracking and written sentence-completion tasks in English to examine online as well as offline use of remention bias information in Spanish-English bilinguals. In the eye-tracking task, they measured participants' eye gaze on images depicting NP1 and NP2 while they listened to linguistic stimuli such as those in (8). The sentence-completion task involved the main clause portion of the sentences in the eye-tracking task followed by *because he*.

(8a) NP1 Verb-Congruent: Kevin apologized to Dave in the evening (pause) because he was scared and because he had insulted him.

(8b) NP1 Verb-Incongruent: Kevin apologized to Dave in the evening (pause) because he was scared and because he was insulted.

(8c) NP2 Verb-Congruent: Kevin believed Dave yesterday (pause) because he was kind and because he showed him the photograph of the crime.

(8d) NP2 Verb-Incongruent: Kevin believed Dave yesterday (pause) because he was kind and because he had seen a photograph of the crime.

(Contemori & Dussias, 2018, p. 164)

In the sentence-completion task, they found a clear preference for continuations consistent with the verb's remention bias for both NP1-bias and NP2-bias conditions for the bilingual group as well as the monolingual group. Specifically, both groups provided more subject than object reference following NP1-biased verbs, and more object than subject reference following NP2-biased verbs. Contemori and Dussias noted that unlike the L2 learners in Cheng and Almor (2017), who showed weaker sensitivity to remention biases with NP2-bias verbs in a written sentence-completion task, their bilingual participants successfully employed remention bias information in the same type of task, which they attributed to their participants' high L2 English proficiency and/or the typological similarities between their L1 (Spanish) and L2 (English).

In the eye-tracking task, Contemori and Dussias examined whether the bilinguals can use remention bias information as efficiently as the monolingual speakers to anticipate upcoming information without aid from disambiguating information. As evidence of listeners' predictive processing, they focused on the difference between looks to the "target" (i.e., the bias-congruent referent) picture and the "competitor" (bias-incongruent) picture prior to disambiguating regions, that is from the onset of the pause immediately following the main clause up to 1500ms after the pause onset. In the measurement of participants' overall proportion of looks, they found earlier divergence between looks to target versus competitor in the monolingual group than in the bilingual group in the NP1 condition, but no differences were found for either group in the NP2 condition. From these results, Contemori and Dussias concluded that the bilingual speakers were able to use remention bias information both for anticipatory processing (at least in the NP1 condition) and referential choices, but that their online processing of this information was delayed compared to monolingual speakers. This was interpreted as due to either the bilinguals' limited ability to activate the verb's remention bias information for anticipatory processing or their reliance on the first-mention bias upon encountering the pronoun.

Although Contemori and Dussias' (2018) findings provide potentially novel evidence on the role of remention bias in L2 processing, their report gives rise to a number of concerns about data analysis and interpretation. First, their primary analysis is based on separate models for the monolingual and bilingual groups. The conclusion of a delay in the bilingual group is based on the comparison between the output of these two models, rather than including group as a factor in a single model; indeed a secondary analysis on looks to target that included group as a factor showed no differences between the two groups. Moreover, their graphs show that the magnitude of differences of looks to the target and the competitor is much greater for the monolingual group than the bilingual group *prior to* the critical region. These substantial baseline effects in the monolingual group are not considered in their analysis and interpretation of eye gaze patterns in the critical region. Furthermore, they found an effect of remention bias only in the NP1 condition, with no evidence of predictive processing for either monolingual or bilingual speakers in the NP2 condition. They argued that the null effect of remention bias in the NP2 condition is consistent with the findings from Cozijn et al. (2011), which found no effect of remention bias prior to the disambiguating word among monolingual speakers in the NP2 condition in their first

experiment. However, Cozijn et al. did find a significant effect of remention bias in the pre-disambiguating segments for both NP1 and NP2 conditions in their second experiment, where they addressed potential problems of participants' strategic processing by removing a secondary task that asked participants to name the referent of the pronoun in the dependent causal clause and including filler sentences that contained no remention bias verbs and had connectors other than *because*. As Contemori and Dussias' (2018) do not provide specific information about a secondary task or fillers in their eye-tracking task, it is difficult to account for the null effect of remention bias in their NP2 condition. Taken together, these limitations make it difficult to assess to what extent remention bias was involved in both the monolingual and the bilingual speakers in this study.

In addition, the bilingual group in Contemori and Dussias (2018) consisted of highly proficient speakers with extensive, and mostly early, exposure to English (mean age of onset = 6 years) in an immersion context, and with an L1 (Spanish) typologically similar to the target language (English). This raises the question as to whether remention bias information is also available to less proficient, non-immersed L2 learners who have an L1 typologically distinct from English. Since proficiency and language learning experience are argued to influence L2 learners' engagement in predictive processing (Kaan, 2014), further research is required to take into consideration the role of these factors in L2 processing of remention bias information.

To address these gaps and limitations in existing research, the visual-world eye-tracking experiment in this dissertation (Experiment 4, Chapter 6) investigates the L2 processing of remention bias information, and the role of cross-linguistic activation in this process, in non-immersed Korean-speaking L2 learners of English with varying English proficiency, and analyses their performance in direct comparison with that of native speakers of English.

### 2.2 Cross-linguistic activation in language processing

In the bilingualism literature, there has been ongoing debate whether bilinguals and L2 learners activate only one language selectively or access information of words from both languages in an integrated way. This debate produced two general accounts of how L2 learners process target words – the language-selective account and the language-non-selective account. The language-selective account argues that lexical representations in bilinguals are stored

separately for each language, and hence the processing of words in one language is not affected by their counterparts in the other (Tulving & Colotla, 1970). In contrast, the language-non-selective account claims that learners' lexical memory is organized in an integrated system and thus activation of a word in one language can lead to activation of words in the other language (Dijkstra & Heuven, 2002).

While the debate about language-selective vs. non-selective activation of words remains ongoing, recent research on bilingual and L2 processing has provided substantial evidence that information from both L1 and L2 is accessed in parallel, cross-linguistically influencing each other at various levels of linguistic representation, thus broadly supporting language-non-selective accounts (Altarriba, 1992; De Groot & Nas, 1991; Dijkstra, & van Heuven, 2002; Dijkstra, Wahl, Buytenhuijs, van Halem, Aljibouri, de Korte & Rekké, 2018; Gollan, Forster & Frost, 1997; Hopp, 2016; Kroll, & Stewart, 1994; Prior, Degani, Awawdy, Yassin & Korem, 2017; Sanoudaki & Thierry, 2015; Spivey & Marian, 1999). For example, L1-Korean L2 learners of English are shown to have facilitated processing of English words whose meanings and/or sounds overlap with their Korean counterparts (e.g., Kim & Davis, 2003). In cross-language masked priming tasks, Korean speakers showed faster response time in naming English words preceded by Korean prime words that form interlingual cognates (e.g., *pheyn* 'pen' – pen) or homophones (e.g., *phwul* 'grass' – pull) as compared to control word pairs that have no semantic or sound overlap across the languages. The fact that a wide range of words that overlap in form and meaning can be activated in parallel across languages indicates that L2 learners have an integrated system of representations devoted to all the languages known to them (Dijkstra & van Heuven, 2002).

In this section, I will overview some characteristics of cross-linguistic activation observed in previous studies, and discuss how these characteristics shed light on the investigation of the influence of cross-linguistic activation between Korean and English remention bias predicates on Korean speakers' referential processing in English, the critical focus of this dissertation. I will also discuss how examining such influence affords insight into whether cross-linguistic activation goes beyond a word- and construction-level and influences discourse-level processing, an issue that has not been investigated in previous studies.

*2.2.1 Characteristics of cross-language activation*

One theoretical model that attempts to capture the widely-attested effects of cross-linguistic activation at various linguistic levels is the Bilingual Interactive Activation (BIA) + model (Dijkstra & Van Heuven, 2002). According to this model, L2 learners' or bilingual speakers' access to words and phrases across languages occurs in a parallel, non-selective manner such that L2 learners experience cross-linguistic interference at various levels of representation (i.e. orthographic, phonological, syntactic, semantic). In line with this model, numerous studies have demonstrated that L2 learners may be affected by cross-linguistic activation that operates in parallel across L1 and L2 at various levels, including words in isolation (e.g., Dijkstra, 2005; Duyck, 2005; Duyck, Van Assche, Drieghe & Hartsuiker, 2007; Gollan et al., 1997; Kim & Davis, 2003; Prior et al., 2017; Van Assche, Duyck & Brysbaert, 2013; Zhou, Chen, Yang & Dunlap, 2010), words in sentence contexts (e.g., Chambers & Cooke, 2009; Elston-Güttler, Gunter & Kotz, 2005; Hopp, 2016; Libben & Titone, 2009; Schwartz & Kroll, 2006; Van Assche, Drieghe, Duyck, Welvaert & Hartsuiker, 2011), lexical tones (e.g., Wang, Wang & Malins, 2017), and syntactic structures (e.g., Dussias & Sagarra, 2007; Hartsuiker, Pickering & Veltkamp, 2004; Jacob, Katsika, Family & Allen, 2017; Prior et al., 2017; Sanoudaki & Thierry, 2015; Schoonbaert, Hartsuiker & Pickering, 2007).

Studies that were motivated by the BIA+ model have identified several characteristics of cross-linguistic activation, which allow me to formulate predictions about cross-linguistic influence of causative marking in Korean predicates on Korean speakers' reference processing in English. One characteristic of cross-linguistic influence is that the effect of cross-language interference is enhanced when learners process an L2 rather than an L1 (Altarriba & Basnight-Brown, 2007; Gollan et al., 1997; Jared & Kroll, 2001; Marian & Spivey, 2003; van Hell & Dijkstra, 2002; Weber & Cutler, 2004). For example, Gollan et al. (1997) found a significant asymmetry in terms of the direction of priming (i.e., from L1 to L2 vs. from L2 to L1) in a masked translation priming task with Hebrew-speaking L2 learners of English. The learners responded significantly faster in their lexical decision when the target L2 words were cognates of L1 prime words than when the target and prime words were non-cognates. However, such priming effects were not found for L1 target words primed by L2 words, even when the target and prime words were cognates. The directional asymmetry in the priming effects led Gollan and

colleagues to conclude that influence from L1 to L2 may be stronger than the other way around. Although it is beyond the scope of this dissertation to assess cross-linguistic influence bidirectionally, the larger effect of cross-language activation coming from L1 to L2 than the other way around indicates that the predicted cross-linguistic influence on remention bias is most likely to emerge when Korean-speaking learners of English process L2 English sentences.

Previous research found that parallel activation of words can also occur between languages that are typologically different, particularly in terms of the orthographic system (e.g., Kim & Davis, 2003, for Korean and English; Dimitropoulou, Duñabeitia, & Carreiras, 2011a, for Greek and Spanish; Nakayama, Verdonschot, Sears, & Lupker, 2014, for Japanese and English; Prior et al., 2017, for Arabic and Hebrew). For example, Prior and colleagues (Prior et al., 2017) conducted a cross-modal semantic similarity judgment task in Hebrew with adult learners speaking Arabic as an L1 and Hebrew as an L2, languages that have different scripts. They found that when learners were asked to make a decision whether a written Hebrew target was semantically related to the prime that they previously listened to, they showed a slower response to the target word when the prime word was a false-cognate between Hebrew and Arabic compared to when it was a control word that did not share any semantic or phonological information with the target word. The learners' response pattern was distinguished from that of Hebrew monolingual speakers who responded to target words equally fast following false cognates versus control prime words. These results suggest that the phonological and semantic overlap between Hebrew and Arabic cognates affected the L2 learners' processing of Hebrew words, despite the substantially different orthographic systems between the languages. It is assumed that this type of transfer is possible as a result of previously established links between the syntactic and semantic properties of a word and the concept it relates to in learners' integrated mental lexicons (Jarvis, 2009; Kroll & Stewart, 1994). In light of the findings that cross-language activation occurs across typologically distinct languages with different orthographic systems, including Korean and English (e.g., Kim & Davis, 2003), it appears plausible to hypothesize that cross-linguistic influence between Korean and English in terms of remention bias information may occur.

Cross-language activation occurs not only within each level of representation but also across diverse levels of representation, such as sounds, word forms, word meaning, and syntax. Several

studies have demonstrated that transfer at the word level can affect bilinguals' or L2 learners' syntactic processing (e.g., Cai, Pickering, Yan, & Branigan, 2011; Helms-Park, 2001; Hopp, 2017; Schoonbaert et al., 2007). For example, Hopp (2017) tested effects of cognate words in cross-language syntactic activation in L1 German learners of English. In eye-tracking-while-reading experiments, participants read English sentences including relative clauses, as in (9), whose surface word order either coincides with the possible word order of a German sentence (9a) or does not constitute a possible word order in German (9b, 9c, 9d). Across the conditions, the status of the relative clause verb was manipulated by including English-German cognates (*ignored*) and non-cognate control words (*avoided*).

(9a) When the doctor Sarah ignored/avoided tried to leave the room the nurse came in all of a sudden. (Reduced Relative Clause – Embedded Clause)

(9b) The doctor Sarah ignored/avoided tried to leave the room when the nurse came in all of a sudden. (Reduced Relative Clause – Main Clause)

(9c) When the doctor who Sarah ignored/avoided tried to leave the room the nurse came in all of a sudden. ([Full] Relative Clause – Embedded Clause)

(9d) The doctor who Sarah ignored/avoided tried to leave the room when the nurse came in all of a sudden. ([Full] Relative Clause – Main Clause)

(Hopp, 2017, p. 105)

Results from fixation analyses revealed a significant effect of cognate status only for L2 learners, who showed significantly shorter first fixation durations on cognate verbs than on non-cognate verbs, whereas native speakers were not influenced by cognate status, confirming cognate facilitation among L2 learners. Notably, the effect of cognates found in the L2 group also interacted with relative clause type in such a way that L2 learners showed longer first-pass reading times for the second verb (e.g., *tried*) when the relative clause consisted of a possible word order in German (9a) than when it did not, and this effect emerged only when the embedded verbs were non-cognates. From these findings, Hopp (2017) concluded that cognate facilitation can free more resources, allowing learners to effectively inhibit L1 syntax during their L2 sentence processing.

The results from Hopp (2017) provide evidence of indirect effects of co-activation at a word level on syntactic processing by means of cognate facilitation freeing up resources for syntactic processing. Other evidence relevant to potential connections between cross-language activation at a word level and syntactic processing comes from experiments on cross-language priming. Schoonbaert et al. (2007) found a stronger priming effect from Dutch (L1) to English (L2) dative constructions when the sentences included Dutch-English translation-equivalent verbs (e.g., *gooien* and *throw*) than when they included non-translation-equivalent verbs (e.g., *gooien* and *give*). Similarly, Cai et al. (2011) showed that both translation-equivalent and cognates between Mandarin and Cantonese boosted the effects of syntactic priming in the dative constructions across the two languages. The magnified effects of syntactic priming by virtue of cross-language activation of cognate verbs (e.g., Cai et al., 2011) or translation equivalents of verbs (e.g., Schoonbaert et al., 2007) suggest that cross-language activation at the word level can influence L2 processing in a higher domain such as syntax. However, while these existing studies have focused on the phrase- and sentence-level as a target domain influenced by word-level co-activation, little is known about whether the effect of cross-linguistic activation at a word level extends to the discourse domain. Similarly, little is known about cross-language activation when the relevant correspondents in the two languages involve potentially different levels of representation, as is the case with lexical remention bias verbs in English whose closest translation correspondents in Korean consist of a phrase-level construction. In the following two sections, I address these unexplored domains in more detail.

### *2.2.2 Cross-linguistic activation between differently sized units*

While previous research has provided evidence for cross-language activation *within* diverse levels of representation and for the effects of co-activation at one level on processing at a different level, as discussed above, it is an understudied issue whether cross-language activation can occur between units of different sizes across languages, such as words and phrase-level constructions. A helpful concept in this context is Jarvis' (2009) notion of *lemmatic transfer*. Jarvis breaks down the traditional concept of lexical transfer into *lemmatic transfer* and *lexemic transfer*. Adopting earlier definitions of 'lemma' as a word's syntactic-semantic properties (Kempen & Huijbers 1983; Levelt, 1989) and 'lexeme' as a word's morphophonological

properties (Roelofs, Meyer & Levelt, 1998), and using 'transfer' as an interchangeable term for cross-linguistic influence, Jarvis refers to cross-linguistic influence associated with syntactic/semantic properties of words as *lemmatic transfer* and to cross-linguistic influence involving the phonological and orthographic form of a word as *lexemic transfer*. The notion of *lemmatic transfer* appears particularly relevant in the context of the translation correspondence between Korean and English predicates, as it involves co-activation of syntactic/semantic properties of words/constructions across the two languages. Jarvis' notion of lemmatic transfer includes a variety of transfer phenomena that relate to "the semantic and syntactic properties of words" (p. 102). As examples of lemmatic transfer, Jarvis lists semantic extensions (e.g., *spin* for *purr*, from Finnish *kehräta* meaning both *spin* and *purr*; Meriläinen, 2006), calques or loan translations (e.g., *youngman* for *bachelor*, from Swedish *ungkarl* literally meaning *young man*; Ringbom, 2001), collocational transfer (e.g., *do children* for *have children*, from Finnish *tehdä lapsia* literally meaning *do*/*make children*; Meriläinen, 2006), and subcategorization transfer (e.g., *late from an appointment* for *late for an appointment*; Meriläinen, 2006). Importantly, in addition to transfer at a level of individual words, the most extensively studied scenario in the bilingualism literature, Jarvis' notion of lemmatic transfer also includes transfer at the level of multi-word units, or constructions. This opens the door for the inclusion of transfer between single- and multi-word units, the case at hand here, and one that to the best of my knowledge has not been addressed in the previous literature.

According to cognitive linguistic theories of Construction Grammar, constructions are defined as form-meaning correspondences where a particular form is conventionally paired with a meaning (Fillmore, Kay & O'Connor, 1988; Goldberg, 1995, 2006). Constructions include linguistic units that contain unique pairing of form and meaning at various levels of abstraction, ranging from specific items, such as simple lexical words (e.g., *dog*, *pencil*) and idiomatic expressions (e.g., *to make a long story short*), to abstract syntactic frames such as phrasal patterns (e.g., *The Xer, the Yer*) and argument structure constructions (e.g., NP-Verb-NP-NP, meaning transferring of an object; Ellis, Römer & O'Donnell, 2016; Hunston & Francis, 2000; Lakoff, 1987; Michaelis & Lambrecht, 1996; Trousdale & Hoffmann, 2013). In contrast to the concepts of lemma and lexeme, which specify a word's semantic, syntactic and morphophological properties, a construction is a more comprehensive term that encompasses a

broader range of linguistic units that carry properties of morphophonological, semantic, and syntactic information (Ellis et al., 2016; Goldberg, 2013).

The Korean SC construction, the target construction investigated in this dissertation, falls into the category of a phrase-level construction, in the sense that it carries a conventionalized pairing of form (V-*key ha*) and meaning (cause-to-do) (Lee, 2017). As illustrated in Section 2.2.1, some English remention bias verbs can only be translated into a Korean SC construction. Thus the effect of co-activation of English verbs and the Korean equivalent SC construction could be regarded as transfer between a word (i.e., English lexical verbs) and a phrase-level construction (i.e., Korean SC construction). Despite the prolific research on cross-linguistic influence at the word-level, this specific case of lemmatic transfer involving the cross-language correspondence between an individual word and a phrase-level construction has received little attention in the literature on cross-language activation.

### *2.2.3 Effects of cross-linguistic influence on discourse-level processing*

Another issue that has hardly been addressed in previous studies is whether cross-linguistic activation at word and construction levels can affect discourse processing. Although numerous studies have investigated cross-linguistic interference beyond the lemma level such as syntactic structures (e.g., Dussias & Sagarra, 2007; Hartsuiker et al., 2004; Jacob et al., 2017; Prior et al., 2017; Sanoudaki & Thierry, 2015; Schoonbaert et al., 2007), none to my knowledge have addressed the question as to whether the effect of cross-linguistic activation goes beyond the word and construction levels to potentially influence learners' pragmatic inferences as reflected in their referential choices in a separate clause. Investigating effects of cross-language activation in the L2 processing of remention bias information helps address this gap, since the Korean speakers' processing of English remention bias verbs creates a potential for transfer of Korean equivalents including an SC construction (i.e., cross-language co-activation between a word and a construction); and computation of the inference of remention bias calls upon discourse-level processing, which requires learners to construct the mental models of the events described by the main and causal dependent clauses and identify the causal relations between the two events. To this aim, this dissertation investigates potential effects of cross-linguistic activation at word and (phrasal) construction levels on *discourse-level* processing. Given that Jarvis' notion of lemmatic

transfer covers a broad range of cross-language influence associated with a word's semantic and syntactic properties, and these properties, particularly in verbs, help to contribute to the overall form and meaning of a sentence, it is expected that effects of lemmatic transfer should go beyond word and construction levels and extend to a higher level such as discourse.

## 2.3 Effects of L2 proficiency and learning experience on cross-linguistic activation

While previous findings provide consistent evidence of non-selective, parallel activation of words from both languages at multiple levels of representation, additional factors may modulate the extent to which L2 learners activate words from their L1 during L2 processing, particularly those known to affect L2 processing in general, such as L2 proficiency and the amount/type of exposure to a target language (Kaan, 2014). In this section, I will review what is known about how two potential factors, L2 proficiency and learning experience (classroom vs. natural exposure), influence cross-language activation.

### *2.3.1 Proficiency and cross-language activation*

While several studies investigated the role of proficiency in cross-language activation, results regarding its effect remain less than conclusive. Some studies report that increased L2 proficiency can either facilitate or attenuate lexical or syntactic co-activation (e.g., Bernolet, Hartsuiker, & Pickering, 2013; Brenders, Van Hell & Dijkstra, 2011; Libben & Titone, 2009), while others show pervasive effects of cross-language activation regardless of learners' L2 proficiency (e.g., Chambers & Cooke, 2009; Duyck, Diependaele, Drieghe, & Brysbaert, 2004; Haigh & Jared, 2007; Jared & Kroll, 2001; Jared & Szucs, 2002; van Hell & Tanner, 2012; Zhou et al., 2010). For example, Libben and Titone (2009) found a stronger cognate effect for lower-proficiency participants in a sample of generally highly proficient L1 French learners of L2 English. In an eye-tracking-while-reading study, they presented participants with English sentences containing English-French cognates (e.g., *piano*), homographs (e.g., *coin* meaning 'corner' in French), and matched control words (e.g., *wedding*). Their results showed a negative correlation between learners' English proficiency (measured by self-reported ratings) and cognitive facilitation (i.e., reading time differences for cognates versus matched control words). While learners in general read sentences containing cognates faster than sentences with control

words, the effect of cognate facilitation was stronger for the lower proficiency subgroup than the higher proficiency subgroup.

Brenders et al. (2011) also found effects of L2 proficiency in child L1 Dutch L2 English learners' English word recognition, yet the proficiency effect in their study occurred in the opposite direction to what was reported in Libben and Titone (2009). In a lexical decision task, English learners at beginner ($5^{th}$ and $6^{th}$ graders) and more advanced ($7^{th}$ and $9^{th}$ graders) levels were presented with English words that constitute either cognate (e.g., *ankle-enkel*) or non-cognate (e.g., *air-lucht*) relations with the Dutch correspondents, and asked to determine whether each item is an English word or non-word. Brenders et al. found a cognate facilitation effect for more advanced learners, who made a faster lexical decision for cognates than for non-cognates, but not for beginner-level learners who demonstrated slower response times for cognates than for non-cognates. From these results, the researchers concluded that cognate facilitation effects are robust in young L2 learners at higher proficiency, yet the effects may be weaker or even absent for beginner and intermediate L2 learners because cognates may create ambiguity regarding language membership for less proficient learners and thus delay their response.

Although results from Libben and Titone (2009) and Brenders et al. (2011) seemingly conflict with each other with respect to the role of proficiency in L2 lexical processing, both findings can be captured by the Revised Hierarchical model (RHM, Kroll & Stewart, 1994). The RHM assumes that word meanings are more strongly associated with L1 than L2 word forms, and accessing meanings from L2 words is mediated by the L1 translation equivalents. According to the model, the link between L2 word forms and meanings remains weaker when the lexical representations in L2 words are not fully developed, as in the case for less proficient L2 learners. As a result, the relatively weak link between L2 word forms and meanings may lead to more delayed processing of L2 words in less proficient compared to more proficient L2 learners because lower proficiency requires learners to rely more on L1 translation to access semantic information of target words. The delayed access to word meanings in less proficient learners may engender two contrasting predictions for L2 lexical processing.

On the one hand, slower lexical access may result in longer time to activate words from both languages, allowing for a greater chance for cross-language co-activation (van Hell & Tanner, 2012). Libben and Titone's (2009) results may support this prediction since they demonstrate a

stronger cognate facilitation effect for less proficient learners. It is conceivable that the less proficient learners in their study may have been more reliant on L1 translation in reading the target words, experiencing greater effects of cognate facilitation than the more highly proficient learners, who are assumed to rely less on L1 translation links to access word meanings.

On the other hand, slower access to L2 words can cause a smaller extent of word co-activation in the course of processing because learners have greater difficulty to retain information of activated words in their working memory. As van Hell and Tanner (2012) noted, less proficient learners may be vulnerable to a greater decay rate of L2 words, particularly when they engage in a secondary task after processing target words, such as in a lexical decision task. This prediction is consistent with the findings from Brenders et al.'s (2011) lexical decision study in which beginner-level learners failed to show cognate facilitation. The absence of cognate facilitation in these beginner learners may have occurred as a function of their greater difficulties in retrieving previously activated words.

In sum, previous studies demonstrate that the degree of cross-linguistic activation may be modulated by proficiency and the timing of lexical access such that lower proficiency allows for more chances of activating words in both languages right around the time when target words are processed, but as time elapses afterwards, learners may be more susceptible to a decay of activated words, possibly leading to reduced co-activation of target words.

Although the RHM predicts different degrees of strength for connections between L2 word forms and meanings depending on L2 proficiency, the model also postulates that association strength between them still remains relatively weaker as compared to between L1 word forms and meanings, no matter how proficient a learner becomes in an L2, leaving open the possibility that even highly proficient learners may not entirely escape from the influence of their L1 when they process L2 words. Indeed, numerous studies report persistent effects of cross-language activation for highly proficient L2 learners as well as less proficient learners (e.g., Chambers & Cooke, 2009; Duyck et al., 2004; Haigh & Jared, 2007; Jared & Kroll, 2001; Jared & Szucs, 2002; van Hell & Tanner, 2012; Zhou et al., 2010). For instance, Duyck and colleagues (2004) tested both highly proficient and moderately proficient Dutch-speaking L2 learners of French in a cross-language phonological priming task. In this study, participants made a lexical decision for a target word (e.g., *crane* 'skull') preceded by a masked Dutch prime word that either

phonologically overlapped with the target (e.g., *kraan* 'tap') or shared no phonological representation with the target (e.g., *graan* 'steam'). Duyck et al. found consistent phonological priming effects across both proficiency groups, who showed faster response times for phonologically related target words than for control words. Similarly, Dimitropoulou, Duñabeitia, and Carreiras (2011b) investigated cross-language translation priming in Greek-speaking learners of English who were divided into three proficiency levels (low, medium, high) based on their self-reported proficiency ratings and scores from independent placement tests. Their results revealed that while lexical decision accuracy improved with increasing proficiency, translation priming effects were consistently observed across all proficiency groups, indicating that cross-language activation of words can occur in learners with a wide range of L2 proficiency.

Taken all together, while previous research provides some evidence of the modulating role of L2 proficiency in cross-linguistic activation, several studies show that L2 learners at various proficiency levels are still subject to interference from cross-language co-activation. These inconsistent outcomes may come from differences across the studies in terms of learners' background, such as age and L1, and the task types employed, which makes it difficult to draw firm conclusions from these findings. To better characterize the modulatory role of proficiency in cross-linguistic activation, it is necessary to test learners from the same age and L1 using a variety of experimental tasks including both offline and online measures. Moreover, as no definitive answer has been given to the issue of effects of cross-linguistic activation at word and construction levels in L2 discourse processing, as reviewed in the previous section, it also remains an open question how proficiency modulates the way that cross-language activation at word and construction levels influences sentence- and discourse-level processing. For these reasons, this dissertation included learners' scores on a lexical decision task and a cloze test as measures of L2 proficiency (Experiments 3 and 4, Chapters 5 and 6).

### *2.3.2 L2 learning experience and cross-language activation*

L2 learners' language learning experience is one of the most important indicators of both the quality and amount of input learners receive. Most often, L2 learners who learn a target language in an immersive situation are exposed to an abundant amount of naturalistic language input, whereas learners in a non-immersive context receive a restricted amount of target input, mostly

confined to textbooks and formal (written) language (Saito, 2017; Yang, 2010). It is argued that increased frequency of exposure to a target language permits a learner to form a stronger link between a target word's form and its meaning (Javis, 2009; Michael & Gollan, 2005) and to inhibit properties of learners' L1 more efficiently (Dussias, 2003; Dussias & Piñar, 2010; Frenck-Mestre, 2002; Pliatsikas & Marinis, 2013; Selinker, 1972).

Despite its purported role in L2 learning and processing, L2 learning experience has received relatively little attention in the literature on cross-linguistic influence. To my knowledge, there have been few studies (except for Zhao, Li, Liu, Fang & Shu, 2011, see below) that directly tested for the effects of L2 learning experience on cross-language activation. Instead, most studies focused on either immersed or instructed learners and investigated whether cross-language activation occurs within each learner group separately. These studies reported cognate facilitation both for learners who had only classroom instruction (e.g., Brenders et al., 2011; Jacob et al., 2017; Kantola & van Gompel, 2011; Lemhöfer, Dijkstra & Michel, 2004; Weber & Cutler, 2004) and for those who had been immersed in the L2 for at least one year (e.g., Costa, Caramazza & Sebastián-Gallés, 2000; Haigh & Jared, 2007; Poarch & Van Hell, 2012; Wang & Foster, 2015; Zhao et al., 2011).[7] However, it is impossible to determine the precise role of L2 learning experience in cross-language activation based on outcomes from these studies, since they employ a wide range of methodologies with learners from different L1 backgrounds and age. To evaluate how L2 learning experience moderates cross-language activation, ideally one should look at how learners with different types of learning experience but with closely matched learner characteristics, such as L1 background, age, and proficiency, perform in the same type of tasks that tap into cross-language activation.

This issue was partially addressed by Zhao and colleagues (2011), who investigated priming effects for translation equivalents and semantically related words between Chinese and English among two groups of advanced Chinese-speaking learners of English with different L2 learning experience (study-abroad vs. classroom experience) but with closely matched English proficiency. Their results showed that learners with only classroom experience showed

[7] This list of references was selected among those studies reporting participants' language experience information. Many other studies on cross-language interference do not provide specific information on the amount/type of participants' L2 learning experience, and thus are not included here.

translation priming effects only from the direction of L1 to L2. In contrast, the group with study-abroad experience demonstrated significant translation priming effects in both directions of L1 to L2 and L2 to L1. Zhao et al. interpreted this priming asymmetry as a result of more knowledge of L2 words and thus stronger associations between L1 and L2 words in the mental representations of the learners with immersion experience. Despite the group differences, these results suggest that both immersed and non-immersed L2 learners experienced cross-linguistic interference from their L1 prime words when they processed L2 target words (L1 to L2 direction).

While Zhao et al.'s (2011) study found no evidence of modulating effects of L2 learning experience in cross-language activation when L2 target words are primed by L1 words, these outcomes might not necessarily generalize to cross-linguistic activation in other contexts, such as processing verbs in a sentence or discourse context. Different scenarios can be envisaged. On the one hand, it is possible that strong effects of cross-language co-activation occur regardless of learners' L2 learning experience, as in the case of Zhao et al. (2011), affecting learners' sentence- and discourse-level processing equally across the board. Alternatively, in light of previous findings that L2 learners with extensive immersion experience are less likely to be influenced by their L1 properties than non-immersed learners in sentence processing (e.g., Dussias & Sagarra, 2007; Frenck-Mestre, 2002; Pliatsikas & Marinis, 2013), it is possible that L2 learners with immersion experience may be more likely than instructed learners to adopt target-like processing strategies, showing reduced effects of cross-language interference in sentence and discourse processing.

### 2.4 Research questions and predictions for experiments

The overall objective of this dissertation is to investigate effects of cross-linguistic activation on referential processing at the sentence- and discourse-level by examining the referential biases of Korean-speaking L2 learners of English in sentences containing English translation equivalents of Korean SC and non-SC predicates. More specifically, it explores the extent to which differences in bias strength associated with different verb types and constructions in their L1 affect Korean-speaking L2 learners' offline and online reference processing in

English. The specific research questions (RQs) to be addressed by the four experiments in this dissertation are stated below.

RQ1) *Do Korean speakers show stronger subject-bias for Korean SC (-key ha) than non-SC predicates?* (Experiment 1)

As a prerequisite for testing for effects of cross-language activation in Korean speakers' processing of English sentences containing remention bias verbs, it is important to establish that cross-linguistic differences do exist between the English verbs and their Korean translation counterparts in terms of remention bias strength. Testing cross-linguistic differences between English remention bias verbs and their Korean translation counterparts is motivated by Hartshorne et al.'s (2013) hypothesis that explicit marking of causality may give rise to stronger remention biases than implicit causality in the lexical verbs typically examined in research on remention biases. They raised this hypothesis in the context of Japanese, noting that some experiencer-object verbs in Japanese are realized by inserting a causative morpheme –(*s*)*ase*, which denotes cause-to-do. To test this hypothesis, two parallel written sentence-completion experiments are conducted in Korean (Experiment 1a) and English (Experiment 1b). Based on Hartshorne et al.'s (2013) hypothesis, I predict that the Korean SC (*-key ha*) construction, which like the Japanese –(*s*)*ase* construction involves explicit marking of causality, will lead to stronger subject bias than non-SC predicates in Experiment 1a, whereas no differences will be observed between their English counterparts, which are all lexical verbs, in Experiment 1b.

RQ2) *Do the cross-linguistic differences in syntactic and semantic structure of predicates affect Korean-speaking L2 learners' referential choices in English?* (Experiment 2)

(RQ2-1) Do Korean learners of English carry over remention bias from Korean predicates while making referential choices in English causal dependent clauses? (Effects of cross-linguistic influence)

(RQ2-2) Does completing a translation task preceding the sentence-completion task enhance the extent to which learners carry over remention bias from Korean predicates? (Effects of translation priming)

Building on the results of Experiment 1, which provides empirical support for the relevant cross-linguistic difference, Experiment 2 investigates the effects of cross-linguistic activation of remention predicates and their associated constructions in Korean-speaking L2 learners' referential choices in English (RQ2-1). An English written sentence-completion task is conducted with Korean-speaking L2 learners of English and English native speakers where the materials contain English remention bias verbs that best translate into either an SC or non-SC predicate in Korean. If cross-linguistic differences of referential bias between Korean and English predicates affect Korean-speaking learners' referential processing in English, L2 learners will produce more continuations with subject reference in the SC type than in the non-SC type sentences, whereas native speakers will show little difference between the predicate types. In addition, previous studies found that exposure to the other language (e.g., L1) immediately prior to an experiment can enhance the activation of that language during the processing of the target language (e.g., L2) (e.g., Canseco-Gonzalez, Brehm, Brick, Brown-Schmidt, Fischer & Wagner, 2010; Elston-Güttler et al., 2005). The present study extends these previous findings and examines whether completing the translation before the sentence-completion task leads to more activation of Korean predicates, inducing a stronger effect of cross-linguistic activation in learners' referential choices (RQ2-2). To this aim, L2 learners will be randomly assigned to two sub-groups: a translation-first (T1) group who completed the translation task before the sentence-completion task, and a translation-second (T2) group who completed it after.

As previous research on cross-linguistic activation has focused predominantly on word- or sentence-level processing, addressing this research question will contribute new evidence regarding whether cross-language activation occurs between words and constructions and whether such co-activation goes beyond the word- and construction-level and affects sentence- and discourse-level processing.

RQ3) *To what extent do L2 proficiency and learning experience modulate the effects of cross-linguistic activation at the word- and construction-level on referential choices?* (Experiment 3)

Although L2 proficiency and learning experience are assumed to play an important role in L2 lexical and sentence processing (Basnight- Brown & Altarriba, 2007; Kaan, 2014; Prior et al., 2017), these two factors have not been sufficiently investigated in the literature of cross-

linguistic activation. Previous studies employing a wide variety of methodologies with learners of various L1 backgrounds and age have produced inconsistent results regarding the roles of proficiency and learning experience in moderating cross-linguistic activation. Moreover, little is known about whether these factors also influence the way that cross-language activation at word and construction levels affects sentence- and discourse-level processing. A written English sentence-completion task is conducted with learners who have different L2 proficiency and learning experiences. To obtain a wider range of proficiencies and investigate to what extent learners' learning experience and proficiency influence effects of cross-linguistic activation in L2 referential choices, learners' scores on a lexical decision task and a cloze test are included as measures of L2 proficiency, and half of the learners are recruited in the U.S. and half in Korea. If L2 proficiency and learning experience modulate effects of cross-linguistic activation in sentence and discourse processing, there will be a significant interaction between each of these two factors and differences of remention bias strength in the SC and non-SC predicates in L2 learners.

RQ4) *To what extent do cross-linguistic differences in bias strength affect learners' referential biases during real-time processing?* (Experiment 4)

Experiment 4 examines to what extent cross-linguistic activation of remention predicates affects the L2 use of remention bias information in online processing by employing the visual-world eye-tracking paradigm. During the task, Korean-speaking learners of English and native English speakers will see visual scenes containing images of two characters while listening to sentences where the two characters appear as the subject and the object along with a remention bias verb that corresponds to either an SC or non-SC predicate in Korean. If cross-linguistic activation of Korean SC constructions affects L2 learners' processing, there will be significantly more fixations on the target image that denotes the referent consistent with the bias of the verb relative to the bias-inconsistent competitor image with SC than with non-SC predicates for L2 learners, but no such difference will be observed for native speakers. Experiment 4 also tests modulating effects of proficiency in cross-linguistic activation. If proficiency affects the size of effects from cross-linguistic activation, there will be an interaction of proficiency and verb type (SC vs. non-SC).

# CHAPTER III
# EXPERIMENT 1: REFERENTIAL BIASES IN NATIVE KOREAN AND ENGLISH

## 3.1 Introduction

Experiment 1 investigates cross-linguistic differences in remention bias strength between Korean and English predicates by testing whether explicit causative marking in some Korean predicates would induce stronger remention bias than in other Korean predicates without causative marking. Experiment 1 consists of two sentence-completion experiments conducted in Korean (Experiment 1a) and English (Experiment 1b), respectively. As reviewed in Section 2.1.1, some Korean predicates contain explicit causative marking that their English counterparts do not have. Experiment 1a tests whether native Korean speakers show stronger subject biases for Korean predicates with explicit causative marking than (subject-biased) predicates without causative marking. Experiment 1b is conducted to establish that there are no differences in bias-strength for the English translations of Korean predicates with versus without causative marking among native English speakers, since (almost) all of the translations consist of lexical verbs with no explicit encoding of causality information. In Experiment 1a, native speakers of Korean completed written sentence fragments containing predicates with/without causative marking in Korean. In Experiment 1b, native speakers of English completed the same items translated into English.

## 3.2 Experiment 1a: L1 Korean written sentence-completion task

### *3.2.1 Methods*

#### *3.2.1.1 Participants*

A total of 36 native speakers of Korean (age 20–22) participated in this experiment. To minimize any interference of languages other than Korean, all participants were recruited from a college in Korea. A language background questionnaire revealed that these participants had started learning English at the mean age of 12, and that none of them reported having stayed outside Korea longer than 3 months. All participants received monetary compensation for their participation.

*3.2.1.2 Materials and design*

As materials for the sentence-completion task, 80 Korean predicates (40 SC, 40 non-SC) were selected based on the following steps. First, a pool of English verbs was taken from previous studies that investigated the effects of remention biases in offline and online monolingual comprehension (Garnham et al., 1996; Kasof & Lee, 1993; Long & De Ley, 2000; Rohde & Ettlinger, 2011; Rohde, Levy, & Kehler, 2011; Stewart, Pickering, & Sanford, 1998). From this verb pool, I selected 40 subject-biased and 40 object-biased verbs among those that were reported as showing bias toward subject or object more than 60% of the time. In the selection of subject-biased verbs, the status of the Korean translation counterparts of these verbs was considered, such that half of the subject-biased verbs corresponded most closely to a Korean syntactic causative construction containing *-key ha* (SC construction, e.g., *culkep-key ha* 'amuse'),[8] and the other half corresponded to a non-SC construction in Korean (e.g., *hyeppakha* 'threaten'). At this stage in the item selection process, English-to-Korean translations were conducted as closely as possible by using the NAVER English-Korean dictionary (http://dic.naver.com/). For each English verb, the Korean translation that appeared as the first entry in the dictionary was chosen. As a result, the experimental stimuli included three types of Korean predicates in terms of the direction of the bias that the predicate was expected to create, and the presence of *-key ha* in the predicate: 20 subject-biased non-SC predicates, 20 subject-bias SC predicates, and 40 object-biased predicates. The predicates used in the Korean sentence-completion task are listed in Appendix A.

For each predicate, a sentence fragment was created as illustrated in (10).

(10) Eceyspamey Hyesoo-ka Younghee-lul
last night Hyesoo-NOM Younghee-ACC
mwusep-key hay-ess-nuntey waynyahamyen _____
be frightened-COMP do-PAST-connective because
"Last night, Hyesoo frightened Younghee because _____."

[8] There was one predicate, *mayhoksikhita*, among the SC items that did not include *-key ha* (see Appendix A). It instead contains a lexical causative verb *–shiki*, which means 'to cause/force'.

Each sentence fragment was introduced by an adverbial phrase, which served as providing temporal (e.g., *last night*) or locational (e.g., *in the room*) information for the event being introduced in the first clause. After the adverbial phrase, two protagonists appeared as the subject and the object followed by the main verb, constituting the canonical word order of SOV in Korean. The subject NP was marked by the nominative case marker -*ka* or -*i*,[9] and the object NP was marked by the accusative case marker -*ul* or -*lul*. The gender for the NP protagonists was kept identical within each clause (both male or both female), because gender-contrasting referents may provide participants with additional cues for referential choices (Long & De Ley, 2000; Stewart et al., 2000). As such, half of the items included only male protagonists and half included only female protagonists. The gender was signalled by common Korean proper names. Two native speakers of Korean, who did not participate in the sentence-completion task, confirmed that all names sounded natural and that their gender was easily detectable.

The main predicate in each sentence fragment was presented in the past tense. In addition, a relational connective *nuntey* was attached to the verb in order to denote the discourse coherence relation between the first and the second clauses. This connective functions as a background builder (Lee, 1993; Park, 1999) which combines the two clauses into a sentence by marking the previous clause as background information for the following clause.

[9] There are two ways of realizing a case for subject in Korean. One is to use the nominative case marker -*i* or -*ka*, and the other is to use the topic marker -*un* or -*nun*. The present study avoided the topic marker for the NP in the subject position, since as Walker, Iida and Cote (1994) point out, the Korean/Japanese topic marker can influence a referent choice in discourse due to their high salience such that topic marked NPs are more likely than nominative-marked ones to be mentioned in the subsequently following utterance. The potential influence of the topic marker in referential choices was tested by Ueno and Kehler (2016), who asked adult Japanese speakers to produce a story after reading a Japanese sentence fragment containing a subject (either topic- or nominative-marked), an object, a remention bias verb, and the connector *because*. Although the results showed no statistical difference between the two conditions of topic- and nominative-marked subjects with regard to the likelihood of participants' choice of the subject referent in the second clause, the Japanese speakers mentioned subject in the topic-marked condition more frequently than they did in the nominative-marked condition. The present experiment therefore constrained the subject to be marked by nominative -*i* and –*ka* to avoid any potential influence that may come from using the topic marker for the subject NP.

After the first clause, the second clause was introduced by the connector *waynyahamyen* ('because') and a blank line.[10] Unlike some studies that presented an overt pronoun immediately following the conjunction in sentence-completion (e.g., Commandeur, 2010; Cozijn et al., 2011; Garnham et al., 1996; Koornneef & Van Berkum, 2006), the experimental stimuli in this study did not include a pronoun or any referential expressions in the second clause for two reasons. First, the overt pronoun system in Korean is underused (Lee, Lee, & Chae, 1997; Han, 2006), rendering the use of a pronoun in a context like (10) less felicitous. For example, the use of the 3rd-person feminine singular pronoun *kunye* for the second-clause subject in (10) may degrade the overall naturalness of the sentence.[11] Instead, the use of a repeated name in the subject position may be the most natural in this context (see Miyao, 2017, for Japanese). Second, the use of a pronoun in the subject position of the second clause may be inappropriate considering that Korean, a null-subject language, often prefers a null subject in a following clause. This null-subject phenomenon in Korean is observed particularly when there are requirements for continuity of topic and avoidance of redundancy in discourse (Kim, 1999; Roh & Lee, 2003). Given that a null subject can be used in a context like (10) for continuation of the topic, including a pronoun prompt in this context may hamper the natural production of sentence continuations. For these reasons, the current experiment did not provide any prompt in the second clause after *waynyahamyen* ('because'), allowing participants to freely choose the form of the subject (e.g., overt pronoun, null pronoun, repeated name) as well as its referent.

#### *3.2.1.3 Procedure*

Participants completed a language background questionnaire and the sentence-completion task in a quiet room. Each participant was given a booklet, which contained questions about language background, instructions for the task, and 80 experimental sentence fragments, all written in Korean. Participants first completed language background questions on a single page,

[10] Even with this connective, the first clause is interpreted as ending with a full stop, followed by the next clause in an independent clause. Unlike the English *because*, which entails a subordinate clause, the Korean *waynyahamyen* is an adverb, rather than a conjunction, that helps explicate the cause of a preceding sentence (Kweon, 2008).

[11] Five Korean informants mentioned that it is unnatural to include the 3rd-person feminine singular pronoun *kunye* appearing in the second-clause subject position. They instead pointed out that the sentence is more natural with a repeated name (e.g., *Hyesoo* or *Younghee*).

and then they started the sentence-completion task that appeared from the second page, with 20 items presented on each page. Before starting the experiment, participants received oral and written instructions on the task. They were asked to read each sentence fragment and provide a written continuation in the most natural way according to their Korean intuitions. They were also told to avoid any humor and not to look back and/or correct their previous answers. The entire experiment took approximately 30-40 minutes.

*3.2.1.4 Coding*

Participants' responses were coded by two native speakers of Korean, who were blind to the study purpose. Following the coding criteria established by Rohde, Kehler and Elman (2006), the coders annotated the participants' responses for intended reference of the subordinate subject in terms of *form* and *type*.

Reference form was coded as falling into one of the categories of "Pronoun," "Name," "Null subject," "Full NP" and "Other." A response was annotated as "Pronoun" when the subject referent was a pronoun. The reference form "Name" indicated a proper name in the subject position, referring to one of the names (e.g., *Phathieyse Heejungi Eunheelul kippukey hayessnuntey, waynyahamyen* ***Heejung****i Eunheeeykey senmwulul cwuesski ttaymwunita*, 'At the party, Heejung pleased Eunhee because **Heejung** gave Eunhee a present') or both names in the previous clause (e.g., *Cinancwuey Soheeka Jinmilul ccacungnakey hayessnuntey, waynyahamyen* ***Jinmi****wa* ***Sohee****nun angswukiki ttaymwunita*, 'Last week, Sohee annoyed Jinmi because **Jinmi** and **Sohee** are on bad terms with each other'). A response was coded as "Null subject" when an overt subject was missing (e.g., *Hoyuyhwuey Sangheeka Yejilul kkwucicessnuntey, waynyahamyen,* ***Ø*** *hoyuycwungey colasski ttaymwunita*, 'After the meeting, Sanghee scolded Yeji because **Ø** fell asleep during the meeting'). A response was coded as "Full NP" when the subject was a noun phrase other than a pronoun or a name (e.g., *Sophwungeyse Kihoka Minsoolul pwulewehayessnuntey, waynyahamyen* ***Minsoouy tosilak****i masisse poyesski ttaymwunita*, 'At the picnic, Kiho envied Minsoo because **Minsoo's lunch** looked delicious'). When a reference form corresponded to none of the forms listed above, it was coded as "Other."

Intended reference type was annotated as one of the categories of "Subj," "Obj," "Psub," "Pobj," "Ambi" and "Other." A response was coded as "Subj" when the subject referred to the

subject in the first clause (e.g. *Cinancwuey Soheeka Jinmilul ccacungnakey hayessnuntey, waynyahamyen **Sohee**ka sikkulewesski ttaymwunita*, 'Last week, Sohee annoyed Jinmi because **Sohee** was noisy') and "Obj" when the subject referred to the previous object (e.g. *Siktangeyse Eunsoki Hyunukul piwusessnuntey, waynyahamyen **Hyunuk**i papokathasski ttaymwunita*, 'At the restaurant, Eunsok ridiculed Hyunuk because **Hyunuk** looked like an idiot'). "Psub" and "Pobj" were used for subjects referring to a property or attribute of, respectively, the subject or the object of the previous clause. Referential expressions were annotated as "Psub" when they referred to the property of the previous subject (e.g. *Kongyencwungey Junghyeka Chanmilul kekcenghakey hayssnuntey, waynyahamyen **Junghyeuy khentisyen**i choyakiesski ttaymwunita*, 'During the performance, Junghye worried Chanmi because **Junghye's condition** was very bad') and as "Pobj" when they indicated a property of the previous object (e.g. *Pyengweneyse Kyungsooka Jinmolul pwukkulewehayessnuntey, waynyahamyen **Jinmouy oschalim**i nemwu helumhaysski ttaymwunita*, 'At the hospital, Kyungsoo was ashamed of Jinmo because ***Jinmo's clothes*** were shabby'). Responses were coded as "Ambi" when it was not clear which character the referent indicates (e.g. *Palphyocwungey Sumika Yujinul hweypangnwassnuntey, waynyahamyen **Ø** nemwu kuphaysski ttaymwunita*, 'During the presentation, Sumi interrupted Yujin because **Ø** was too hasty). This reference type was further divided into three subtypes of "Ambi-ambi", "Ambi-subj" and "Ambi-obj" depending on the degree of the ambiguity perceived by the annotator. The subtype "Ambi-ambi" indicated that it was totally ambiguous which protagonist the subject in the second clause referred to. The subtypes "Ambi-subj" and "Ambi-obj" indicated when the reference was ambiguous, yet it seemed more likely to refer to the previous subject (Ambi-subj) or previous object (Ambi-obj). Finally, the type "Other" included cases when the subject of the second clause referred to both subject and object in the previous clause (e.g., *Kyosileyse Seyoungi Suzylul culkepkey hayessnuntey, waynyahamyen **twul**un celchiniki ttaymwunita,* 'In the classroom, Seyoung amused Suzy because **both** were good friends') or to neither of them (e.g., *Eceyspamey Hyesooka Youngheelul mwusepkey hayessnuntey, waynyahamyen **ecey**nun hallowiniesski ttaymwunita*, 'Last night, Hyesoo frightened Younghee because **yesterday** was Halloween'). Participants' responses were also inspected in terms of semantic coherence based on criteria adopted from Cheng and Almor (2017).

After the two coders annotated the participants' responses, data were trimmed in the following steps. First, semantically incoherent or incomplete continuations were eliminated (4% of the entire data set). For analysis purposes, "Subj" and "Ambi-subj" responses, and "Obj" and "Ambi-obj" responses, were treated the same, respectively. That is, when a referent was annotated as "Ambi-subj" by one coder and as "Subj" by the other, it was categorized as "Subj." Likewise, when a referent was coded as "Ambi-obj" and "Obj" by the annotators respectively, it was finalized as "Obj."[12] Responses were excluded from further analysis if both coders annotated them as 'totally ambiguous' (0.03% of data), or if coders disagreed on reference (0.10%). Finally, given the small number of tokens annotated as "Psub" or "Pobj," these categories were collapsed with "Other" (7.3%). Inter-coder reliability was high ($\kappa = .998$).

*3.2.2 Results*

The overall distributions of referential form and intended reference (type) are summarized in Tables 3.1 and 3.2 respectively.

Table 3.1. Distribution of reference form in Experiment 1a

| | Overt pronoun | Name | Null subject | Full NP | Other | Disagree | Total |
|---|---|---|---|---|---|---|---|
| Token | 0 | 2373 | 156 | 236 | 0 | 0 | 2765 |
| Percentage | 0 | 85.8 | 5.6 | 8.5 | 0 | 0 | 100.0 |

Table 3.2. Distribution of reference type in Experiment 1a

| | Subject | Object | Other | Total |
|---|---|---|---|---|
| Token | 1142 | 1377 | 246 | 2765 |
| Percentage | 41.3 | 49.8 | 8.9 | 100.0 |

For referential form, the vast majority of responses (86%) consisted of a name. The remaining responses contained full NPs (9%) and null subjects (6%). Not a single overt pronoun

[12] Responses of this type accounted for 1% or less of the data in all experiments reported in this dissertation.

was produced in this task, confirming that overt pronouns would be unnatural in this context in Korean. Participants supplied a null pronoun response only about 6% of the time. Although this percentage is not high, it nevertheless indicates that a null pronoun is one possible referent form in this position, justifying the use of a free prompt in the sentence completion task. Among null referential expressions, the majority referred to the previous subject (89%), confirming that null subjects in Korean are strongly biased towards a subject antecedent (Kweon, 2011). Among names, reference to the previous object (57.3%) and subject (42.3%) were more evenly distributed.

For intended reference, the vast majority of responses (91%) referred to either the previous subject (41.3%) or object (49.8%). Only these responses are included in the following analyses, in which the proportion of subject reference out of all responses with either subject or object reference constitutes the measure of interest. Figure 3.1 illustrates subject bias thus calculated for subject-biased (SC and non-SC predicates) and object-biased (OB) items.

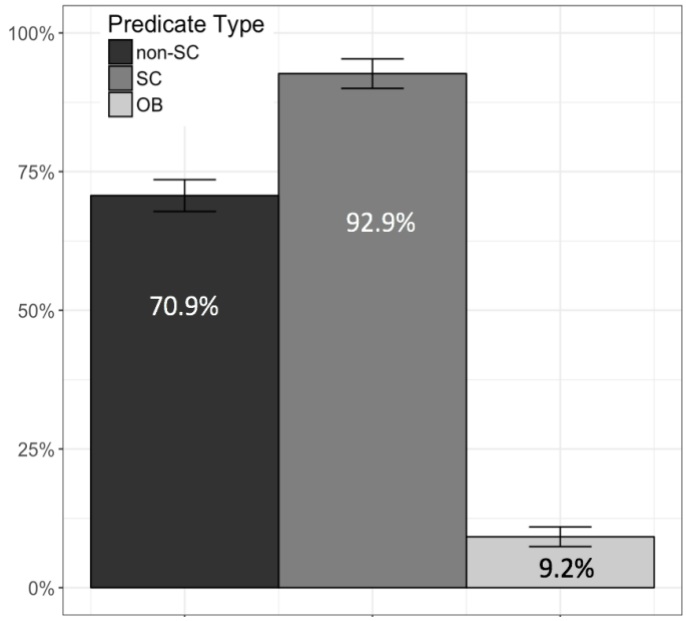


Figure 3.1. Mean percentage of subject bias in Experiment 1a; error bars indicate 95% CIs

Since the goal of the experiment is to test whether Korean speakers produce more subject reference following SC than non-SC predicates, only the subject-biased items (SC, non-SC) were included in statistical analyses. Proportion of subject bias was modelled using a mixed-effects

logistic regression model with the maximal random effects structure that converged (Baayen, 2008; Barr, Levy, Scheepers & Tily, 2013; Jaeger, 2008). The model included predicate type (SC vs. non-SC) as a fixed effect and participant and item as random effects (see Table 3.3 for model statement). All fixed effects were contrast-coded and centered. The model was created in R version 3.5.0 (R Development Core Team, 2009) using the lme4 package.[13]

As shown in Table 3.3, the model revealed a significant main effect of *Predicate type* ($b$ = 2.499, $SE$ = 0.505, $p < .001$), with more subject reference following SC (M = 93%, SD = 6%) than non-SC predicates (M = 71%, SD = 6%).

Table 3.3. Results of the mixed-effects logistic regression in Experiment 1a

| | *b* | *SE* | *p* |
|---|---|---|---|
| (Intercept) | 2.308 | 0.266 | < .001 |
| **Predicate type** | **2.499** | **0.505** | **< .001** |

*Note*. Model formula: glmer(type ~ predicate.type + (1+ predicate.type|participant) + (1|item)

*3.2.3 Discussion*

To address the first research question (RQ1), this experiment tested Hartshorne et al.'s (2013) hypothesis by investigating whether SC predicates in Korean, which contain explicit causative marking *-key ha*, induce a stronger remention bias toward a subject antecedent than do non-SC sentences, which lack such causative marking. The results of the sentence-completion task demonstrated that the Korean speakers did indeed mention previous subjects more frequently in the subject position of the following clause when they encountered SC predicates than when they read non-SC predicates in the preceding clause. However, these results should be interpreted with some caution. It is possible that the observed subject bias difference between the two predicate types may be due to differences in the lexical semantics of these predicates rather than the presence or absence of *-key ha*. It is conceivable, for example, that semantic properties of the predicates in the SC category unrelated to causative marking might have led to a stronger subject bias with these items. In order to further test whether it is the presence of explicit causative

[13] All the modelling in the experiments throughout this dissertation was carried out using the same package in the same version of R.

marking that drives this difference, all materials from Experiment 1a were translated into English as closely as possible. In Experiment 1b, native English speakers completed the same task with these English materials. If it is the causative marking in the SC predicates that led to the stronger remention bias in Experiment 1a, not other properties of these predicates, then the English translation counterparts of the Korean SC and non-SC sentences should exhibit little difference in their remention bias strength, since none of these English predicates contains causative marking.

### 3.3 Experiment 1b: L1 English written sentence-completion task

#### *3.3.1 Methods*

##### *3.3.1.1 Participants*

Thirty-five native speakers of English (age 18–29) participated in the English sentence-completion task. They were recruited among graduate and undergraduate students at the University of Hawai‘i. The language background questionnaire revealed that all participants had no or only basic knowledge of Korean, ruling out any possibilities that they had influence from the distinct bias patterns between SC and non-SC predicates in Korean when they completed the English version of the task.

##### *3.3.1.2 Materials and design*

Eighty sentence fragments served as the stimuli for Experiment 1b. As in Experiment 1a, the stimuli were presented as (11), in English this time.

(11) Last night, Eliza frightened Natalie because _____.

The experimental stimuli for the English sentence-completion task were obtained by translating the Korean sentences used in Experiment 1a into English as closely as possible. Since it was important to obtain translations for the verb in the particular sentence frame used in the experimental materials, independent translators translated the Korean sentences into English instead of simply using the English items originally entered into the NAVER dictionary when

creating materials for Experiment 1a. The translations were completed by four speakers fluent in both languages. A translation for each sentence was selected based upon the following criteria.

i. When at least three translators agreed upon the same translation (47/80 items).
ii. When two translators agreed and the other two each provided different translations, the translation that two agreed on was selected (20/80).
iii. In the case of a two-two tie (6/80) or disagreement among all four translators (7/80), the author used his own judgment to select among the translations provided.

As in Experiment 1a, the translated predicates were distributed in three types – 20 SC, 20 non-SC, 40 object-biased predicates. The predicates used for the stimuli in Experiment 1b are presented in Appendix A.

It should be noted that these predicate types were determined based on the status of their Korean counterparts. For the purpose of the current experiment, which is to probe whether the difference in remention bias between Korean SC and non-SC predicates found in Experiment 1a may have been due to properties other than explicitness of causative marking, it was important to maintain the English predicates in the current experiment as close as possible to their Korean translation counterparts. As a result, some predicates were used twice (e.g., *worry, embarrass, threaten, beg*) when the translators translated different items into the same English predicates. Also, some were multi-word predicates (e.g., *put someone in a bad mood, do better than someone*) as they were determined by the translators as the semantically closest translations for the Korean predicates.

Each predicate was presented with two NPs, one each in the subject position and the object position. As in Experiment 1a, the same gender of NPs appeared in each sentence fragment. Half of the stimuli included male protagonists and half included female protagonists. For the names of the protagonists, English names were selected from a previous remention bias study (Rohde & Attlinger, 2011) and among the most popular names for boys and girls provided by the U.S. Social Security Administration.

After the English stimuli were constructed based on the translation results and the name selection, the main clause in each sentence fragment (i.e. NP1 verbed NP2) was rated for

naturalness by two native speakers of English, who were blind to the purpose of the study and did not participate in any of the tasks. In addition to the 80 experimental sentences, 25 sentences were included as fillers for the naturalness test, which were adapted from previous studies (Hahn, 2011; Hawkins & Chan, 1997; Warren, McConnell, & Rayner, 2008). These fillers involved sentences with semantic violations (e.g., ?*At the restaurant, Jordan used a straw to drink a large lobster*), and sentences with grammatical violations including passivization of unaccusative verbs (e.g., **A few hours ago, my package was arrived*), resumptive pronouns in relative clauses (e.g., **A few days ago, the aunt whom I received a parcel from her left America*), and null subjects in subordinate clauses (e.g., **Everyday, the children played games when attended lessons*). The raters judged the naturalness of the experimental items and fillers on a scale of 1 (very unnatural) to 4 (very natural).

Results from the naturalness judgment demonstrated that all the fillers were rated as 1 or 2, indicating that the raters were able to detect the semantic and syntactic violations for the fillers. For the experimental items, most were rated as 3 or 4, confirming that they were natural. However, there were eight sentences that at least one rater judged to be unnatural (rated as 1 or 2). When a sentence was rated as 4 by one rater but as 1 by the other (1 case), this sentence was checked again with the rater who gave the score of 1, which turned out to be the rater's mistake. When a sentence was rated as 1, 2 or 3 by one rater and as 1 or 2 by the other (7 cases), it was replaced by one of the translations that I judged to be closest to their Korean counterparts. Then the same raters judged the naturalness of the replaced sentence using the same scale. The naturalness scores on these alternative sentences were either 3 or 4.

*3.3.1.3 Procedure*

The English sentence-completion task was conducted via a web-based interface. The web interface was used for the current experiment because this method allowed for more convenient data collection. Moreover, previous studies on remention bias report consistent results between traditional paper-and-pencil tasks and online survey tasks (e.g., Ferstl et al., 2011; Hartshorne & Snedeker, 2013).

Prior to the task, participants completed a language background questionnaire. During the task, participants were instructed to read sentence fragments and provide a natural continuation

in writing, avoiding humor. Sentence fragments were presented one by one on each page on the screen, and participants could advance to the next item at their own pace. The entire task took approximately 30-40 minutes.

*3.3.1.4 Coding*

Two native English speakers who were blind to purpose of the study annotated participants' responses for referential form and intended reference of the syntactic subject in the subordinate clause, using the same criteria as in Experiment 1a. Incomplete or incoherent responses (2% of all data), responses coded as 'totally ambiguous' (0.79%), and those where coders disagreed on intended reference (6%) were excluded from further analysis. Inter-coder reliability was high ($\kappa$ = .902).

*3.3.2 Results*

Tables 3.4 and 3.5 show the overall distributions of referential form and intended reference, respectively.

Table 3.4. Distribution of reference form in Experiment 1b

| | Pronoun | Name | Null subject | Full NP | Other | Disagree | Total |
|---|---|---|---|---|---|---|---|
| Token | 1332 | 1038 | 0 | 145 | 48 | 0 | 2563 |
| Percentage | 52.0 | 40.5 | 0 | 5.7 | 1.9 | 0 | 100.0 |

Table 3.5. Distribution of reference type in Experiment 1b

| | Subject | Object | Other | Total |
|---|---|---|---|---|
| Token | 1173 | 1163 | 227 | 2563 |
| Percentage | 45.8 | 45.4 | 8.9 | 100.0 |

For referential form, about half of all responses (52%) involved a pronominal subject, which is in sharp contrast to Experiment 1a where there were no responses with overt pronouns but 6% null subjects. These results are consistent with the cross-linguistic difference between Korean and English in terms of system of referential expressions: Korean prefers a null subject or a

repeated name over an overt pronoun in the subject position of the following clause in the context of the current experimental items, whereas this position is preferentially occupied by an overt pronoun in English. Among the responses with pronoun subjects, there was more reference to the previous subject (66.6%) than to the previous object (30.6%).[14] The large number of pronominal referents co-referenced with the subject antecedent suggest that overt pronouns in English usually refer to more salient and more accessible entities such as a previous subject (Arnold, 1998; Givón 1983; Gundel, Hedberg & Zacharaski, 1993). The pattern was different when the reference form was "repeated name," with 27.4% referring to the previous subject and 72.4% to the previous object.[15] This is consistent with previous work showing form-related referential biases in (L1 and L2) English (Arnold, 2001, 2010; Grüter et al., 2017; Kehler & Rohde, 2015).

Turning to intended reference, a majority of referents in participants' responses (91%) were coded either as subject (46%) or as object (45%). As in Experiment 1a, only these responses are included in the following analyses, in which the proportion of subject reference out of all responses with either subject or object reference constitutes the measure of interest. Figure 3.2 illustrates subject bias thus calculated for subject-biased (SC and non-SC predicates) and object-biased (OB) items.

[14] The remaining 2.8% of referents include "others."
[15] The remaining 0.2% of referents include "others."

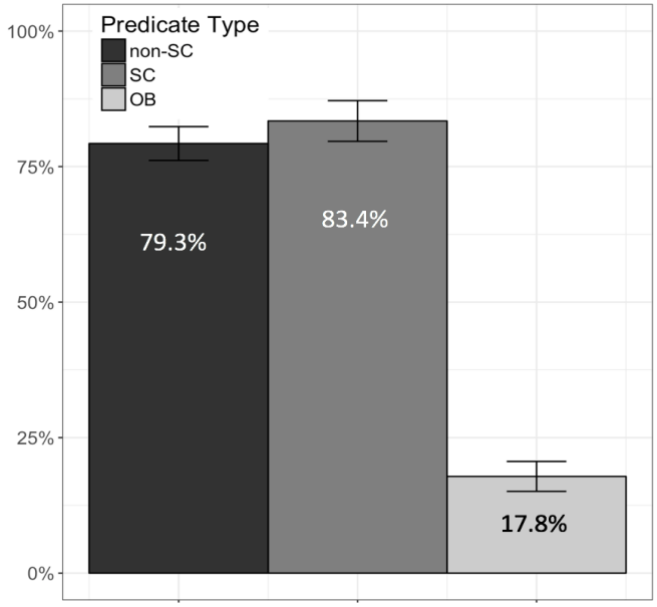


Figure 3.2. Mean percentage of subject bias in Experiment 1b; error bars indicate 95% CIs

As in Experiment 1a, a mixed-effects logistic regression model containing the maximal random effects structure that converged was fitted to these data. The model contained *Predicate type* (SC, non-SC) as a fixed effect (contrast-coded and centered), and participants and items as random effects. As shown in Table 3.6, there was no main effect of *Predicate type* ($b = 0.230$, $SE = 0.393$, $p = .558$). These findings contrast with the results of Experiment 1a, which exhibited a significant difference in subject bias between SC and non-SC predicates.

Table 3.6. Results of the mixed-effects logistic regression in Experiment 1b

| | *b* | *SE* | *p* |
|---|---|---|---|
| (Intercept) | 1.985 | 0.252 | < .001 |
| Predicate type | 0.230 | 0.393 | .558 |

*Note*. Model formula: glmer(type ~ predicate.type + (1+ predicate.type|participant) + (1|item)

In order to further investigate the differences between Experiments 1a and 1b, an additional mixed-effects logistic regression analysis was conducted with predicate type (SC, non-SC) and experiment (or language; 1a/Korean, 1b/English) as fixed factors (contrast-coded and centered),

and participants and items as random effects.[16] As shown in Table 3.7, the model revealed a main effect of *Predicate type* ($b = 1.238$, $SE = 0.313$, $p < .001$), indicating more subject reference following SC than non-SC predicates across the experiments, no main effect of experiment/language ($b = 0.227$, $SE = 0.211$, $p = .281$), indicating that the total subject bias did not differ significantly between experiments, and critically an interaction between these two factors ($b = 1.664$, $SE = 0.281$, $p < .001$). This interaction between predicate type and experiment provides further support for the interpretation of the findings from Experiments 1a and 1b, namely that SC and non-SC predicates in Korean differ in the strength of their subject bias whereas their English translation equivalents do not.

Table 3.7. Results of the mixed-effects logistic regression including Experiments 1a and 1b

| | *b* | *SE* | *p* |
|---|---|---|---|
| (Intercept) | 1.961 | 0.176 | < .001 |
| language | 0.227 | 0.211 | .281 |
| **Predicate type** | **1.238** | **0.313** | **< .001** |
| **Language × Predicate type** | **1.664** | **0.281** | **< .001** |

*Note.* Model formula: glmer(type ~ language*predicate.type + (1+ predicate.type|participant) + (1|item)

*3.3.3 Discussion*

To test the hypothesized role of explicit causative marking in the Korean SC construction in remention bias strength (RQ1), Experiment 1b was carried out to explore whether the stronger subject bias in Korean SC than Korean non-SC conditions observed in Experiment 1a could be attributable to differences between the items in the two conditions unrelated to causative marking. For Experiment 1b, the Korean sentences from Experiment 1a were closely translated into English, and English speakers provided written continuations for the English sentence fragments. The finding that there was little difference in terms of subject bias strength between the English

[16] When the model did not converge with the full random effects structure justified by the design, dropping the slope for language from the by-item random effects allowed the model to converge. The formula for the model is presented below in Table 3.7.

counterparts of Korean SC and non-SC predicates indicates that there appear to be no obvious differences apart from explicitness of causality marking between items in the two conditions that could have induced the difference of subject bias between SC and non-SC predicates in Experiment 1a. These findings suggest that the causative marking *-key ha* in the Korean SC predicates is the most likely cue that led to the difference, supporting Hartshorne et al.'s (2013) conjecture that explicit causative marking, as in Japanese –(*s*)*ase* or Korean *-key ha*, would elicit stronger remention biases.

However, it does not necessarily follow that causative marking is the *only* driving factor that affects remention bias: It remains possible that other factors beyond explicit causative marking in the predicates may have contributed to the observed differences. One such factor concerns potential differences at the level of *intentionality* among the predicates involved.[17] More specifically, it is possible that SC and non-SC predicates, categorized based on the explicitness of causality marking in Korean, also differ along the degree of intentionality encoded by the predicates in each language and condition, potentially affecting bias strength. Previous research shows that subject-biased predicates including Stimulus-Experiencer (SE, e.g., *surprise*) and Agent-Patient (AP, e.g., *threaten*) verbs have different remention bias strength by virtue of the fact that only an Agent, but not a Stimulus, can *intentionally* cause an event. For example, in a written sentence-completion study of remention bias in German and Norwegian, Bott and Solstad (2014) found a significantly stronger subject bias for SE than for AP verbs, which was taken to indicate that verbs associated with greater intentionality of the subject may induce a *weaker* subject bias. This bias pattern contingent on intentionality was also reflected in different types of causal explanations following SE versus AP predicates. AP verbs that involve an agent with an intention to act induced significantly more external or internal reasons than simple causes, whereas psych verbs involving a stimulus argument that has no volitional control of an evoked event elicited significantly more simple causes than external or internal reasons in sentence continuations (Bott & Solstad, 2014, see Section 2.1 for descriptions of each explanation type).

[17] This point was raised by an anonymous reviewer of a manuscript reporting Experiments 1 and 2 (Kim & Grüter, 2018).

Since this potential confound was not considered in the design of the materials for Experiment 1, it is difficult to rule out the possibility that SE and AP verbs were unequally distributed between SC and non-SC conditions as well as between Korean and English predicates. Considering that even cognates in typologically related languages like German and Norwegian can differ in terms of their status as SE or AP predicates (Bott & Solstad, 2014), it is plausible that the predicates employed in this study may also differ along the distribution of SE and AP predicates. Thus, in order to examine this possibility, I additionally investigated whether each verb in the experimental stimuli classifies as SE, AP, or both, based on Bott and Solstad's (2014) diagnostic tests. Following Bott and Solstad (2014), verbs were classified as AP when they allowed for the insertion of the adverbial *deliberately/ilpwule* in the frame *X Verbed Y,* and as SE when the *X Verbed Y* frame can be replaced with *it Verbed Y that,* where the proper name in the relevant position is replaced with a proposition that could otherwise be expressed in a subordinate *because*-clause (e.g., *Peter annoyed Mary because he sang loudly* → *It annoyed Mary that Peter sang loudly*). Verbs that passed both tests were categorized as ambiguous. It should be noted that using these diagnostic tests can only provide a coarse approximation to the issue of intentionality since the verbs categorized as ambiguous do not provide a clear clue as to the degree to which they are associated with AP or SE interpretations. It is also difficult to determine how strongly AP verbs are interpreted to be intentional, as they allow both intentional and unintentional interpretations. It would therefore be necessary to employ independent rating studies (in both languages), in which participants provide acceptability judgments on sentences with these predicates combined with the adverb "un/intentionally," to more precisely evaluate potential differences between the predicates with regard to intentionality.

Table 3.8 summarizes the results of the classification tests (see Appendix A for classification of individual items).

Table 3.8. Distribution of thematic verb types by Language/Experiment (Korean, English) and Predicate Type (SC, non-SC) following Bott and Solstad's (2014) diagnostics

| | | SC (*k*=20) | non-SC (*k*=20) |
|---|---|---|---|
| Experiment 1a (Korean) | AP: | 0 | 16 |
| | AP/SE: | 13 | 4 |
| | SE: | 7 | 0 |
| Experiment 1b (English) | AP: | 0 | 14 |
| | AP/SE: | 12 | 6 |
| | SE: | 8 | 0 |

*Note*. AP = Agent-Patient verbs; SE = Stimulus-Experiencer verbs

This analysis showed that for both languages, the majority of predicates in the non-SC condition were categorized as AP verbs, whereas all predicates in the SC condition allow an SE interpretation (SE or AP/SE). Given the stronger subject bias in SE verbs than AP verbs observed in Bott and Solstad, the different bias strength between SC and non-SC predicates in Experiment 1a (Korean) could thus be accounted for by different degrees of intentionality associated with these verbs, rather than by the presence of causative marking. On this explanation, however, it is difficult to interpret the results from Experiment 1b (English). Although the imbalance between AP and SE predicates was equally observed in both Korean and English predicates, there were no significant differences in bias strength between English SC and non-SC type predicates in Experiment 1b. If different degrees of intentionality associated with these predicates are a contributing factor to their remention bias strength, one should see the same effect of predicate type in English as observed in the Korean stimuli. It appears that although the level of intentionality is a likely factor that may have contributed to the differences observed in Experiment 1a, this factor alone is insufficient to explain the pattern of results across both Experiments 1a and 1b. Therefore, it may be concluded that explicitness of causality marking is an additional factor, leading to the cross-linguistic difference between the Korean and English predicates in this study. Future work is needed to tease apart these potential factors and determine how each of them plays a role in referential biases.

While it is difficult to pinpoint the precise reason(s) for the stronger subject bias in SC than non-SC predicates observed in Experiment 1a, the results from both experiments clearly demonstrated cross-language differences between Korean and English in the bias strength of the predicates investigated, which is critical for the investigation of cross-linguistic activation in L2 processing. Previous cross-linguistic work on rementionbiases has found quite consistent patterns across languages in the way that rementionverbs give rise to a certain direction of bias (subject- vs object-bias). For example, Hartshorne et al. (2013) demonstrated consistent patterns of rementionbiases for emotion verbs (stronger subject biases for experiencer-object verbs; stronger object biases for experiencer-subject verbs) across eight languages including Japanese, Mandarin, Russian, English, Spanish, Finnish, Dutch, and Italian. In contrast, more subtle differences have been observed between Norwegian and German, two closely related languages (Bott & Solstad, 2014). The findings from Experiment 1 provide further evidence of cross-linguistic differences, thus offering an ideal testing ground for probing the second research question (RQ2), namely whether Korean SC predicates are activated and affect referential choices when Korean learners of English process sentences in English. More specifically, it is predicted that if rementionbiases from Korean SC predicates cross-linguistically affect L2 referential choices in English, then Korean-speaking learners of English, but not native English speakers, will show a similar pattern observed in Experiment 1a, namely, a stronger bias to rementionthe subject following English translation correspondents of Korean SC predicates than English counterparts of Korean non-SC predicates. Results supporting this prediction will afford insight into the question of whether effects of cross-linguistic activation go beyond the word and construction level and affect sentence- and discourse-level processing (RQ 2). This issue was addressed in Experiment 2 where Korean-speaking learners of English provided written continuations in an English sentence-completion task.

# CHAPTER IV
# EXPERIMENT 2: REFERENTIAL BIASES IN L2 ENGLISH

## 4.1 Introduction

Drawing on the results from Experiment 1, Experiment 2 explores effects of cross-linguistic influence on L2 learners' referential choices in English with the following research question (RQ2, repeated below).

(RQ2) Do the cross-linguistic differences in syntactic and semantic structure of predicates affect Korean-speaking L2 learners' referential choices in English?

This research question is explored by investigating effects of cross-linguistic influence and effects of translation priming, as reflected in the two sub-questions:

(RQ2-1) Do Korean learners of English carry over remention bias from Korean predicates while making referential choices in English causal dependent clauses? (Effects of cross-linguistic influence)

(RQ2-2) Does completing the translation task preceding the sentence-completion task enhance the extent to which these learners carry over remention bias from Korean predicates? (Effects of translation priming)

To address these questions, a written English sentence-completion task analogous to Experiment 1b was conducted with Korean-speaking learners of English and a control group of native English speakers. The learners additionally completed a translation task, either preceding (T1 group) or following (T2 group) the sentence-completion task, in which they were asked to translate the English sentence fragments from the sentence-completion task into Korean. The purpose of the translation task was three-fold. First, it allows for a close examination of participants' understanding of the stimuli (e.g., Brysbaert, van Dyck, & van de Poel, 1999; Dijkgraaf, Hartsuiker & Duyck, 2017; Midgley, Holcomb, & Grainger, 2009). On the basis of participants' performance on the translation task, any responses provided in the sentence-

completion task with incorrect translations of the stimuli could be eliminated from analysis. The second purpose of the translation task was to investigate whether participants made the specific cross-language associations that I expected, i.e., translating an SC-type English predicate (classified based on the NAVER dictionary) into a Korean SC predicate with *-key ha*, and translating a non-SC type English predicate into a Korean predicate that does not contain *-key ha*. Finally, the translation task helps to explore the role of translation priming by examining whether completing the translation task immediately before the sentence-completion task leads to a stronger effect of cross-linguistic activation in learners' referential choices (RQ2-2).

Experiment 2 addresses these research questions by comparing group performance in the sentence-completion task in the following manners. On the one hand, effects of cross-linguistic activation at word and construction levels by Korean learners of English was investigated through a comparison between the control group and the T2 group (RQ2-1). Since both groups completed the same sentence-completion task in English, and the T2 group completed the sentence-completion task *before* the translation task, evidence of bias strength difference between non-SC and SC predicates found in T2, but not in the control group, would be an indication of effects of cross-linguistic influence for the learner group. On the other hand, comparing the T2 group with the T1 group in their performance on the sentence completion task enables us to inspect effects of translation priming (RQ2-2). Since only the T1 group had an opportunity to translate the stimuli from English to Korean prior to the sentence-completion task, results showing a greater difference between non-SC versus SC types for the T1 relative to T2 group will be taken as an effect of translation priming, i.e., enhanced L1 influence on referential choices in English due to cross-linguistic associations primed through a preceding translation task.

By testing whether cross-linguistic differences of remention bias strength between Korean and English predicates affect L2 learners' referential choices in English causal dependent clauses, this experiment extends the scope of research on L2 lexical access by exploring whether the effect of cross-linguistic activation at word and construction levels potentially influences L2 learners' pragmatic inferences as reflected in their referential choices in a separate clause. This experiment thus probes to what extent L2 learners' shared representations at word and constructional levels affect their processing at a discourse level.

## 4.2 Methods

### *4.2.1 Participants*

A total of 72 adult Korean-speaking learners of English (age 20–26) along with 34 adult native speakers of English (age 19–33) participated in this study. None of them had participated in Experiment 1a or Experiment 1b. Participants in the native speaker (NS) group were recruited among graduate and undergraduate students at the University of Hawai‘i. They reported that English was their first language, used dominantly since childhood. Participants in the learner group were recruited from colleges in Seoul, Korea. The mean onset of their exposure to English was 9.1 years, ranging from 6 to 12 years. Four of them reported having some experience living in English-speaking countries such as the USA, Canada and the Philippines, yet their length of stay in these countries was less than a year at the time of testing.[18] The remaining L2 participants had been exposed to English only in a classroom setting in Korea. The 72 L2 learners were randomly assigned to two subgroups, for which the order of the sentence-completion task and the translation task was manipulated. Half of the learners ($n$ = 36) completed the translation task first and then did the sentence-completion task (T1), and the other half ($n$ = 36) completed the sentence-completion task first followed by the translation task (T2). Details of the participants in each group are summarized in Table 4.1.

Table 4.1. Experiment 2: Participant information

| Group | Mean age | Mean years of studying English | TOEIC score (max= 990) | Self-ratings of overall English proficiency (1-10) |
|---|---|---|---|---|
| NS ($n$=34) | 21.1 (3.0) | - | - | 9.7 (0.4) |
| T1 ($n$=36) | 22.3 (1.1) | 9.3 (1.8) | 837.2 (90.6) | 6.1 (1.1) |
| T2 ($n$=36) | 21.8 (1.1) | 9.0 (1.8) | 807.2 (87.1) | 6.1 (1.2) |

*Note*. Numbers in parentheses indicate standard deviations

[18] Two learners had spent 2 months in the Philippines, one had spent 6 months in the USA, and the other had spent 9 months in Canada.

Learners' English proficiency was estimated based on the mean length of studying English, self-reported TOEIC® (*Test of English for International Communication*™) scores, and self-ratings of English proficiency. For the self-rated English proficiency, participants rated their overall proficiency in English on a scale of 1 (lowest) to 10 (highest). Independent samples t-tests on each of these measures revealed that the two L2 groups (T1, T2) did not differ significantly in the mean years of studying English ($t(70) = 0.794$, $p = .430$), TOEIC scores ($t(70) = 1.432$, $p = .157$), or self-ratings of English proficiency ($t(70) = 0.051$, $p = .960$). These results indicate that while there was a sizable gap in self-reported ratings between the NS and L2 groups, the two L2 groups were comparable in their general English proficiency.

### *4.2.2 Materials*

#### *4.2.2.1 Sentence-completion task*

For the experimental stimuli in the sentence-completion task, 36 English verbs (12 in SC, 12 in non-SC, 12 in OB) were selected in the following steps. To begin with, the 80 English predicates from Experiment 1b were screened for any multi-word predicates and predicates that were used more than once. Multi-word predicates are defined as a unit of one verb and one or more other words, which behaves as a single verb unit (Quirk, Greenbaum, Leech, Svartvik & Crystal, 1985). Among the 80 predicates, 9 multi-word predicates (e.g., *put someone in a bad mood*, *do better than*, *be suspicious of*) were removed, so that experimental stimuli include only lexical verbs (e.g., Bott & Solstad, 2014; Ferstl et al., 2011; Hartshorne & Snedeker, 2013). Further removed were predicates that occurred more than once in Experiment 1b. Recall that the stimuli in Experiment 1b were constructed by closely translating the Korean predicates in Experiment 1a into English, and as a result, there were 10 cases where two different Korean predicates were translated into the same English predicate. For these predicates, only one token was selected to make sure that none of the predicates for the current experiment appeared more than once. I also eliminated predicates that showed a weak remention bias strength either in Experiment 1a or in Experiment 1b, including NP1-biased predicates with a subject bias rate less than 60% and NP2-biased predicates with a subject bias rate above 40%. The Korean predicates that had a weak remention bias in Experiment 1a were three subject-biased predicates, *hyeppakhata* 'threaten' (subject bias rate of 35%), *apwuhata* 'flatter' (subject bias rate of 24%),

*keyekhata* ‘disobey’ (subject bias rate of 41%), and the object-biased predicate *cwukita* ‘kill’ (subject bias rate of 44%), and thus their English translation counterparts were eliminated. The English predicates that had a weak remention bias in Experiment 1b included three subject-biased predicates, *discourage* (subject bias rate = 41%), *threaten* (subject bias rate of 40%) and *interrupt* (subject bias rate of 56%), and five object-biased predicates, *help* (subject bias rate of 43%), *accuse* (subject bias rate of 44%), *kill* (subject bias rate of 55%), *complain* (subject bias rate of 57%), *stop* (subject bias rate of 50%), which were also eliminated. Finally, the remaining predicates were inspected for their inclusion in the vocabulary list in the English textbooks used in Korean middle and high schools and in the vocabulary list for the Korean SAT test. This criterion served to maximize the chances that L2 participants understood the lexical meaning of the English verbs in the current experiment. As a result, 5 predicates (*aggravate, humiliate, reproach, mistrust, belittle*) that are not included in the vocabulary list were further removed. Appendix B contains a list of all items in Experiment 2.

As in Experiment 1b, the two types of subject-biased verbs, SC and non-SC, were determined according to the status of their Korean translation counterparts by me, using the NAVER dictionary. When the Korean translation of an English verb contained *-key ha*, the English verb was classified as SC; when the Korean translation did not contain *-key ha*, the verb was classified as non-SC.

For the 24 subject-biased English predicates, bias strength was compared between SC and non-SC type predicates based on the results of the Korean sentence-completion task in Experiment 1a and the English sentence-completion task in Experiment 1b, in order to make sure that this subset of 24 predicates represent the stimuli in Experiments 1a and 1b. For this purpose, I conducted a mixed-effects logistic regression (glmer) on the subject bias rates in this subset of the data from Experiments 1a (Korean) and 1b (English) respectively, with predicate type (SC, non-SC) as a fixed factor (contrast-coded and centered) and participant and item as random factors. Table 4.2 summarizes results of the models in each language.

Table 4.2. Results of the mixed-effects logistic regression on the subset of SC and non-SC predicates selected for Experiment 2

| | | *b* | *SE* | *p* |
|---|---|---|---|---|
| Korean | (Intercept) | 2.288 | 0.314 | < .001 |
| | **Predicate type** | **1.352** | **0.568** | **.018** |
| English | (Intercept) | 2.047 | 0.283 | < .001 |
| | Predicate type | –0.094 | 0.410 | .818 |

*Note*. Formula for each model: glmer(type ~ predicate.type + (1+ predicate.type|participant) + (1|item)

The output of these models replicated the results of Experiments 1a and 1b: The Korean SC predicates had a stronger subject bias than the non-SC predicates did ($b$ = 1.352, $SE$ = 0.568, $p$ = .018), whereas the English counterparts of these predicates showed little difference in their subject bias rates ($b$ = –0.094, $SE$ = 0.410, $p$ = .818). These results established that the selected subset of predicates is representative of the items in Experiments 1a and 1b.

In addition to the 24 subject-biased and 12 object-biased predicates selected for Experiment 2, 12 non-remention-bias verbs adopted from Rohde et al. (2011) were included as distractors. The complete list of predicates used in Experiment 2 is provided in Appendix B.

These 48 English predicates were presented as in (12) in the sentence-completion task. Unlike the Korean and English sentence-completion tasks in Experiments 1a and 1b, the sentence fragments in Experiment 2 did not contain any adverbial adjunct. For each fragment, NPs were of the same gender: Half of the items included only male characters, and half included only female characters.

(12) Jacob amused Bill because _____.

*4.2.2.2 Translation task*

Experimental items for the translation task were obtained by taking the main clause from the sentence fragments in the sentence-completion task (without connective or continuation). A total of 36 items (12 in SC, 12 in non-SC, and 12 in OB) were presented as in (13).

(13) Jacob amused Bill.

### *4.2.3 Procedure*

The sentence-completion and translation tasks were completed via a web-based interface provided by Google Forms. Prior to the tasks, participants completed questions on their history and experience with English. Then they read task directions, which were presented on the screen in the participants' native language (English or Korean). Unlike the NS group, who only completed the sentence-completion task, half of the L2 participants completed the translation task *following* the sentence-completion task (T2 group), and the other half completed the translation task *preceding* the sentence-completion task (T1 group). There was a 5-minute break between the two tasks.

During the sentence-completion task, participants were asked to read a sentence fragment and complete the rest of the sentence. They were told to avoid any humorous responses. For the translation task, the Korean participants were asked to provide a Korean translation for each English sentence as accurately as they could. They were told to translate as much as they could if they did not understand all words. For both tasks, each item was presented on a separate screen, and participants were advised not to go back to previous pages and/or correct their responses once they completed them. Including the language background questionnaire, the entire sessions took approximately 20-30 minutes for the NS group and 60-80 minutes for the L2 groups.

### *4.2.4 Coding*

#### *4.2.4.1 Sentence-completion task*

Participants' responses in the sentence-completion task were annotated for intended reference form and type in the same manner as in Experiments 1a and 1b. Two coders participated in the annotation procedure – the main researcher, who is an advanced Korean-speaking learner of English, and a native speaker of English, who was blind to the purpose of the study. Incoherent or incomplete continuations (1% of all data), responses annotated as 'totally ambiguous' (0.1%), and items with inter-coder disagreement (1%) were excluded from further analysis. Inter-coder reliability was high ($\kappa = .980$).

*4.2.4.2 Translation task*

L2 participants' translations were annotated in terms of accuracy and presence of *-key ha*. For translation accuracy, two Korean speakers with advanced English proficiency – I and another native Korean speaker blind to the study purpose – coded participants' translations as correct or incorrect in terms of semantic similarity between the English predicate and its Korean translation. Translations coded as incorrect (10% of the L2 data, 10.6% in T1, 9.8% in T2) and translations with rater disagreement (1% of L2 data) were removed. For the annotation of *-key ha*, I coded participants' translations based on whether or not their responses contained this marking in the predicates, regardless of which predicate type the item had originally been classified as.

**4.3 Results**

Tables 4.3 and 4.4 present the distributions of referent form and type for the three groups.

Table 4.3. Distribution of reference form for each group in Experiment 2

| Group | | Pronoun | Name | Null subject | Full NP | Other | Disagree | Total |
|---|---|---|---|---|---|---|---|---|
| NS | Token | 665 | 467 | 0 | 33 | 33 | 3 | 1201 |
| (*n*=34) | Percentage | 55.4 | 38.9 | 0 | 2.7 | 2.7 | 0.3 | 100.0 |
| T1 | Token | 96 | 955 | 0 | 33 | 10 | 4 | 1098 |
| (*n*=36) | Percentage | 8.7 | 87.0 | 0 | 3.0 | 0.9 | 0.4 | 100.0 |
| T2 | Token | 119 | 959 | 0 | 38 | 1 | 2 | 1119 |
| (*n*=36) | Percentage | 10.6 | 85.7 | 0 | 3.4 | 0.1 | 0.2 | 100.0 |

Table 4.4. Distribution of reference type for each group in Experiment 2

| Group | | Subject | Object | Other | Total |
|---|---|---|---|---|---|
| NS | Token | 621 | 440 | 140 | 1201 |
| (*n*=34) | Percentage | 51.7 | 36.6 | 11.7 | 100.0 |
| T1 | Token | 538 | 486 | 74 | 1098 |
| (*n*=36) | Percentage | 49.0 | 44.3 | 6.7 | 100.0 |
| T2 | Token | 542 | 506 | 71 | 1119 |
| (*n*=36) | Percentage | 48.4 | 45.2 | 6.4 | 100.0 |

For reference form, the NS group showed a pattern distinct from the L2 groups in the use of pronouns and names. Overall, the NS group produced 55.4% pronouns and 38.9% names. In contrast, the L2 groups produced substantially fewer pronouns (8.7% in T1, 10.6% in T2), while using names about 86% of the time (87.0% in T1, 85.7% in T2). The pattern of the L2 groups is reminiscent of the results from Experiment 1a, where the Korean speakers produced names 85.8% overall. The overwhelming occurrence of names in the L2 data indicates that the learners carried over the pattern from their L1 Korean when they provided a referential expression in the subordinate subject position. The NS group, on the other hand, showed a similar pattern as in Experiment 1b: They used more pronouns (55.4%) than names (38.9%).

For reference type, all three groups demonstrated similar overall patterns. Reference to the previous subject or object constituted about 91% of all reference types across groups (88.3% in NS, 93.3% in T1, 93.6% in T2). Reference type was further analyzed by reference form. As shown in Table 4.5, when using a pronoun, the NS group referred to the previous subject more often (68.7%) than the previous object (20.8%), yet when they used a name, they referred to the previous object more often (64.7%) than the previous subject (35.1%). These results are consistent with previous findings that referents referring to the subject of the previous clause are expressed preferentially with pronominal forms, whereas non-subject referents are expressed preferentially with names (e.g., Stevenson et al., 1994; Arnold, 2001; Miltsakaki, 2007; Kehler, Kertz, Rohde & Elman, 2008; Fukumura & van Gompel, 2010; Rohde & Kehler, 2014). The L2 groups also referred to the previous subject more often (59.4% in T1, 60.5% in T2) than the previous object (11.5% in T1, 16.0% in T2) when they used a pronoun. Unlike the NS group,

however, when the L2 learners used a name, the reference type was evenly distributed across subject (50.4% in T1, 49.0% in T2) and object (49.6% in T1, 50.8% in T2).

Table 4.5. Distribution of reference type by pronoun and name for each group in Experiment 2

| Group | | Pronoun | | | Name | | |
|---|---|---|---|---|---|---|---|
| | | Subject | Object | Other | Subject | Object | Other |
| NS | Token | 457 | 138 | 70 | 164 | 302 | 1 |
| | Percentage | 68.7 | 20.8 | 10.5 | 35.1 | 64.7 | 0.2 |
| T1 | Token | 57 | 11 | 28 | 481 | 474 | 0 |
| | Percentage | 59.4 | 11.5 | 29.2 | 50.4 | 49.6 | 0 |
| T2 | Token | 72 | 19 | 28 | 470 | 487 | 2 |
| | Percentage | 60.5 | 16.0 | 23.5 | 49.0 | 50.8 | 0.2 |

As in Experiments 1a and 1b, subject bias was calculated by dividing the number of items with reference to the previous subject by the total number of items with reference to the previous subject or object.

For detailed analyses of the strength of subject bias in SC and non-SC items, the effects of predicate type (SC, non-SC) and group were investigated in three steps, each involving different types of data: (1) total data, (2) translation-consistent data, and (3) data with participant-driven categories. On analogy to the analyses in Experiments 1a and 1b, the analysis of total data included all participants' responses other than incorrectly translated items, ignoring the individual translations for each item in the translation task. For the analysis of translation-consistent data, items were included only when they had a translation consistent with the expected predicate type. In this process, data were excluded when items classified as SC had a translation without *-key ha* and when items classified as non-SC contained a translation with this causative marking. Finally, in the analysis of data with participant-driven categories, items were regrouped into SC and non-SC solely based on the participant's translation, ignoring the original category that the item had been assigned to. Each step of analysis has advantages and drawbacks. While the first analysis most closely follows the analyses in Experiments 1a and 1b, it does not take into consideration individual variability between learners with regard to cross-language

associations, which is likely to introduce noise in the results due to individual differences in learner translation. The second analysis helps reduce this noise, but the benefit of this analysis comes at the expense of excluding meaningful data points. This problem can be resolved by the third analysis, which includes all data points, but this analysis may lead to greater imbalance in the number of items per condition.

For statistical analyses, two separate models were created in each step of analysis in order to address the research questions about effects of cross-linguistic influence (RQ2-1) and translation priming (RQ2-2), one comparing between the NS and the T2 groups (RQ2-1), and the other between the T2 and the T1 groups (RQ2-2). For each comparison and analysis, a mixed-effects logistic regression model was fitted to the data, with group, predicate type (SC, non-SC), and their interaction as fixed effects (contrast-coded and centered), and participants and items as random effects. As in Experiment 1, all models included the maximal random effects structures allowed by the design (Baayen, 2008; Barr et al., 2013; Jaeger, 2008), but when a convergence problem arose, the random effects were simplified by removing the by-item slope for group. (The model statement for each comparison is provided in the results table for each analysis). Due to the three different analyses for each comparison, the alpha level was adjusted to .017 (.05/3).

### *4.3.1 Analysis 1: Total data*

First, all data were analysed regardless of presence of *-key ha* in participants' translations. Figure 4.1 illustrate these results.

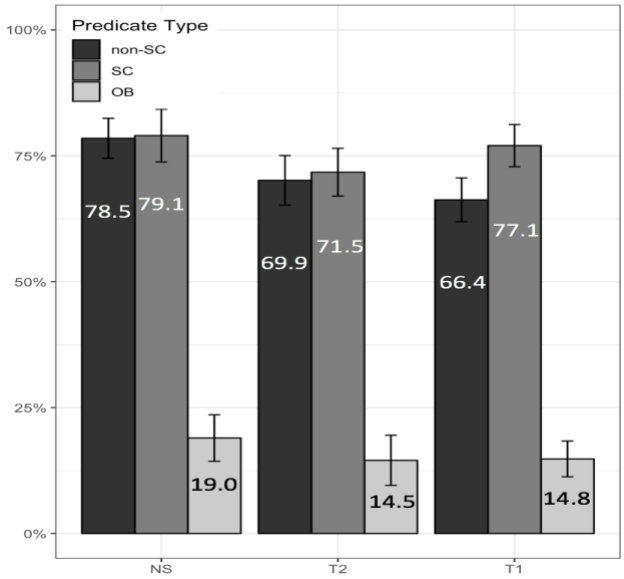


Figure 4.1. Mean percentage of subject bias in Experiment 2 (Analysis 1: Total data); error bars indicate 95% CIs

Mixed-effects logistic regressions (glmer) were conducted to assess the likelihood of the subject in participants' responses referring to the previous subject. Table 4.6 presents the model output for each comparison.

Table 4.6. Results of the mixed-effects logistic regression in Experiment 2: Analysis of total data

| | | *b* | *SE* | *p* ($\alpha$ = .017) |
|---|---|---|---|---|
| NS vs. T2 | (Intercept) | 1.441 | 0.205 | < .001 |
| | Group | –0.685 | 0.312 | .028 |
| | Predicate type | –0.065 | 0.357 | .856 |
| | Group × Predicate type | 0.226 | 0.478 | .636 |
| T2 vs. T1 | (Intercept) | 1.085 | 0.158 | < .001 |
| | Group | 0.045 | 0.231 | .847 |
| | Predicate type | 0.311 | 0.276 | .259 |
| | Group × Predicate type | 0.556 | 0.343 | .105 |

*Note*. Formula for each model: glmer(type ~ predicate.type*group + (1+ predicate.type|participant) + (1+ group|item)

Let us first focus on the comparison between NS and T2. The model showed a marginal effects of *Group* ($b$ = –0.685, $SE$ = 0.312, $p$ = .028) at the adjusted alpha level, with more subject reference in the NS group than in the T2 group. Such an effect was not predicted, but is potentially due to an overall preference for names (vs. pronouns) among Korean learners of English (see Tables 4.3 and 4.5 above), as names are generally less likely to be associated with subject antecedents (e.g., Stevenson et al., 1994; Arnold, 2001; Miltsakaki, 2007; Kehler et al., 2008; Fukumura & van Gompel, 2010; Rohde & Kehler, 2014; see also Grüter et al., 2017, for similar effects of referential form on the choice of reference by Korean and Japanese learners of English in a written sentence-completion task). There was no effect of *Predicate type* ($b$ = –0.065, $SE$ = 0.357, $p$ = .856), nor was there an interaction between the two ($b$ = 0.226, $SE$ = 0.478, $p$ = .636) at the adjusted alpha level. These results indicate that for both groups, the difference between SC and non-SC type predicates in terms of the likelihood of subject reference was not statistically significant, thus showing no evidence of cross-linguistic influence for this learner group in the analysis of total data.

For the comparison between T2 and T1, despite a numerically greater subject bias for SC than non-SC type predicates in the T1 group, as shown in Figure 4.1, there were no significant effects of *Group* ($b$ = 0.045, $SE$ = 0.231, $p$ = .847) or *Predicate type* ($b$ = 0.311, $SE$ = 0.276, $p$ = .259), nor a significant interaction between the two ($b$ = 0.556, $SE$ = 0.343, $p$ = .105). These results suggest no significant effects of cross-linguistic influence for both T1 and T2 groups when the total data is considered.

### *4.3.2 Analysis 2: Translation-consistent data*

For the second analysis, L2 participants' responses in the translation task were coded according to whether they were consistent with the original predicate type determined based on the presence of *-key ha* in their translations. Table 4.7 presents the distribution of consistent and inconsistent translations by predicate type in the L2 groups.

Table 4.7. Number (percentage) of translation-consistent and translation-inconsistent items in Experiment 2

| Group | Original predicate type | Consistent items | Inconsistent items | Total |
|---|---|---|---|---|
| T1 | Non-SC | 353 (97.8) | 8 (2.2) | 361 |
| | SC | 341 (92.9) | 26 (7.1) | 367 |
| T2 | Non-SC | 366 (97.3) | 10 (2.7) | 376 |
| | SC | 333 (90.7) | 34 (9.3) | 367 |

After removing translation-inconsistent items, subject bias for the remaining data was compared between groups in the same manner as in the analysis of total data. Results are illustrated in Figure 4.2, and model output is presented in Table 4.8.

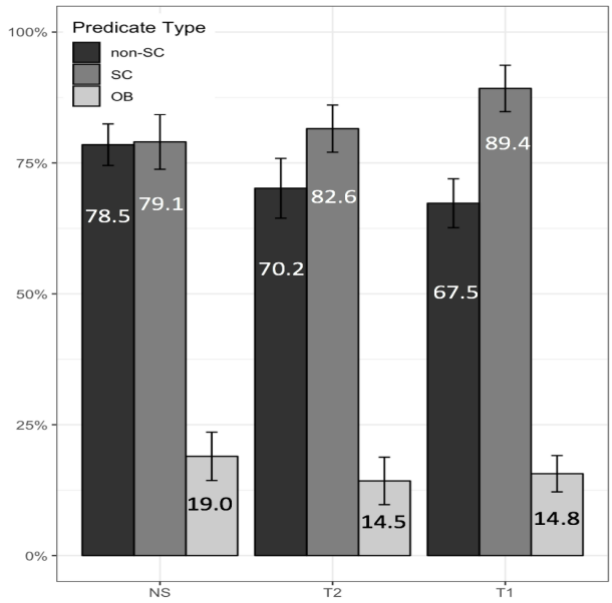


Figure 4.2. Mean percentage of subject bias in Experiment 2 (Analysis 2: Translation-consistent data); error bars indicate 95% CIs

In the model including the NS and the T2 groups, there was a numerical trend towards more subject reference for SC than for non-SC predicates in the T2 but not in the NS group (see Figure 4.2), and a marginal interaction emerged between *Group* and *Predicate type* at the adjusted alpha level ($b = 0.728$, $SE = 0.359$, $p = .042$). However, no effects of *Group* ($b = –0.240$, $SE = 0.253$, $p$

= .344) or *Predicate type* (*b* = 0.267, *SE* = 0.368, *p* = .468) reached significance. In the comparison between T2 and T1, however, the model revealed a significant effect of *Predicate type* (*b* = 1.137, *SE* = 0.364, *p* = .002), with more subject reference in SC than non-SC type predicates. There was no effect of *Group* (*b* = 0.176, *SE* = 0.215, *p* = .414). The interaction between the two factors did not reach significance at the adjusted alpha level (*b* = 0.704, *SE* = 0.349, *p* = .044). However, in light of the marginal interaction and in order to fully explore the research question regarding translation priming (RQ2-2), separate models were created for each group to examine potential differences in the effects of *Predicate type* between groups (with alpha further adjusted to .008; .017/2). These follow-up models yielded a significant effect of *Predicate type* within T1 (*b* = 1.500, *SE* = 0.002, *p* < .001) but not within T2 (*b* = 0.648, *SE* = 0.454, *p* = .153) group.

Taken together, the second analysis, which included translation-consistent items only, indicated differences between SC and non-SC predicates for T1, but not for T2 or the NS group. These findings provide some indication of cross-language influence, but only when cross-language associations were primed through an immediately preceding translation task.

Table 4.8. Results of the mixed-effects logistic regression in Experiment 2: Analysis of translation-consistent data

| | | *b* | *SE* | *p* (*α* = .017) |
|---|---|---|---|---|
| NS vs. T2 | (Intercept) | 1.579 | 0.206 | < .001 |
| | Group | –0.240 | 0.253 | .344 |
| | Predicate type | 0.267 | 0.368 | .468 |
| | Group × Predicate type | 0.728 | 0.359 | .042 |
| T2 vs. T1 | (Intercept) | 1.428 | 0.193 | < .001 |
| | Group | 0.176 | 0.215 | .414 |
| | **Predicate type** | **1.137** | **0.364** | **.002** |
| | Group × Predicate type | 0.704 | 0.349 | .044 |

*Note*. Formula for each model: glmer(type ~ predicate.type*group + (1+ predicate.type |participant) + (1|item)

*4.3.3 Analysis 3: Participant-driven analysis*

In this third analysis, L2 data were re-categorized into SC and non-SC depending on whether participants' translations were SC or non-SC, ignoring the original classification of predicate type. Table 4.9 presents the distribution of the recategorized data.

Table 4.9. Distribution of data after reassignment to participant-driven predicate type categories in Experiment 2

| Group | Original predicate type | Number (%) of items categorized as Non-SC | Number (%) of items categorized as SC | Total |
|---|---|---|---|---|
| T1 | Non-SC | 353 (97.8) | 8 (2.2) | 361 |
| | SC | 26 (7.1) | 341 (92.9) | 367 |
| T2 | Non-SC | 366 (97.3) | 10 (2.7) | 376 |
| | SC | 34 (9.3) | 333 (90.7) | 367 |

For the reorganized data, subject bias rate was calculated for each predicate type and group. These results are presented in in Figure 4.3.

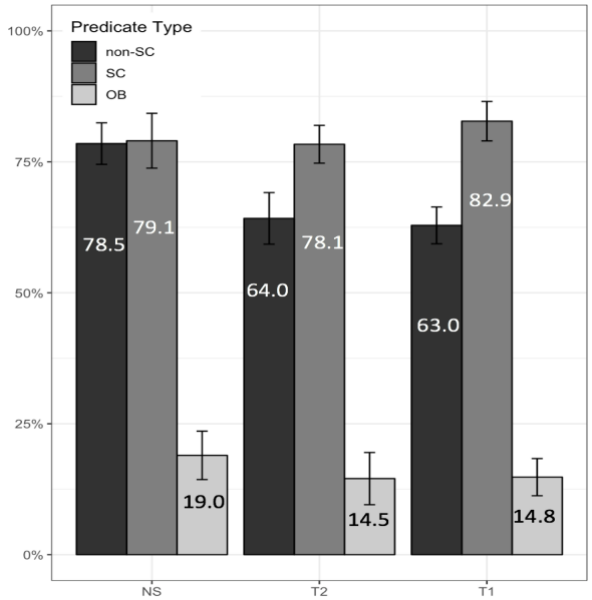


Figure 4.3. Mean percentage of subject reference in Experiment 2 (Analysis 3: Data with participant-driven categories); error bars indicate 95% CIs

As in the previous analyses, the likelihood of subject reference in SC and non-SC predicates was statistically compared between NS and T2 (effect of cross-linguistic influence) and between T1 and T2 (effect of translation priming), using mixed-effects logistic regressions. Table 4.10 summarizes the model output.

Table 4.10. Results of the mixed-effects logistic regression in Experiment 2: Analysis of data with participant-driven categories

| | | *b* | *SE* | *p* ($\alpha$ = .017) |
|---|---|---|---|---|
| NS vs. T2 | (Intercept) | 1.389 | 0.200 | < .001 |
| | Group | –0.505 | 0.249 | .043 |
| | **Predicate type** | **0.798** | **0.230** | **< .001** |
| | **Group × Predicate type** | **0.789** | **0.324** | **.015** |
| T2 vs. T1 | (Intercept) | 1.101 | 0.178 | < .001 |
| | Group | 0.089 | 0.205 | .664 |
| | **Predicate type** | **1.424** | **0.197** | **< .001** |
| | Group × Predicate type | 0.325 | 0.286 | .256 |

*Note*. Formula for each model: glmer(type ~ predicate.type*group + (1+ predicate.type |participant) + (1|item)

In the comparison between NS and T2, there was a marginal effect of Group ($b$ = –0.505, $SE$ = 0.249, $p$ = .043) at the adjusted alpha level, with more subject reference in the NS group than in the T2 group, presumably driven by the greater use of names in the T2 group (see above). There was also a main effect of *Predicate type* ($b$ = 0.798, $SE$ = 0.230, $p$ < .001), which was induced by higher subject bias for SC predicates than for non-SC predicates across these groups. More importantly, this effect interacted with *Group* ($b$ = 0.789, $SE$ = 0.324, $p$ = .015). To unpack this interaction, follow-up analyses were conducted within each group. A main effect of *Predicate type* was found in the T2 ($b$ = 1.230, $SE$ = 0.289, $p$ < .001), but not in the NS group ($b$ = 0.202, $SE$ = 0.527, $p$ = .700), indicating effects of cross-linguistic activation even in the absence of translation priming.

In the comparison between T1 and T2, the effect of *Predicate type* was significant ($b = 1.424$, $SE = 0.197$, $p < .001$), with more subject reference in the SC condition than the non-SC condition. There was no effect of *Group* ($b = 0.089$, $SE = 0.205$, $p = .664$) or interaction ($b = 0.325$, $SE = 0.286$, $p = .256$), suggesting that the effect of predicate type manifested similarly in both T1 and T2 groups. This result was further confirmed by follow-up analyses conducted within each group, which revealed a main effect of *Predicate type* in both the T1 ($b = 1.428$, $SE = 0.259$, $p < .001$) and the T2 group ($b = 1.230$, $SE = 0.289$, $p < .001$).

### 4.4 Discussion

In the current experiment, the effect of cross-linguistic influence was investigated by comparing referential choices in English causal dependent clauses between the native English-speaking control group and the L2 group who completed the sentence-completion task before the translation task, and the effect of translation priming was explored by comparing between the two groups of L2 learners who completed the sentence-completion and translation tasks in different orders. Participants' reference choices were analysed in three different ways, each of which produced somewhat different results. In the analysis of the total data, which was analogous to the analyses in Experiments 1a and 1b and did not take into account participants' actual cross-linguistic associations, no significant effects of cross-linguistic activation or translation priming were observed. In the second analysis, which included translation-consistent items only, there was some indication of the effect of cross-linguistic influence, but only for learners who were primed through a preceding translation task. In the third analysis, where data were reorganized to SC and non-SC categories according to participants' individual translations in the translation task, the effect of predicate type was present for both T1 and T2 groups, but not in the NS group, indicating an effect of cross-linguistic activation.

The analysis of the data reorganized according to participants' individual translations arguably provides a more precise picture than the other analyses with regard to how Korean participants were affected by cross-linguistic influence. This analysis not only included all data points, but also respected participants' individual cross-linguistic associations. Given that the primary goal of this experiment is to investigate the effect of cross-linguistic influence on L2 learners' referential choices in English, the analysis of data with participant-driven categories,

which takes individual variability in translation into account, may better test the role of cross-linguistic activation than the other two analyses. However, it is important to ask to what extent the analysis of data with participant-driven categories conducted here precisely captures L2 learners' cross-linguistic associations in general. One might object, for example, that the (T2 group) learners' translations were primed by their earlier completion of the sentences in the sentence-completion task. Since the stimuli of the translation task were constructed by taking the main clause portion of the items in the sentence-completion task, and there was only a 5-minute gap between the two tasks, it is possible that the participants who completed the translation task after the sentence-completion task may have been influenced by their prior exposure to the items in the sentence-completion task when they translated these items. This concern could be addressed by conducting a (substantially) delayed translation task, and comparing the results with those on the task conducted in the original test session. I implemented this additional measure in Experiment 3.

The findings of Experiment 2 indicate that properties of Korean predicates with regard to remention bias are activated when Korean learners of English process the English translation equivalents of these predicates, leading to stronger subject bias for SC-type predicates than for non-SC-type predicates even in English. These findings support the prediction that cross-linguistic activation at the word and construction level can affect learners' discourse processing.

# CHAPTER V
# EXPERIMENT 3: REFERENTIAL BIASES IN L2 ENGLISH: CONTROLLING VERB CLASS AND EXPLORING EFFECTS OF PROFICIENCY AND L2 LEARNING EXPERIENCE

## 5.1 Introduction

Building on the results of Experiment 2, which has provided empirical support for the effect of cross-linguistic activation of remention bias verbs in L2 referential choices, Experiment 3 seeks to address two potentially modulating factors that were not investigated in Experiment 2: L2 proficiency and L2 learning experience (RQ3). At the same time, it aims to replicate the findings from Experiment 2 with linguistic materials comprised of a more carefully selected set of interpersonal predicates.

While Experiment 2 provided evidence for effects of cross-linguistic activation on L2 reference choices, it has some methodological limitations, which I sought to address in Experiment 3. First, target verbs were systematically selected based on independently defined verb classes (e.g., Bott & Solstad, 2014; Hartshorne & Snedeker, 2013). In Experiment 2, subject- and object-biased items were simply chosen from a range of published studies on remention biases. For the current experiment, verbs were selected from consistent VerbNet classes (Kipper, Korhonen, Ryant & Palmer, 2008), an extended version of Levin's (1993) taxonomy of verb argument structure, to maintain verbs' semantic structures as similar as possible across both SC and non-SC conditions. Second, two independent norming studies – fill-in-the-blank and translation tasks – were conducted before finalizing the experimental stimuli. Through the fill-in-the-blank task, some verbs used in Experiment 2 were eliminated because they were not rated highly natural in conjunction with two human arguments (e.g., *cheat*). The translation task was used to gain a better understanding of L2 learners' general knowledge and interpretation of the target verbs, and to better estimate how often these verbs would be translated into SC or non-SC predicates in Korean. Third, all L2 participants completed the sentence completion task prior to the translation task. Since the results of Experiment 2 showed that the predicted effect of cross-linguistic activation in L2 referential choices emerged even in the absence of translation priming, with no significant interaction between translation priming

and predicate type, the translation priming factor was eliminated in Experiment 3. Fourth, in addition to a translation task conducted immediately after the sentence-completion task, a delayed translation task was administered to a subset of the L2 participants 9 months later. As in Experiment 2, there was a only 5-minute gap between the sentence-completion and translation tasks, which might induce between-task interference. For instance, it is possible that participants' translations for English sentences may have been influenced by their prior experience of completing these sentences in the sentence-completion task. In order to obtain participants' translations independent of possible influence from an immediately preceding sentence-completion task, participants' translations from both periods (5 minutes and 9 months after the completion task) were compared, and considered in data analysis.

In addition to these modifications, Experiment 3 included two additional factors – L2 proficiency and L2 learning experience – to investigate their potentially modulating effects on cross-linguistic activation of remention bias. Inclusion of these factors is particularly motivated by the gaps found in previous studies: There is little evidence of how proficiency and/or learning experience influence the way that cross-linguistic activation at word and construction levels affects discourse-level processing. To address these gaps, I measured participants' L2 proficiency and L2 learning experience, and included the measures in the modelling. For English proficiency, L2 participants completed two tasks: a lexical-decision test (LexTALE, Lemhöfer & Broersma, 2012) and a cloze test (Brown, 1980). To investigate the effect of L2 learning experience, two groups of L2 learners were recruited with different English learning experience – one group with only classroom exposure in Korea and the other with at least six months of natural exposure to English.

## 5.2 Methods

### *5.2.1 Participants*

Sixty-two Korean-speaking learners of English (NNS) participated in Experiment 3. To obtain variability in learner proficiency and English learning experience, 31 of the participants were recruited in Korea (NNS-KR) and 31 in the U.S. (NNS-US). The two L2 groups are closely matched in their age and onset of learning English, but exhibit significant differences in terms of English learning experience (see Table 5.1). Results from the language background questionnaire

showed that none of the participants in the NNS-KR group had any experience of staying in English-speaking countries, indicating that their English learning experiences are restricted to classroom instruction. In contrast, participants in the NNS-US group had been staying in the U.S. at least for 6 months by the time of testing (mean months of staying in the U.S = 44.2, Median = 20, Range: 6 -156). I set 6 months of staying in the U.S. as a criterion for inclusion in the NNS-US group in order to increase the effect of language experience, and since it has been found that among study-abroad students, the shift towards the L2 and inhibition of the L1 is not complete within a six-month period (Kroll, van Hell, Tokowicz & Green, 2010).

As measures of English proficiency, participants completed an English lexical decision task (LexTALE, Lemhöfer & Broersma, 2012) and a written cloze test (Brown, 1980). In the LexTALE, participants saw a series of letter strings (60 items in total) on a computer screen and were asked to decide whether or not the strings were an existing English word by clicking on the "yes" or "no" button on the screen. LexTALE provides participants' scores as the percentage of correct responses, "corrected for the unequal proportion of words and nonwords by averaging the percentages correct for these two item types" (LexTALE, Lemhöfer & Broersma, 2012, p. 329).[19] The cloze test consisted of three English paragraphs in which a total of 50 phrases had been replaced with blanks. Participants were asked to write the correct word or phrase for each blank during the cloze test. Following Brown (1980), participants' responses in the cloze test were scored using the "acceptable-answer scoring" method, which accepts not only the exact answer but also a finite set of contextually appropriate and grammatically correct answers provided in an answer key as correct (p. 311). In addition to LexTALE and the cloze test, participants were asked to rated their English proficiency on a scale of 1 to 10 for speaking, listening, reading, and writing. Independent samples t-tests on these scores revealed that the two learner groups were matched in each of their self-ratings for the four language skills (speaking: $t(60) = -1.090$, $p = .280$; listening: $t(60) < 0.001$, $p = .999$; reading: $t(60) = 0.393$, $p = .696$; writing: $t(60) = -1.422$, $p = .160$), but the NNS-US group had significantly higher scores in the cloze test ($t(60) = -2.566$, $p = .013$) and marginally higher scores in LexTALE ($t(60) = -1.781$, $p = .080$) than the NNS-KR group. While the averaged self-rating scores significantly, though only

[19] This measure is calculated as follows: ((number of words correct/40*100) + (number of nonwords correct/20*100)) / 2 (see http://www.lextale.com/scoring.html for more information of LexTALE scoring).

weakly to moderately, correlated both with the cloze-test scores ($r$ = .272, $p$ = .032) and with the LexTALE scores ($r$ = .417, $p$ = .001), the LexTALE scores only marginally correlated with the cloze-test scores ($r$ = .231, $p$ = .070). LexTALE scores correlated moderately and significantly with ratings of all four skills (speaking: $r$ = .432, $p$ < .001; listening: $r$ = .360, $p$ = .004; reading: $r$ = .370, $p$ = .003; writing: $r$ = .346, $p$ = .006), whereas the cloze-test scores only correlated significantly with ratings for speaking ($r$ = .279, $p$ = .028) and writing ($r$ = .303, $p$ = .017).

In addition to the learner groups, 40 native speakers of English (NS) were recruited from the student population at the University of Hawai'i and served as a control group. Information of participants in Experiment 3 is summarized in Table 5.1.

Table 5.1. Experiment 3: Participant information

| Group | Age | Onset of learning English | Self-ratings averaged over four skills (1-10) | LexTALE score | Cloze test score (max = 50) |
|---|---|---|---|---|---|
| NS ($n$=40) | 21.2 (3.1) | - | 9.8 (0.5) | - | - |
| NNS-KR ($n$=31) | 19.6 (0.8) | 8.9 (2.2) | 7.0 (1.5) | 70.8 (14.3) | 25.0 (8.8) |
| NNS-US ($n$=31) | 21.2 (2.0) | 8.7 (2.6) | 7.3 (1.7) | 76.7 (12.0) | 30.5 (8.1) |

*Note*. Numbers in parentheses indicate standard deviations

### *5.2.2 Materials*

#### *5.2.2.1 Sentence completion task*

Materials for the English sentence completion task consist of 24 experimental items with subject-biased verbs varying predicate type as determined by their Korean translation counterparts (SC, non-SC) along with 24 object-biased (OB) verbs as distractors. The target verbs were carefully selected to satisfy the following criteria. First, for subject-biased verbs, I started with the 220 verbs in class 31.1 (Stimulus-Experiencer verbs, e.g., *frighten*, *surprise*) in VerbNet (Kipper et al., 2008) and Levin's (1993) classification, and selected the 104 interpersonal verbs that are listed in English text books used in Korean middle and high schools and in the vocabulary list for the Korean SAT test. From this pool, 50 verbs were further removed because they allow for more than one possible interpretation, including, for instance,

verbs that can also be understood as nouns (e.g., *alarm, bug, delight, move*). In addition to these 54 subject-biased verbs, 26 OB verbs were selected from class 31.2 (Experiencer-Stimulus verbs, e.g., *admire, love*) and class 33 (Judgment verbs, e.g., *blame, thank*) verbs in VerbNet, which are also listed in the learner vocabulary list, making for a total of 80 verbs (54 verbs in class 31.1, 14 in class 31.2, and 12 in class 33). These verbs were inspected for naturalness through two norming studies: a fill-in-the-blank task and an English-to-Korean translation task.

a. Fill-in-the-blank task: This task was conducted to establish whether the selected verbs allow human referents in the subject and object positions. Twenty-three self-identified native speakers of English read sentences of the type illustrated in (14) and provided written answers in the blank. The position of the blank (subject vs. object) was counterbalanced within items and between participants, with each participant seeing only one version of each item, with an equal number of items across conditions. The task was completed via a web-based interface.

(14a) _____ amused Anna. (Subject-blank condition)
(14b) Leah amused _____. (Object-blank condition)

From the results of the fill-in-the-blank task, I selected verbs for which at least 3 out of 23 participants provided human referents in both subject and object positions, leaving 48 verbs in class 31.1, 12 in class 31.2, and 12 in class 33.

b. Translation task: This task was conducted (a) to assess whether L2 learners can understand the selected English verbs, and (b) to estimate how often the verbs are translated into SC or non-SC predicates in Korean. Twenty Korean learners of English (10 from U.S., 10 from Korea), who did not participate in any of the other experiments presented in this dissertation, translated English sentences containing the target verb and two human referents in the subject and object positions (e.g., *Eliza surprised Natalie*) into Korean. The task was completed on a web interface.

Participants' translations were coded by me in terms of translation accuracy and presence of *-key ha*. Translations were coded as accurate when they appeared as entries in NAVER English-Korean dictionary. Only verbs translated accurately at least 80% of the time in this task were

included in the experimental materials. In this process, 11 verbs in class 31.1, 2 in class 31.2, and 1 in class 33 were excluded, leaving 37 in class 31.1, 10 in class 31.2, and 11 in class 33. Also, the verbs in class 31.1 were categorized either as SC or non-SC type depending on the presence of *-key ha* in the participants' translations. Verbs translated into Korean predicates that contained *-key ha* 50% of the time or more were labelled as SC, and the remaining verbs whose Korean translation included *-key ha* less than 50% of the time were labelled as non-SC. From this process, 25 verbs were classified as SC and 12 as non-SC type.

From the verbs that were retained after the two norming tasks (25 in the SC, 12 in the non-SC, and 21 in the OB condition), I selected 12 out of the 25 remaining SC items, using information from published studies on strength of subject bias of individual items (e.g., Hartshorne & Snedeker, 2013) to try and approximate the mean (estimated) bias strength of the 12 remaining non-SC items. I thus removed 13 potential SC verbs that had a strong bias toward subject (more than 73%) in a previous study (Hartshorne & Snedeker, 2013). In order to attain the required 24 items for the object-biased condition, I also included three OB verbs that had been excluded due to translation accuracies lower than 80% but had accuracy scores close to the criterion (*despise*: 70% accuracy, *condemn*: 75%, *fear*: 75%). Thus the final item set consisted of 24 subject-biased verbs (12 in SC, 12 in non-SC), which served as experimental items, and 24 object-biased verbs, which served as distractors (see Appendix C for a complete list).[20] In order to assess whether the verbs in the experimental items are matched in terms of the degree of intentionality across conditions, the two diagnostic tests from Bott and Solstad (2014), '*deliberately*-insertion test' and '*that*-clause replacement test', were applied to the selected verbs in the same manner as in Experiment 1, classifying them as SE, AP, or both (see Section 3.3.3). Results showed that all the subject-biased verbs in the SC and non-SC types allowed both SE and AP interpretations, suggesting that the verbs in the two conditions are roughly matched in the degrees of intentionality.[21]

[20] Among these 48 verbs, 12 overlap with verbs used in Experiment 2 (4 in SC, 2 in non-SC, 6 in object-biased). See Appendix C for further details.

[21] Again, the results have to be taken with caution because these diagnostic tests allow for only a coarse-grained estimation of intentionality, as they do not provide specific information regarding preferences for SE versus AP interpretations.

These verbs were presented in a sentence frame, as illustrated in (15).

(15a) Leah amused Anna because _____. (SC type)
(15b) Rachel impressed Elizabeth because _____. (non-SC type)

*5.2.2.2 Translation task*

The items for the translation task consisted of the main clause portion of the 48 items (24 experimental items and 24 distractors) from the sentence completion task, presented as in (16).

(16) Leah amused Anna.

*5.2.3 Procedure*

Both NS and NNS groups completed the written sentence completion task, and the NNS group additionally completed the translation task, LexTALE and the cloze test. All tasks were completed via a web-based interface. The NS group completed only the sentence completion task individually at home. Participants in the NNS-KR group completed the tasks collectively in a quiet classroom under my supervision. Participants in the NNS-US group individually completed the tasks presented via a web-based interface. NNS completed all tasks in the following order: sentence completion, translation, lexical decision task, and cloze test. There was a 5-minute break between each task. In addition, some participants in the NNS-KR group completed a delayed translation task 9 months after the main tasks (see Section 5.3.5 for more information). Prior to these tasks, all participants completed a language background questionnaire.

*5.2.4 Coding*

*5.2.4.1 Sentence-completion task*

Participants' responses in the sentence-completion task were annotated for intended reference form and type in the same manner as in Experiments 1a, 1b and 2. Two coders participated in the annotation process – I and a native speaker of English, who was highly proficient in Korean. Incoherent or incomplete continuations (1.9% of all data), responses whose reference type was annotated as 'totally ambiguous' (0.2%), and items with inter-coder

disagreement (1.5%) were excluded from further analysis. Inter-coder reliability was high (κ = .970).

*5.2.4.2 Translation task*

The same annotators as in the sentence-completion task coded L2 participants' translations in terms of translation accuracy in the same manner as in Experiment 2. Participants' translations were coded as correct when they appeared as entries in the NAVER English-Korean dictionary. Translations coded as incorrect (4.7% of the L2 data, 5.7% in NNS-KR, 3.6% in NNS-US) and translations with rater disagreement (3.3% of L2 data, 3.8% in NNS-KR, 2.7% in NNS-US) were excluded from further analyses. In addition, I coded participants' translations for the presence of *-key ha*, regardless of which predicate type the item had originally been classified as.

**5.3 Results**

The distributions of referent form and type for the three groups are presented in Tables 5.2 and 5.3, respectively.

Table 5.2. Distribution of reference form for each group in Experiment 3

| Group | | Pronoun | Name | Null subject | Full NP | Other | Disagree | Total |
|---|---|---|---|---|---|---|---|---|
| NS | Token | 1074 | 647 | 0 | 61 | 77 | 9 | 1868 |
| | Percentage | 57.5 | 34.6 | 0 | 3.3 | 4.1 | 0.5 | 100.0 |
| NNS-KR | Token | 263 | 950 | 0 | 55 | 24 | 4 | 1296 |
| | Percentage | 20.3 | 73.3 | 0 | 4.2 | 1.9 | 0.3 | 100.0 |
| NNS-US | Token | 250 | 1016 | 0 | 57 | 16 | 2 | 1341 |
| | Percentage | 18.6 | 75.8 | 0 | 4.3 | 1.2 | 0.1 | 100.0 |

Table 5.3. Distribution of reference type for each group in Experiment 3

| Group | | Subject | Object | Other | Total |
|---|---|---|---|---|---|
| NS | Token | 678 | 942 | 248 | 1868 |
| | Percentage | 36.3 | 50.4 | 13.3 | 100.0 |
| NNS-KR | Token | 377 | 812 | 107 | 1296 |
| | Percentage | 29.1 | 62.6 | 8.3 | 100.0 |
| NNS-US | Token | 426 | 803 | 112 | 1341 |
| | Percentage | 31.8 | 59.9 | 8.3 | 100.0 |

For reference form, the NS and the NNS groups demonstrated similar patterns as those in Experiment 2. The NS group produced more pronouns (57.5%) than names (34.6%), whereas the NNS groups produced more names (73.3% in NNS-KR, 75.8% in NNS-US) than pronouns (20.3% in NNS-KR, 18.6% in NNS-US). The greater number of names than pronouns in the L2 data indicates transfer of the way that the learners choose referential form in their L1.

For reference type, all three groups mostly produced referents referring to previous subjects and objects (86.7% in NS, 91.7% in NNS-KR, 91.7% in NNS-US). Table 5.4 shows the distribution of subject or object reference by each reference form of pronoun and name. The NS group used a slightly greater number of pronouns referring to the previous subject than pronouns referring to the previous object, yet the number of continuations with subject reference greatly decreased with the use of names in the embedded subject position. In contrast, the L2 groups referred to the previous object slightly more often than the previous subject when they used a pronoun, and the gap between the object and subject reference was greater when they used a name.

Table 5.4. Distribution of reference type by pronoun and name for each group in Experiment 3

| Group | | Pronoun | | | Name | | |
|---|---|---|---|---|---|---|---|
| | | Subject | Object | Other | Subject | Object | Other |
| NS | Token | 506 | 468 | 100 | 172 | 474 | 1 |
| | Percentage | 47.1 | 43.6 | 9.3 | 26.6 | 73.3 | 0.2 |
| NNS-KR | Token | 104 | 135 | 24 | 273 | 677 | 0 |
| | Percentage | 39.5 | 51.3 | 9.1 | 28.7 | 71.3 | 0 |
| NNS-US | Token | 105 | 108 | 37 | 321 | 695 | 0 |
| | Percentage | 42.0 | 43.2 | 14.8 | 31.6 | 68.4 | 0 |

As in Experiments 1a, 1b and 2, the analysis focused on the subject-biased verbs (SC, non-SC types), and only responses with subject or object reference were included in data analysis. Participants' responses were entered into two mixed-effect logistic regression models for two comparisons. The first analysis compared NS and NNS groups in terms of differences between predicate types, including Predicate type (SC, non-SC) and Group (NS, NNS) as binary fixed effects. On analogy to Experiment 2, the strength of subject bias was analysed in three steps: (1) analysis of total data, (2) analysis of translation-consistent data, and (3) analysis of data by participant-driven category. For multiple comparisons corresponding to three different analyses, the alpha level was adjusted to .017 (.05/3). The second analysis focused on L2 participants only and tested for potentially modulating effects of English learning experience and proficiency by including either Group (NNS-KR, NNS-US) and proficiency measures (self-ratings, cloze-test scores, LexTALE scores, and combined z-scores averaged over the three proficiency measures) as fixed effects along with *Predicate type*. All models included the maximum random effects structure allowed by the design.

### *5.3.1 Analysis 1: Total data*

Figure 5.1 illustrates results from the analysis of the total data, which did not consider presence of *-key ha* in participants' translations.

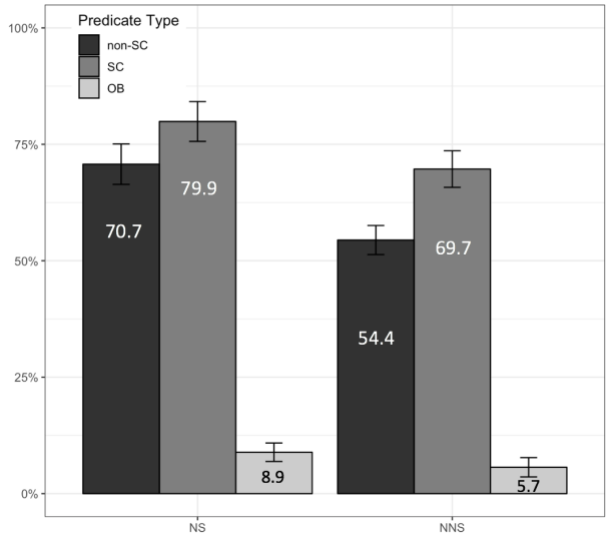


Figure 5.1. Mean percentage of subject bias in Experiment 3 (Analysis 1: Total data); error bars indicate 95% CIs

A mixed-effects logistic regression (glmer) was conducted to assess the likelihood of the subject in participants' responses referring to the previous subject. The results of the model are presented in Table 5.5.

Table 5.5. Results of the mixed-effects logistic regression in Experiment 3: Analysis of total data

| | | *b* | *SE* | *p* ($\alpha$ = .017) |
|---|---|---|---|---|
| NS vs. NNS | (Intercept) | 0.972 | 0.215 | < .001 |
| | **Group** | **0.717** | **0.222** | **.001** |
| | Predicate type | –0.770 | 0.413 | .063 |
| | Group × Predicate type | 0.324 | 0.376 | .388 |

*Note*. Model formula: glmer(type ~ predicate.type*group + (1 + predicate.type|participant) + (1 + group|item)

The model revealed a main effect of *Group* (*b* = 0.717, *SE* = 0.222, *p* = .001), with more subject reference in the NS than NNS group. As in Experiment 2, this effect of *Group* is likely due to the different distributions of referential forms between the two groups: The NNS group used more names and fewer pronouns compared to the NS group. Considering that names are

more strongly associated with non-subject reference, the greater use of names in the NNS group may explain the lower percentage of subject reference compared to the NS group (see Table 5.4).

Turning to other effects, although there was a numerical trend towards a stronger subject bias for SC versus non-SC type predicates in the NNS group (see Figure 5.1), no significant effect of *Predicate type* ($b$ = –0.770, $SE$ = 0.413, $p$ = .063) or interaction of the two factors ($b$ = 0.324, $SE$ = 0.376, $p$ = .388) was found, indicating that subject bias was not significantly different between SC and non-SC type predicates for either group. A separate analysis within each group (with alpha further adjusted to .008; .017/2) showed that there was no effect of *Predicate type* either for the NS ($b$ = –0.493, $SE$ = 0.506, $p$ = .330) or the NNS group ($b$ = –0.944, $SE$ = 0.424, $p$ = .026) at the adjusted alpha level. The absence of a robust predicate type effect in the NNS group (despite a numerical trend in the expected direction) is consistent with the results of Analysis 1 in Experiment 2.

### *5.3.2 Analysis 2: Translation-consistent data*

This analysis only included items for which L2 participants provided translations consistent with the predetermined predicate type (i.e., translations for SC type predicates with *-key ha*, and translations for non-SC type predicates without *-key ha*). Table 5.6 presents the distribution of consistent and inconsistent translations for predicate type in each learner group.

Table 5.6. Number (percentage) of translation-consistent and translation-inconsistent items in Experiment 3

| Group | Original membership of predicate type | Consistent items | Inconsistent items | Total |
|---|---|---|---|---|
| NNS-KR | Non-SC | 246 (81.5) | 56 (18.5) | 302 |
| | SC | 199 (71.8) | 78 (28.2) | 277 |
| NNS-US | Non-SC | 236 (77.9) | 67 (22.1) | 303 |
| | SC | 235 (78.6) | 64 (21.4) | 399 |

Subject bias after removing translation-inconsistent items was compared between NS and NNS in the same manner as in Analysis 1. Results are presented in Figure 5.2.

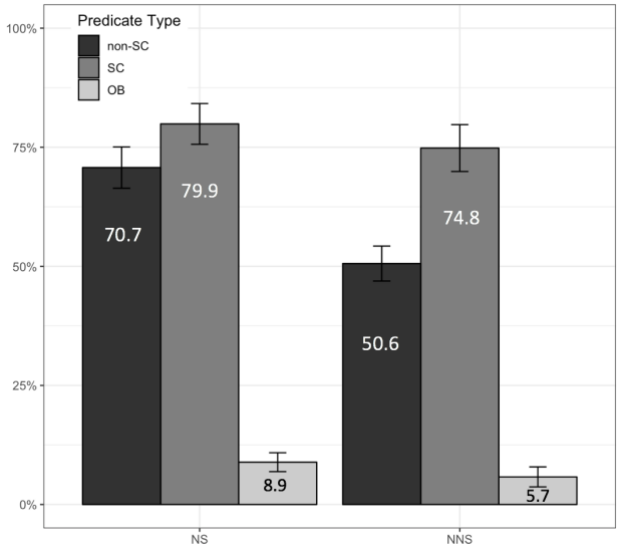


Figure 5.2. Mean percentage of subject bias in Experiment 3 (Analysis 2: Translation-consistent data); error bars indicate 95% CIs

A mixed-effects regression was fitted to this reduced dataset. Results of the model are summarized in Table 5.7.

Table 5.7. Results of the mixed-effects logistic regression in Experiment 3: Analysis of translation-consistent data

| | | *b* | *SE* | *p* ($\alpha$ = .017) |
|---|---|---|---|---|
| NS vs. NNS | (Intercept) | 1.014 | 0.212 | < .001 |
| | **Group** | **0.689** | **0.178** | **< .001** |
| | Predicate type | –0.973 | 0.409 | .017 |
| | **Group × Predicate type** | **0.685** | **0.282** | **.015** |

*Note*. Model formula: glmer(type ~ predicate.type*group + (1 + predicate.type|participant) + (1|item)

The model showed a main effect of *Group* (*b* = 0.689, *SE* = 0.178, *p* < .001), with more subject reference in NS than NNS, suggesting influence of referential forms in the L2 learners' choice of reference type. Also, a marginal effect of *Predicate type* (*b* = –0.973, *SE* = 0.409, *p* = .017) was found due to more subject reference in the SC than non-SC condition. Crucially,

there was a significant interaction between *Group* and *Predicate type* ($b$ = 0.685, $SE$ = 0.282, $p$ = .015). To unpack the interaction, separate models were created for each group to examine potential effects of *Predicate type* (with alpha further adjusted to .008; .017/2). A significant effect of *Predicate type* was found for the NNS ($b$ = –1.387, $SE$ = 0.494, $p$ = .005) but not the NS ($b$ = –0.493, $SE$ = 0.506, $p$ = .330) group.

### *5.3.3 Analysis 3: Participant-driven analysis*

Table 5.8 presents the distribution of the reorganized data in each predicate type for the L2 groups after re-categorizing their responses into SC and non-SC depending on whether participants' translations were SC or non-SC, irrespective of the original SC/non-SC categories.

Table 5.8. Distribution of data after reassignment to participant-driven predicate type categories in Experiment 3

| Group | Original predicate type | Number (%) of items categorized as Non-SC | Number (%) of items categorized as SC | Total |
|---|---|---|---|---|
| NNS-KR | Non-SC | 246 (81.5) | 56 (18.5) | 302 |
| | SC | 78 (28.2) | 199 (71.8) | 277 |
| NNS-US | Non-SC | 236 (77.9) | 67 (22.1) | 303 |
| | SC | 64 (21.4) | 235 (78.6) | 399 |

Figure 5.3 presents subject bias rates for each predicate type per group calculated based on the reorganized data.

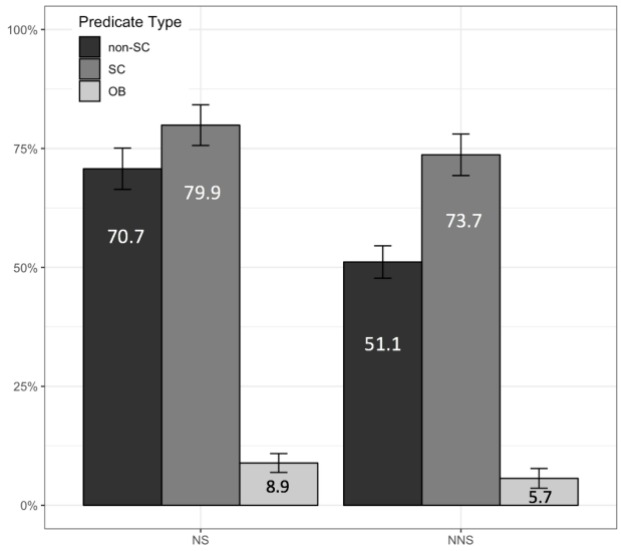


Figure 5.3. Mean percentage of subject reference in Experiment 3 (Analysis 3: Data with participant-driven categories); error bars indicate 95% CIs

Again, a mixed-effects logistic regression was used to compare the subject bias rates between SC and non-SC predicates across NS and NNS. The model results are summarized in Table 5.9.

Table 5.9. Results of the mixed-effects logistic regression in Experiment 3: Analysis of data with participant-driven categories

| | | *b* | *SE* | *p* ($\alpha$ = .017) |
|---|---|---|---|---|
| NS vs. NNS | (Intercept) | 0.968 | 0.207 | < .001 |
| | **Group** | **0.656** | **0.169** | **< .001** |
| | **Predicate type** | **–0.656** | **0.180** | **< .001** |
| | **Group × Predicate type** | **0.698** | **0.271** | **.009** |

*Note*. Model formula: glmer(type ~ predicate.type*group + (1 + predicate.type|participant) + (1|item)

The model revealed a significant effect of *Group* ($b$ = 0.656, $SE$ = 0.169, $p$ < .001) with more subject reference in NS than in NNS, again suggesting an effect of referential form. There was also a main effect of *Predicate type* ($b$ = –0.656, $SE$ = 0.180, $p$ < .001) with higher subject

rates for SC than for non-SC. These main effects were qualified by a significant interaction between the two factors ($b$ = 0.698, $SE$ = 0.271, $p$ = .009). Follow-up analyses within each group (with alpha further adjusted to .008; .017/2) showed that a main effect of *Predicate Type* was found in the NNS ($b$ = –0.960, $SE$ = 0.208, $p$ < .001), but not in the NS group ($b$ = –0.493, $SE$ = 0.506, $p$ = .330). These findings indicate that the NNS group, but not the NS group, produced more continuations with subject reference in the SC type than in the non-SC type sentences. The current results provide reaffirming evidence that the effect of cross-linguistic activation of remention predicates at a word and construction level extends to referential choices at a discourse level, replicating the findings from Experiment 2.

Based on the data with participant-driven categories, I further investigated modulating effects of L2 proficiency and language learning experience in the data from the L2 learners.

### *5.3.4 Effects of L2 proficiency and L2 learning experience*

To assess the role of L2 proficiency and English learning experience in L2 participants' referential choices, different types of proficiency measures – LexTALE scores, cloze-test scores, self-ratings (averaged across the scores in the four domains), and a single measure averaged over the three scores – and learning experience were added as additional factors to a series of exploratory models of the data with participant-driven categories. Learning experience was added as a categorical variable (NNS-KR, NNS-US), and each proficiency measure was added either as a continuous or as a categorical variable. When a proficiency measure was added as a continuous variable, the scores were normalized to z-scores before adding them to the model. When a proficiency measure was added as a categorical variable, participants were split into two proficiency groups based on a median split: Participants were assigned to the higher proficiency (NNS-H) group when their scores were the same as or higher than the median score and to the lower proficiency (NNS-L) group when their scores were below the median score for the measure under consideration. Table 5.10 shows results of grouping by the median score of each proficiency measure.

Table 5.10. Results of median-split for each proficiency measure

| | NNS-H | | NNS-L | |
|---|---|---|---|---|
| | Number of participants | Mean score (SD) | Number of participants | Mean score (SD) |
| LexTALE (median = 71.25) | 34 | 83.4 (7.5) | 28 | 62.0 (8.8) |
| Cloze test (median = 28) | 33 | 34.0 (6.0) | 29 | 20.7 (5.7) |
| Self-ratings (median = 7.5) | 32 | 8.4 (0.7) | 30 | 5.8 (1.1) |
| Combined z-score (median = –0.03) | 32 | 0.57 (0.42) | 30 | –0.60 (0.46) |

Before conducting statistical analyses, to obtain a general picture of the roles of proficiency and learning experience in participants' reference choices across conditions, the mean proportion of subject reference by predicate type was inspected for each proficiency group (NNS-H vs. NNS-L, based on the median split of the combined z-scores), as well as for each experience group (NNS-KR vs. NNS-US). As illustrated in Figures 5.4 and 5.5, both proficiency groups and both experience groups showed greater subject bias for SC than for non-SC type predicates.

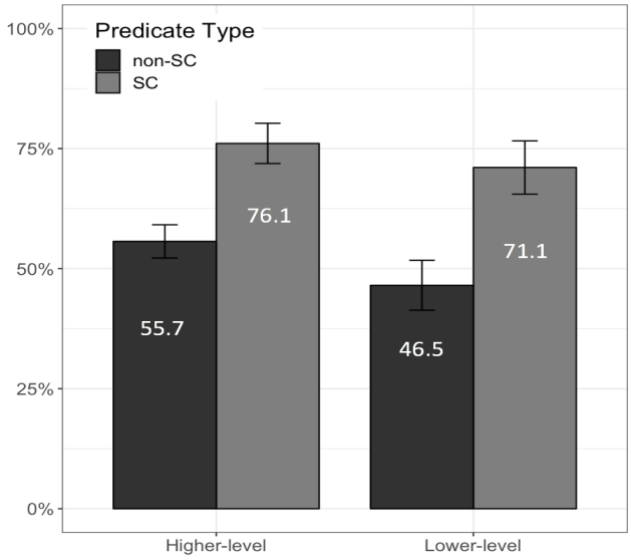


Figure 5.4. Mean percentage of subject reference for the higher- and lower-proficiency groups in Experiment 3 (Analysis of data with participant-driven categories); error bars indicate 95% CIs

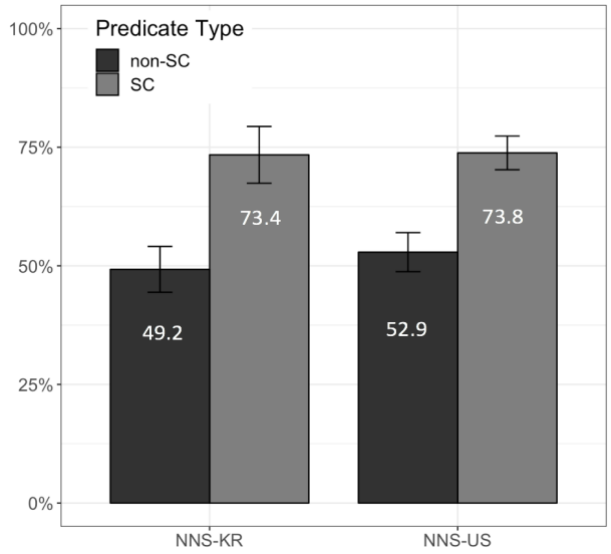


Figure 5.5. Mean percentage of subject reference for the NNS-KR and NNS-US groups in Experiment 3 (Analysis of data with participant-driven categories); error bars indicate 95% CIs

Each mixed-effects model included *Predicate type* (SC, non-SC), a proficiency measure, and *Experience* as fixed effects along with the maximal random effects structure allowed by the design. The by-item slope for *Experience* was dropped in case of a convergence problem. In addition to the models including all three factors, I also ran further exploratory analyses looking at each factor separately. These models included either one proficiency measure or experience as a fixed factor, in addition to predicate type.

Tables 5.11 and 5.12 show results of the three-factor models including LexTALE scores as a continuous (Table 5.11) or as a categorical variable (Table 5.12).

Table 5.11. Results of the mixed-effects logistic regression in Experiment 3: Language experience added as a categorical variable and LexTALE scores added as a continuous variable

| | *b* | *SE* | *p* |
|---|---|---|---|
| (Intercept) | 0.647 | 0.221 | .003 |
| LexTALE | 0.006 | 0.113 | .955 |
| **Predicate type** | **–0.906** | **0.206** | **< .001** |
| Experience | 0.162 | 0.215 | .451 |
| LexTALE × Predicate type | 0.284 | 0.161 | .078 |
| LexTALE × Experience | 0.321 | 0.224 | .152 |
| Experience × Predicate type | –0.003 | 0.305 | .992 |
| LexTALE × Predicate type × Experience | –0.295 | 0.319 | .354 |

*Note.* Model formula: glmer(type ~ lextale*predicate.type*experience + (1 + predicate.type |participant) + (1|item)

Table 5.12. Results of the mixed-effects logistic regression in Experiment 3: Language experience and LexTALE scores added as categorical variables

| | *b* | *SE* | *p* |
|---|---|---|---|
| (Intercept) | 0.633 | 0.223 | .005 |
| LexTALE | –0.142 | 0.226 | .529 |
| **Predicate type** | **–0.977** | **0.212** | **< .001** |
| Experience | 0.110 | 0.224 | .624 |
| LexTALE × Predicate type | –0.480 | 0.328 | .144 |
| LexTALE × Experience | –0.642 | 0.451 | .155 |
| Experience × Predicate type | –0.042 | 0.324 | .897 |
| LexTALE × Predicate type × Experience | –0.279 | 0.660 | .671 |

*Note.* Model formula: glmer(type ~ lextale*predicate.type*experience + (1 + predicate.type |participant) + (1|item)

The model output showed a main effect of *Predicate type* when LexTALE scores were included either as a continuous ($b$ = –0.906, $SE$ = 0.206, $p$ < .001) or a categorical variable ($b$ = –0.977, $SE$ = 0.212, $p$ < .001), with stronger subject bias in the SC than non-SC condition. There were no effects of *LexTALE* and *Experience*, or interactions between factors, except for a marginal interaction between *LexTALE* added as a continuous variable and *Predicate type* ($b$ = –0.284, $SE$ = 0.161, $p$ = .078). Separate exploratory models that only included LexTALE scores and predicate type as fixed factors showed a main effect of *Predicate type* with LexTALE scores added either as a continuous ($b$ = –0.955, $SE$ = 0.205, $p$ < .001) or a categorical variable ($b$ = –0.960, $SE$ = 0.206, $p$ < .001). There was no effect of *LexTALE*, but a marginal interaction was found between *LexTALE* added as a continuous variable and *Predicate type* ($b$ = 0.306, $SE$ = 0.156, $p$ = .050). However, separate analyses by each proficiency group split based on LexTALE scores showed a main effect of *Predicate type* both for NNS-H ($b$ = –0.706, $SE$ = 0.268, $p$ = .009) and for NNS-L group ($b$ = –1.441, $SE$ = 0.310, $p$ < .001). Also, adding an interaction of *LexTALE* and *Predicate type* to the model containing the two fixed effects did not improve overall model fit. The model including experience and predicate type as fixed factors showed a main effect of *Predicate type* ($b$ = –0.939, $SE$ = 0.209, $p$ < .001), with more subject reference in SC than non-SC condition. There was no effect of *Experience* ($b$ = 0.198, $SE$ = 0.226, $p$ = .382) nor any significant interaction of *Experience* and *Predicate type* ($b$ = 0.154, $SE$ = 0.314, $p$ = .625). A separate analysis within each experience group revealed a main effect of *Predicate type* both for the NNS-KR ($b$ = –1.196, $SE$ = 0.321, $p$ < .001) and NNS-US ($b$ = –0.839, $SE$ = 0.284, $p$ = .003) groups. Also, adding an interaction of *Experience* and *Predicate type* to the model containing the two fixed effects did not improve overall model fit. The main effect of predicate type without its interaction with LexTALE or experience indicate that the effect of cross-linguistic activation emerged in the L2 participants in general, irrespective of their LexTALE scores or learning experience.

Similar results were obtained when the cloze test scores were added as a proficiency measure. As shown in Tables 5.13 and 5.14, the models revealed only the main effect of *Predicate type*, driven by more subject reference in SC than non-SC condition, regardless of whether cloze test scores were added as a continuous ($b$ = –0.939, $SE$ = 0.213, $p$ < .001) or categorical variable ($b$ = –0.918, $SE$ = 0.207, $p$ < .001). There were no main effects of *Cloze test*

*scores* and *Experience*, and neither *Cloze test scores* nor *Experience* interacted with *Predicate type*. An exploratory model including only cloze test scores and predicate type as fixed effects showed only the main effect of *Predicate type*, regardless of whether cloze-test scores were added as a continuous ($b$ = –0.958, $SE$ = 0.210, $p$ < .001) or categorical variable ($b$ = –0.945, $SE$ = 0.206, $p$ < .001). As in the case of LexTALE scores, there was no main effect of proficiency, and the cloze test scores did not interact with *Predicate type*. Also, adding an interaction of *Cloze test scores* and *Predicate type* to the model containing the two fixed effects did not improve overall model fit. These results indicate that the difference of subject bias between SC and non-SC types in the L2 data was not significantly affected by the learners' cloze-test scores.

Table 5.13. Results of the mixed-effects logistic regression in Experiment 3: Language experience added as a categorical variable and cloze test scores added as a continuous variable

| | *b* | *SE* | *p* |
|---|---|---|---|
| (Intercept) | 0.672 | 0.222 | .002 |
| Cloze | 0.077 | 0.113 | .493 |
| **Predicate type** | **–0.939** | **0.213** | **< .001** |
| Experience | 0.116 | 0.223 | .603 |
| Cloze × Predicate type | 0.117 | 0.162 | .469 |
| Cloze × Experience | 0.127 | 0.226 | .574 |
| Experience × Predicate type | 0.067 | 0.319 | .834 |
| Cloze × Predicate type × Experience | –0.083 | 0.324 | .796 |

*Note.* Model formula: glmer(type ~ cloze*predicate.type*experience + (1 + predicate.type |participant) + (1|item)

Table 5.14. Results of the mixed-effects logistic regression in Experiment 3: Language experience and cloze test scores added as categorical variables

| | *b* | *SE* | *p* |
|---|---|---|---|
| (Intercept) | 0.674 | 0.220 | .002 |
| Cloze | –0.121 | 0.215 | .574 |
| **Predicate type** | **–0.918** | **0.207** | **< .001** |
| Experience | 0.136 | 0.214 | .524 |
| Cloze × Predicate type | –0.308 | 0.305 | .313 |
| Cloze × Experience | –0.469 | 0.429 | .275 |
| Experience × Predicate type | 0.099 | 0.304 | .745 |
| Cloze × Predicate type × Experience | 0.728 | 0.613 | .235 |

*Note.* Model formula: glmer(type ~ cloze*predicate.type*experience + (1 + predicate.type |participant) + (1|item)

The models including self-ratings (see Tables 5.15 and 5.16) revealed a main effect of *Predicate type*, whether the self-ratings were added as a continuous variable ($b = –0.934$, $SE = 0.205$, $p < .001$) or a categorical variable ($b = –0.941$, $SE = 0.208$, $p < .001$). There was no main effect of *Self-ratings* or *Experience*. No interaction emerged between any pair of factors, but there was a three-way interaction among *Self-ratings*, *Experience* and *Predicate type* when self-ratings were added as a continuous variable. To unpack this interaction, separate mixed-effects models were created for the NNS-KR and NNS-US groups, each including self-ratings (as a continuous variable) and predicate type as fixed effects. The alpha level was adjusted to .025 for the multiple comparisons. The model for the NNS-KR group showed a main effect of *Predicate type* ($b = –1.216$, $SE = 0.314$, $p < .001$) and no main effect of *Self-ratings* ($b = –0.149$, $SE = 0.159$, $p = .347$). Crucially, a marginal interaction was found between *Predicate type* and *Self-ratings* ($b = 0.654$, $SE = 0.295$, $p = .027$) at the adjusted alpha level, with a stronger effect of predicate type as participants' self-ratings were lower. These results indicate that the effect of cross-linguistic activation was greater for the learners with lower self-ratings than for those with higher self-ratings in the NNS-KR group. In contrast, the model for the NNS-US group revealed a main effect of *Predicate type* only ($b = –0.842$, $SE = 0.284$, $p = .003$), without any main effect

of *Self-ratings* ($b = 0.049$, $SE = 0.162$, $p = .763$) or interaction between *Predicate type* and *Self-ratings* ($b = 0.002$, $SE = 0.202$, $p = .993$).

Table 5.15. Results of the mixed-effects logistic regression in Experiment 3: Language experience added as a categorical variable and self-ratings added as a continuous variable

| | *b* | *SE* | *p* |
|---|---|---|---|
| (Intercept) | 0.692 | 0.222 | .002 |
| Self-ratings | –0.043 | 0.111 | .702 |
| **Predicate type** | **–0.934** | **0.205** | **< .001** |
| Experience | 0.144 | 0.216 | .503 |
| Self-ratings × Predicate type | 0.292 | 0.158 | .065 |
| Self-ratings × Experience | 0.158 | 0.223 | .479 |
| Experience × Predicate type | 0.109 | 0.299 | .716 |
| **Self-ratings × Predicate type × Experience** | **–0.643** | **0.316** | **.042** |

*Note.* Model formula: glmer(type ~ self-ratings*predicate.type*experience + (1 + predicate.type |participant) + (1|item)

Table 5.16. Results of the mixed-effects logistic regression in Experiment 3: Language experience and self-ratings added as categorical variables

| | *b* | *SE* | *p* |
|---|---|---|---|
| (Intercept) | 0.693 | 0.221 | .002 |
| Self-ratings | 0.114 | 0.216 | .598 |
| **Predicate type** | **–0.941** | **0.208** | **< .001** |
| Experience | 0.151 | 0.215 | .484 |
| Self-ratings × Predicate type | –0.314 | 0.310 | .311 |
| Self-ratings × Experience | –0.135 | 0.432 | .755 |
| Experience × Predicate type | 0.127 | 0.307 | .678 |
| Self-ratings × Predicate type × Experience | 0.886 | 0.615 | .150 |

*Note.* Model formula: glmer(type ~ self-ratings*predicate.type*experience + (1 + predicate.type |participant) + (1|item)

Finally, an integrated score combining the three proficiency measures was added to the original model, along with fixed factors of language experience and predicate type. For the combined score, z-scores from LexTALE, cloze-test and self-ratings were averaged for each participant (cf. Grüter et al., 2017). As shown in Tables 5.17 and 5.18, the models showed a main effect of *Predicate type*, with more subject reference in the SC condition than the non-SC condition, with z-scores included as a continuous variable ($b = –0.875$, $SE = 0.207$, $p < .001$) or a categorical variable ($b = –0.931$, $SE = 0.208$, $p < .001$). An interaction was only found between *Combined z-scores* and *Predicate type* when the scores were added as a continuous variable ($b = 0.440$, $SE = 0.216$, $p = .042$). This interaction indicates that the effect of predicate type was stronger as learners' combined z-scores were lower. Although there was no three-way interaction between *Combined z-scores*, *Predicate type*, and *Experience* in the model, I further explored whether the interaction between *Predicate type* and *Combined z-scores* also emerges for each language experience group. To this aim, a separate mixed-effects model including combined z-scores (added as a continuous variable) and predicate type as fixed effects was fitted to the data for the NNS-KR and NNS-US groups, respectively, with an adjusted alpha level of .025. The model for the NNS-KR group revealed a main effect of *Predicate type* ($b = –1.176$, $SE = 0.309$, $p$

< .001) but no effect of *Combined z-scores* ($b$ = –0.155, $SE$ = 0.204, $p$ = .447). There was a marginal interaction between *Predicate type* and *Combined z-scores* ($b$ = 0.768, $SE$ = 0.373, $p$ = .040) at the adjusted alpha level, indicating a trend toward a stronger effect of predicate type with decreasing combined z-scores in this group. The model for the NNS-US group also showed a main effect of *Predicate type* ($b$ = –0.840, $SE$ = 0.284, $p$ = .003) and no effect of *Combined z-scores* ($b$ = 0.204, $SE$ = 0.240, $p$ = .396). Unlike the NNS-KR group, there was no interaction between the two factors ($b$ = 0.164, $SE$ = 0.301, $p$ = .586) in this group.

Table 5.17. Results of the mixed-effects logistic regression in Experiment 3: Language experience added as a categorical variable and combined z-scores added as a continuous variable

| | *b* | *SE* | *p* |
|---|---|---|---|
| (Intercept) | 0.654 | 0.223 | .003 |
| Combined | 0.015 | 0.154 | .922 |
| **Predicate type** | **–0.875** | **0.207** | **< .001** |
| Experience | 0.152 | 0.220 | .489 |
| **Combined × Predicate type** | **0.440** | **0.216** | **.042** |
| Combined × Experience | 0.351 | 0.307 | .253 |
| Experience × Predicate type | –0.038 | 0.303 | .901 |
| Combined × Predicate type × Experience | –0.692 | 0.430 | .108 |

*Note.* Model formula: glmer(type ~ combined*predicate.type*experience + (1 + predicate.type |participant) + (1|item)

Table 5.18. Results of the mixed-effects logistic regression in Experiment 3: Language experience and combined z-scores added as categorical variables

| | *b* | *SE* | *p* |
|---|---|---|---|
| (Intercept) | 0.672 | 0.220 | .002 |
| Combined | –0.316 | 0.212 | .137 |
| **Predicate type** | **–0.931** | **0.208** | **< .001** |
| Experience | 0.122 | 0.210 | .562 |
| Combined × Predicate type | –0.240 | 0.310 | .439 |
| Combined × Experience | –0.558 | 0.419 | .183 |
| Experience × Predicate type | 0.098 | 0.309 | .752 |
| Combined × Predicate type × Experience | 0.249 | 0.617 | .686 |

*Note.* Model formula: glmer(type ~ combined *predicate.type*experience + (1 + predicate.type |participant) + (1|item)

In sum, the analyses exploring modulating effects of L2 proficiency and learning experience showed that although there was a weak trend toward a stronger effect of predicate type as learners' self-ratings and combined z-scores decreased in the NNS-KR group, the effect of predicate type did not interact with L2 proficiency or learning experience. These results indicate that the L2 groups carried over L1 properties in their interpretation of the English sentences, regardless of L2 proficiency and learning experience.[22]

[22] In addition to the effects of L2 proficiency and language learning experience, I also tested for the role of verb frequency in L2 learners' referential choices by including target verbs' frequency scores as an additional predictor. Each verb's frequency score was measured by the number of occurrences for the verb in the Corpus of Contemporary American English (COCA, Davies, 2009). The frequency score was added either as a continuous variable (z-transformed) or a categorical variable (median split). A logistic mixed-effects regression model including predicate type and word frequency (either as a continuous or categorical variable) as fixed effects, with the maximal random effects structure allowed by the design revealed only the main effect of *Predicate type* whether the model included verb frequency as a continuous ($b = 0.95$, $SE = 0.21$, $p < .001$) or as a categorical ($b = 0.97$, $SE = 0.22$, $p < .001$) variable. In both models, there was no main effect of *Word frequency*, and no interaction emerged between *Predicate type* and *Word frequency*. These results show that L2 participants consistently provided more subject reference in the SC than non-SC condition regardless of the frequency of the target verbs.

*5.3.5 Analysis of delayed translations*

In the current experiment, L2 participants completed the translation task only 5 minutes after the sentence-completion task. One may thus suspect that their translations may have been affected by their prior completion of the same items in the completion task. In order to test whether participants' translations were influenced by the preceding completion task, participants in the NNS-KR subgroup were contacted and invited to complete the same translation task again approximately 9 months after the original test session. Twenty-four of the original 31 participants in the NNS-KR group completed this delayed translation task.[23] Responses on the delayed task were annotated and analysed in the same manner as those on the original task (see Section 5.2.4.2). Translations coded as incorrect (8.5%) and translations with rater disagreement (4.5%) were excluded from further analyses. Then, I compared individual responses on the original and delayed task in terms of how often they were categorized as the same or different predicate types. Among participants' translations in the two tasks, 40% were both SC, 51% were both non-SC, and 9% belonged to different types between the tasks.

Next, the responses from this subset of participants in the sentence-completion task were entered into the same mixed-effects logistic regression model used in Analysis 3 (participant-driven analysis) above (Section 5.3.3), first coded for predicate type based on translations in the original translation task, and then coded for predicate type in the delayed translation task. Table 5.19 presents the distribution of consistent and inconsistent translations for predicate type in the original and the delayed-translation tasks.

Table 5.19. Number (percentage) of translation-consistent and translation-inconsistent items across original and delayed-translation tasks

| Task | Original predicate type | Consistent items | Inconsistent items | Total |
|---|---|---|---|---|
| Original | Non-SC | 199 (81.2) | 46 (18.8) | 245 |
| | SC | 154 (71.3) | 62 (28.7) | 216 |
| Delayed | Non-SC | 189 (77.1) | 56 (22.9) | 245 |
| | SC | 155 (71.8) | 61 (28.2) | 216 |

[23] Conducting the delayed translation task was not part of the original plan for the experiment. The 24 learners are those who were available 9 months after the original tasks.

As shown in Table 5.20, the results showed a main effect of *Predicate type* for coding based on both the original ($b$ = 1.121, $SE$ = 0.351, $p$ = .001) and the delayed translation tasks ($b$ = 0.801, $SE$ = 0.298, $p$ = .007).

The findings that there was a significant amount of overlap between translations between the original and delayed translation tasks and that the effect of predicate type still emerged in the analysis based on the delayed translation task indicate that there was little influence of the sentence completion task immediately preceding the translation task.

Table 5.20. Results of the mixed-effects logistic regression in Experiment 3: Subset of participants in the original and delayed-translation tasks

| | | *b* | SE | *p* ($\alpha$ = .017) |
|---|---|---|---|---|
| Original task* | (Intercept) | 0.377 | 0.212 | .075 |
| | **Predicate type** | **1.121** | **0.351** | **.001** |
| Delayed task† | (Intercept) | 0.330 | 0.199 | .098 |
| | **Predicate type** | **0.801** | **0.298** | **.007** |

***** Model formula: glmer(type ~ predicate.type + (1 + predicate.type|participant) + (1|item)

† Model formula: glmer(type ~ predicate.type + (1 + predicate.type|participant) + (1|item)

**5.4 Discussion**

The current experiment investigated effects of cross-linguistic differences of remention bias strength between Korean and English predicates in L2 learners' referential choices in English with the same semantic classes of verbs across SC and non-SC conditions, and explored modulating roles of L2 proficiency and L2 learning experience in this process (RQ3).

The results of the sentence-completion task replicated the results from Experiment 2: The NNS group produced more continuations with subject reference in the SC condition than the non-SC condition, whereas the NS group showed little difference between the predicate types, suggesting that the effect of cross-linguistic activation of remention predicates at word and construction levels extends to referential choices at a discourse level. The group difference observed in this experiment is unlikely to be attributable to aspects of verbs' properties other than the cross-linguistic difference, such as different verb classes between SC and non-SC

conditions or across Korean and English. One of the limitations underlying Experiments 1 and 2 was that verb semantics was not properly controlled between SC and non-SC predicates, thus raising the possibility that the difference of remention bias between the conditions in the Korean data (Experiment 1a) and the English L2 data (Experiment 2) may be due to other factors than the causative marker *-key ha* in the SC condition. To address this problem, Experiment 3 adopted verbs from the same semantic class (i.e., class 31.1) based on VerbNet (Kipper et al., 2008). Given that remention bias is assumed to be dependent upon a verb's semantic structure (Bott & Solstad, 2014; Hartshorne & Snedeker, 2013; Hartshorne et al., 2015; Solstad & Bott, 2017) and that the current experiment kept verb class constant across SC and non-SC conditions, the present findings indicate that causative marking is likely to be the driving factor behind the differences between NS and NNS. Since the presence of causative marking in the SC items induced more subject reference in the SC condition than the non-SC condition for the learners, but not for the native control group, the current results provide clear evidence of effects of cross-linguistic activation in discourse-level processing.

Another important finding of this study is that the effects of cross-linguistic influence were found for L2 learners irrespective of L2 proficiency or L2 learning experience. When proficiency measures from LexTALE, cloze-test, self-ratings, or combined z-scores (either as a continuous or categorical variable) and language experience were added to a series of models, most of the models did not show any interactions of proficiency or experience with the L2 learners' continuation patterns. A weak trend toward a stronger effect of predicate type with decreasing proficiency was found in the NNS-KR group, but the effect of predicate type was not modulated by experience.

The robust effects of cross-linguistic influence independent of L2 proficiency indicate that learners at varied proficiency levels carried over L1 properties in their referential choices in English. However, a caveat is required in generalizing the lack of L2 proficiency effect in the current study since it included learners with a limited range of English proficiency. While attempts were made to obtain variability in learner proficiency by including both immersed (NNS-US) and non-immersed (NNS-KR) learners, the participants in the NNS-US group were not highly proficient, as their self-ratings and LexTALE scores did not significantly differ from those in the NNS-KR group. It is possible that effects of proficiency might emerge if learners

with much higher proficiency were included. In this regard, further research is necessary with L2 samples with a wider range of proficiency in order to more clearly assess effects of proficiency on the degree of cross-linguistic influence in L2 referential choices.

In addition, there was no evidence of an effect of L2 learning experience in cross-linguistic activation. Both NNS-US and NNS-KR groups provided more continuations with subject reference in the SC condition than the non-SC condition. These results are in line with other studies demonstrating parallel activation of words between languages either for learners with only classroom experience or for learners immersed in an L2 (e.g., Lemhöfer et al., 2004; Poarch & Van Hell, 2012; Weber & Cutler, 2004). In contrast to some previous findings showing that a substantial amount of naturalistic exposure to a target language permits L2 learners to show more target-like sentence processing (e.g., Dussias, 2003; Dussias & Piñar, 2010; Dussias & Sagarra, 2007; Frenck-Mestre, 2002; Pliatsikas & Marinis, 2013), the current results indicate that the effect of cross-linguistic activation in discourse-level processing applied equally to the learners regardless of their English learning experience.

In summary, results of Experiment 3 replicated the findings of Experiment 2, showing robust effects of cross-linguistic influence in L2 learners' referential choices, yet no evidence was found for interacting effects of L2 proficiency or language experience. Although more research should be conducted to investigate each of these factors more closely by including L2 participants with a wider range of proficiency and L2 learning experience, these findings suggest that the consequences of cross-linguistic activation extend beyond a word and a construction level in L2 learners with varied proficiency and L2 learning experience.

# CHAPTER VI
# EXPERIMENT 4: ONLINE PROCESSING OF REFERENTIAL BIASES IN L2 ENGLISH

## 6.1 Introduction

Experiment 4 draws on findings and insights from the offline data in Experiment 3 and investigates whether the effect of cross-linguistic activation extends to the L2 online processing of remention biases (RQ4). To this end, a visual-world eye-tracking experiment was designed, following the general approach taken by Itzhak and Baum (2015) in a recent visual-world study on L1 speakers' processing of remention bias in English. While the main focus of their study was on testing the effect of prosody on native speakers' use of remention bias information, Itzhak and Baum (2015) also conducted a separate analysis on the data from their 'No-Accent' condition, which focused exclusively on native speakers' processing of remention bias. Experiment 4 follows the basic design and analysis of their 'No-Accent' condition. By including both SC and non-SC items in the present experiment, I aim to test for effects of cross-linguistic activation in L2 learners' use of remention biases in online processing, in addition to contributing to the more general investigation of L2 learners' ability to use remention bias information (NP1- vs NP2-bias) during real-time listening.

As reviewed in Section 2.1.3, previous studies have shown some evidence that L2 learners can use remention bias information during online processing (Contemori & Dussias, 2018; Liu & Nicol, 2010). However, these studies focused exclusively on highly proficient L2 learners, such as highly proficient bilinguals with early exposure to English (Contemori & Dussias, 2018) and advanced ESL learners (Liu & Nicol, 2010), leading to the question as to whether less proficient, non-immersive L2 learners can use remention bias information during online processing, and whether there is any modulating effect of L2 proficiency. Moreover, the extent to which cross-linguistic activation influences L2 processing of remention biases has not been previously investigated. This experiment seeks to address these gaps in the existing literature and investigate whether less proficient L2 learners can use remention bias information, whether co-activation of English and Korean remention bias in predicates affects their use of remention bias information

during real-time processing, and whether there are any modulating effects of L2 proficiency in this process.

## 6.2 Methods

### *6.2.1 Participants*

A total of 56 Korean-speaking L2 learners (NNS) and 56 native speakers (NS) of English participated in this experiment. None of them participated in Experiments 1, 2 or 3. Data from four participants in the NNS and four participants in the NS group were excluded from the analysis, due to eye-tracker calibration difficulty (1 in NS), equipment malfunction (1 in NNS), and low proportion of fixations in the eye gaze record (3 in NS; 3 in NNS, see Section in 6.2.3 for more detail), leaving 52 in the NNS (42 females, mean age = 25 years) and 52 in the NS group (31 females, mean age = 21.9 years) in the final analysis.

Participants in the NS group were recruited among the student population at the University of Hawai‘i. They reported that English was their first language, used dominantly since childhood. They received partial credit toward a course requirement for their participation. The participants in the NNS group were recruited from Ewha Womans University and Seogang University in Seoul, South Korea. Data from the language background questionnaire indicated that these participants had started learning English at the mean age of 8.8 years (SD = 2.8) and that 15 participants had experience of staying in English-dominant countries more than 6 months (mean = 26 months, Range: 6–60). Since I found no interaction between L2 learning experience and cross-linguistic activation in Experiment 3, I did not consider L2 participants' length of stay abroad in this experiment. Detailed participant information is presented in Table 6.1.

Table 6.1. Experiment 4: Participant information

| Group | Age | Onset of learning English | Self-ratings of overall English proficiency (1-10) | LexTALE score | Cloze test score (max = 50) |
|---|---|---|---|---|---|
| NS (*n* = 52) | 21.9 (2.8) | - | 9.6 (0.5) | - | - |
| NNS (*n* = 52) | 25.0 (4.0) | 8.9 (2.8) | 6.8 (1.4) | 70.6 (12.7) | 27.3 (9.5) |

*Note*. Numbers in parentheses indicate standard deviations

The L2 learner sample comprised participants with varying L2 proficiency, as indicated by their self-reported ratings of English proficiency (mean for listening = 7.5 out of 10, SD = 1.6; mean for speaking = 6.0, SD = 2.0; mean for reading = 7.4, SD = 1.4; mean for writing = 6.1, SD = 1.9), LexTALE scores (mean = 70.6, SD = 12.7), and cloze test scores (mean = 27.3 out of 50, SD = 9.5). As shown in Table 6.2, proficiency measures significantly correlated with each other, except for cloze test scores and self-ratings for listening.

Table 6.2. Correlation matrix for the proficiency measures in Experiment 4

| | speaking | reading | writing | LexTALE | Cloze test |
|---|---|---|---|---|---|
| listening | .620*** | .570** | .409** | .472** | .263 |
| speaking | | .471*** | . 772** | 448** | .471** |
| reading | | | .600** | .513** | .335* |
| writing | | | | .508** | .563** |
| LexTALE | | | | | .543** |

* $p < .05$, ** $p < .01$, *** $p < .001$

To estimate participants' proficiency in comparison to the proficiency levels of the L2 learners in Experiment 3, planned pairwise comparisons were performed for the proficiency scores between the L2 learners in this study and the NNS-KR and NNS-US groups from Experiment 3, respectively, using independent samples t-tests. The NNS group in this experiment was not significantly different from the NNS-KR group in terms of self-ratings averaged over the four domains ($t(81) = –0.816$, $p = .417$), LexTALE scores ($t(81) = –0.060$, $p = .953$), and cloze test scores ($t(81) = 1.089$, $p = .279$). This group also did not differ from the NNS-US group in the averaged self-ratings ($t(81) = –1.471$, $p = .145$) and cloze test scores ($t(81) = –1.582$, $p = .118$), but had significantly lower scores than the NNS-US group in LexTALE ($t(81) = –2.174$, $p = .033$). Overall, the L2 participants in Experiment 4 have English proficiency closely matched with that of the L2 learners in Experiment 3, allowing for comparing results of the L2 participants across Experiments 3 and 4. The L2 learners were paid the Korean equivalent of $20 for their participation.

*6.2.2 Materials*

*6.2.2.1 Linguistic stimuli for the eye-tracking experiment*

The same set of English verbs used in Experiment 3 appeared in 48 unique experimental items, consisting of 24 items with NP1-biased verbs (12 SC, 12 non-SC) and 24 items with NP2-biased verbs. Each item was comprised of three sentences, following Itzhak and Baum (2015): A context sentence, a critical sentence, and a follow-up question, as in (17). In order to prevent participants from strategically associating the target pronoun in the *because* clause with a bias-consistent referent, half of the critical sentences had a bias-consistent ending, and half a bias-inconsistent ending (e.g., Cozijn et al., 2011; Izhak & Baum, 2015).

(17) Examples of linguistic stimuli from the eye-tracking task (NP1-biased)

Bias-consistent ending

(Context) Justin and Steve met each other at a Halloween party last year..

(Critical) Justin frightened Steve at first sight because he was wearing a ghost costume.

(Question) Who was wearing a ghost costume?

Bias-inconsistent ending

(Context) Nicolas and Dean had a debate on politics.

(Critical) Nicolas convinced Dean during the debate because he was easily persuaded.

(Question) Who was easily persuaded?

In addition to the 48 items with remention bias verbs, 48 items with predicates with no known remention biases were included as fillers, using connectors other than *because* (e.g., *before*, *although*, *while*, *but* and *yet*) in the critical sentence (e.g., *Linda talked to Jennifer about the change of plans before the trial began*). Appendix D contains all items.

Linguistic stimuli were recorded in a sound booth by a female native speaker of English using broad-focus intonation. Each sentence was recorded two times in one sitting. To minimize variability in the average length of critical sentences, the auditory input needed to be presented in a consistent speed across conditions. To this aim, three highly proficient Korean-speaking learners of English, who were blind to the purpose of the experiment, rated the speed of the

critical sentences on a scale from 1 (very slow) to 5 (very fast). For each item, the recording with the average score among rators that was closer to 3 was selected for the experiment. In order to ensure that the recorded sound files were closely matched in terms of the durations across conditions, independent samples t-tests were used to compare durations and onset times for several regions of interest in the critical sentences between each condition (NP1, NP2; SC, non-SC). Results showed that the critical sentences in the NP1 condition were not significantly different from those in the NP2 condition in terms of total duration ($t(46) = 1.570$, $p = .123$), duration of the NP2 region in the main clause (from NP2 onset to the onset of *because*) ($t(46) = -1.387$, $p = .172$), onset of *because* ($t(46) = 0.865$, $p = .392$), and onset of the pronoun in the *because*-clause ($t(46) = 0.909$, $p = .368$). Likewise, the critical sentences in the SC condition did not significantly differ from those in the non-SC condition in total duration ($t(46) = 1.140$, $p = .266$), duration of the NP2 region in the main clause ($t(46) = -0.318$, $p = .753$), onset of *because* ($t(46) = -0.114$, $p = .910$), and onset of the pronoun in the *because*-clause ($t(46) = -0.212$, $p = .834$). These results indicate that the duration of these regions of interest is closely matched between the two bias type conditions (NP1 vs. NP2) as well as between the two predicate type conditions (SC vs non-SC).

Experimental and filler items were pseudorandomized so that no experimental items in the same bias (NP1, NP2) or predicate type (SC, non-SC) condition occurred more than once in a row. Two lists were created in reversed orders, and participants were randomly assigned to one of the two lists.

*6.2.2.2 Visual stimuli for the eye-tracking experiment*

For the visual stimuli, 96 scenes, each comprised of a set of two human faces, were created (Figure 6.1). Half of the scenes for the experimental sentences contained male faces and half female faces, in accordance with the linguistic stimuli. Each visual scene contained two areas of interest (AOIs), one corresponding to the main-clause subject (NP1) and the other to the main-clause object (NP2) in the linguistic stimuli. Names were printed below each face, as illustrated in Figure 6.1. The position of bias-consistent and bias-inconsistent referents was counterbalanced between items.


Figure 6.1. Sample of visual stimuli

*6.2.2.3 Stimuli for the translation task*

A translation task was conducted to serve the same purposes as in Experiment 3: assessing L2 learners' understanding of the target verbs and examining how often learners translate the target verbs into Korean SC or non-SC predicates. The items for the translation task were the main clause portion (without the adverbial phrase) of the 48 experimental items (24 NP1 and 24 NP2 items) from the eye-tracking task, presented as in (18).

(18) Justin bored Steve.

*6.2.3 Procedure*

Participants in the NS group only completed the eye-tracking task. NNSs completed tasks in the following order: eye-tracking task, lexical decision task (LexTALE), cloze test, and translation. Except for the eye-tracking task, all tasks were conducted via a web-based interface. There was a 5-10 minute break between tasks. Prior to these tasks, all participants completed a language background questionnaire.

Participants completed the visual-world eye-tracking task during a single visit to the lab at the University of Hawai'i (for NS) or in a quiet room at Ewha Womans University in Korea (for NNS). During the task, participants were seated at a comfortable distance from a laptop that was equipped with a remote eye-tracker below the screen. Prior to the experiment, they received written and oral instructions to listen to the sentences and answer a question. Eye movements

were recorded from the right eye with an SMI REDn Scientific eye-tracker with a sampling rate of 60 Hz. The experiment began with calibration and three practice trials. On each trial, participants saw visual scenes for 1000ms before the sound began to play, following Itzhak and Baum (2015). This was done to ensure that participants had time to examine the characters and their names in the scene. The images remained on the screen for the duration of the context, critical and question sentences. There was a 1000ms pause between the context and critical sentences and between the critical and question sentences. The question, which served primarily to keep participants engaged in the task, queried about the referent of the ambiguous pronoun in the critical sentence (experimental items) or one of the referents in the main clause (fillers). Participants responded to the question by clicking on one of the two images. No feedback was provided. After answering the question, participants proceeded to the next trial by pressing the spacebar. After finishing half of the trials, participants took a 5-10 minute break. The next block started with recalibration.

Following the eye-tracking task, NNS additionally completed the LexTALE task, cloze test and translation task. The entire experiment took approximately 40 minutes for NS and 80-90 minutes for NNS.

### *6.2.4 Data treatment and analysis*

#### *6.2.4.1 Eye-tracking experiment*

Using the eye-tracking software's default settings, eye gaze data categorized as fixations, saccades and blinks were obtained for the critical sentences (excluding context and question sentences). Data points were aggregated into 20ms time bins. Trials with insufficient fixation data were identified by calculating the proportion of sample points over the entire trial containing fixations. Trials with fixation data accounting for a number of sample points less than 2 standard deviations below the mean across all trials were excluded (1.5% in NS; 2.0% in NNS). Averaging over items, there was no item with a proportion of fixations below 2 standard deviations from the mean of items; thus no items were excluded. Six participants with an overall proportion of fixations below 2 standard deviations of the mean of all participants were identified and excluded from further analysis (3 NS; 3 NNS). Two additional participants were excluded,

either due to calibration failure (1 NS) or because eye-tracking data were not recorded (1 NNS). This process left 52 participants each in the NS and NNS groups.

Time windows for analysis were determined following Itzhak and Baum (2015), but with some modifications. First, Itzhak and Baum (2015) included the NP2 region in the main clause as an analysis window, but they did not find an effect of remention bias in this early region. Thus, in order to minimize Type I error rates by conducting multiple analyses, I excluded the NP2 region and aligned trials by the onset of *because*. Second, in Itzhak and Baum, the windows of analysis included a series of 200-ms windows following *because* (0-200ms and 200-400ms after the onset of *because*) and the pronoun (0-200, 200-400, 400-600, 600-800ms after pronoun onset). An increased number of time windows with each segment containing fewer sample points may increase chances for Type I error rates and reduce statistical power. Therefore, I decided to include fewer and larger analysis windows in this study. In addition, given that L2 processing is often slower than L1 processing, I decided to analyse eye gaze patterns after 800ms after pronoun onset as well. Based on these considerations, three successive temporal windows were determined as regions for analysis: from onset of *because* to pronoun offset (Because+Pro; 520 ms for all trials), from pronoun offset to 500ms after pronoun offset (Pro+1), and from 500ms to 1000ms after pronoun offset (Pro+2). Considering that it generally takes about 200ms to plan and execute an eye movement (Matin, Shao & Boff, 1993), each analysis window was offset by 200ms. That is, the analysis for the Because+Pro segment included the time frame from 200ms after the onset of *because* to 200ms after pronoun offset. Likewise, the Pro+1 and the Pro+2 windows extended from 200ms to 700ms, and from 700ms to 1200ms after pronoun offset, respectively.

For each of the three analysis windows, participants' preference for fixating on NP2 versus NP1 images was calculated for each trial by subtracting the number of 20-ms bins with looks to NP1 from the number of bins with looks to NP2 (cf. Grüter, Takeda, Rohde & Schafer, 2018).[24] For each time window, two separate analyses were conducted, one testing for an effect of bias type (NP1-biased, NP2-biased), the other testing for an effect of predicate type (SC, non-SC)

[24] This dependent measure was chosen over a weighted empirical logit analysis (Barr, 2008), since the model residuals based on this measure were more normally distributed and thus produced a better model fit. Models based on the empirical logit transformation aggregated by participant and by item were also created and showed similar statistical results.

within the NP1-biased condition. The analysis that assesses listeners' use of remention bias information included *Group* (NS, NNS) and *Verb bias* (NP1-biased, NP2-biased) as fixed effects, and participant and item as random effects.[25] The analysis testing effects of cross-language activation on L2 referential processing included only the data from items with NP1-biased verbs, with *Group* (NS, NNS) and *Predicate type* (SC, non-SC) as fixed effects, and participant and item as random effects. All models were constructed with the maximal random effects structure allowed by the design, including a by-participant slope for *Verb bias* or *Predicate type*, and a by-item slope for *Group*. In case of convergence problems, the random effects structure was simplified by removing the by-item slope for *Group*. For each analysis, participants' proficiency scores (LexTALE scores, cloze test scores, self-ratings, combined scores) were added to the original model of the L2 data in a separate step, either as a continuous or categorical variable, in order to explore the potentially modulating role of L2 proficiency.

#### *6.2.4.2 Translation task*

The coding procedure for the translation task was identical to that in Experiment 3. Two coders, myself and a highly proficient English-speaking learner of Korean, who was blind to the test purpose, annotated participants' translations for accuracy. Trials from the eye gaze data were removed if the participant's response to that item on the translation task was incorrect (9%) or gave rise to intercoder disagreement (1%). For the annotation of *-key ha*, I coded participants' translations based on whether or not their responses contained this marking regardless of the original membership of each item in terms of predicate type.

## 6.3 Results

### 6.3.1 *Inspection of mouse-click responses*

In standard visual-world studies, it is customary to remove from analysis trials in which participants did not click on the named object (e.g., Chambers & Cooke, 2009). In the experimental design here, mouse-click responses do not reflect identification of a named target. Instead, they indicate participants' responses to the question about the referent of the ambiguous

[25] Following the previous visual world studies on remention bias (Contemori & Dussias, 2018; Cozijn et al., 2011; Itzhak & Baum, 2015), the alpha level for analyses in each window was set at .05.

pronoun (in the experimental items). Recall that in half the items, continuations were intended to disambiguate to the bias-consistent referent, in the other half to the bias-inconsistent referent. Thus mouse-click responses in this study indicate to what extent participants interpreted the disambiguating continuation as it was intended. This presents, potentially, a measure of how much attention they paid to the task, how well they understood the sentences, and how successfully the items were constructed to disambiguate reference. Importantly, mouse-click responses are uninformative with regard to participants' use of remention bias, the potential effects of which are expected to emerge only *prior t*o disambiguating information. For this reason, no trials were excluded based on mouse-click responses.

Inspection of mouse-click responses showed that participants chose the referent consistent with the intended disambiguation of the sentence 67.2% of the time (69.9% for NS; 64.5% for NNS), with somewhat higher selection rates of the intended referent (collapsing over bias-consistent and bias-inconsistent continuations) for NP2 (75.2% for NS; 71.9% for NNS) than for NP1 items (64.7% for NS; 57.2% for NNS). The selection rates of the intended referent for bias-consistent continuations were 68.8% (70.8% for NS; 66.8% for NNS) and those for bias-inconsistent continuations were 67.1% (69.8% for NS; 64.5% for NNS). These rates are lower than those observed in Itzhak and Baum (2015), who reported 93% of rates of selecting the intended referent for NP1 items and 87% for NP2 items among native English speakers. Importantly, however, the rates of selection of the intended referent were very similar in the NS and NNS groups. This suggests that the lower percentages compared to Itzhak and Baum are unlikely to be due to the materials being too difficult for the L2 learners. Instead, the most likely reason is that the disambiguating portions of the sentences were more ambiguous than they were intended. This should not impact participants' processing of remention bias in any ways relevant to the questions addressed in this study.

### *6.3.2 Analyses of eye gaze*

I first report results from the analyses that probe whether L2 learners draw upon remention bias information during online processing as efficiently as native speakers. Then, I turn to the results of the analyses testing the influence of cross-linguistic activation in the L2 use of remention bias information. For each analysis, I tested for effects of proficiency by including

each measure (self-ratings, LexTALE, cloze test scores, and combined scores) as an additional fixed factor in the modelling.

*6.3.2.1 Use of remention bias information*

Figure 6.2 illustrates the time-course of fixations over the entire trial for the critical sentences. Visual inspection of the graphs shows that the NS group was more likely to fixate on NP1 referents in the NP1 condition (red solid line) than the NP2 condition (green solid line), and more likely to fixate on NP2 referents in the NP2 condition (green dotted line) than the NP1 condition (red dotted line). Differences between the two condition appear to emerge even before the onset of *because*. In contrast, differences between conditions in the NNS group do not emerge until the onset of Pro+1. In the Pro+1 region, a pattern similar to that in the NS group is beginning to emerge, and becomes stronger in the Pro+2 region. This pattern remains stable throughout the remainder of the sentence.

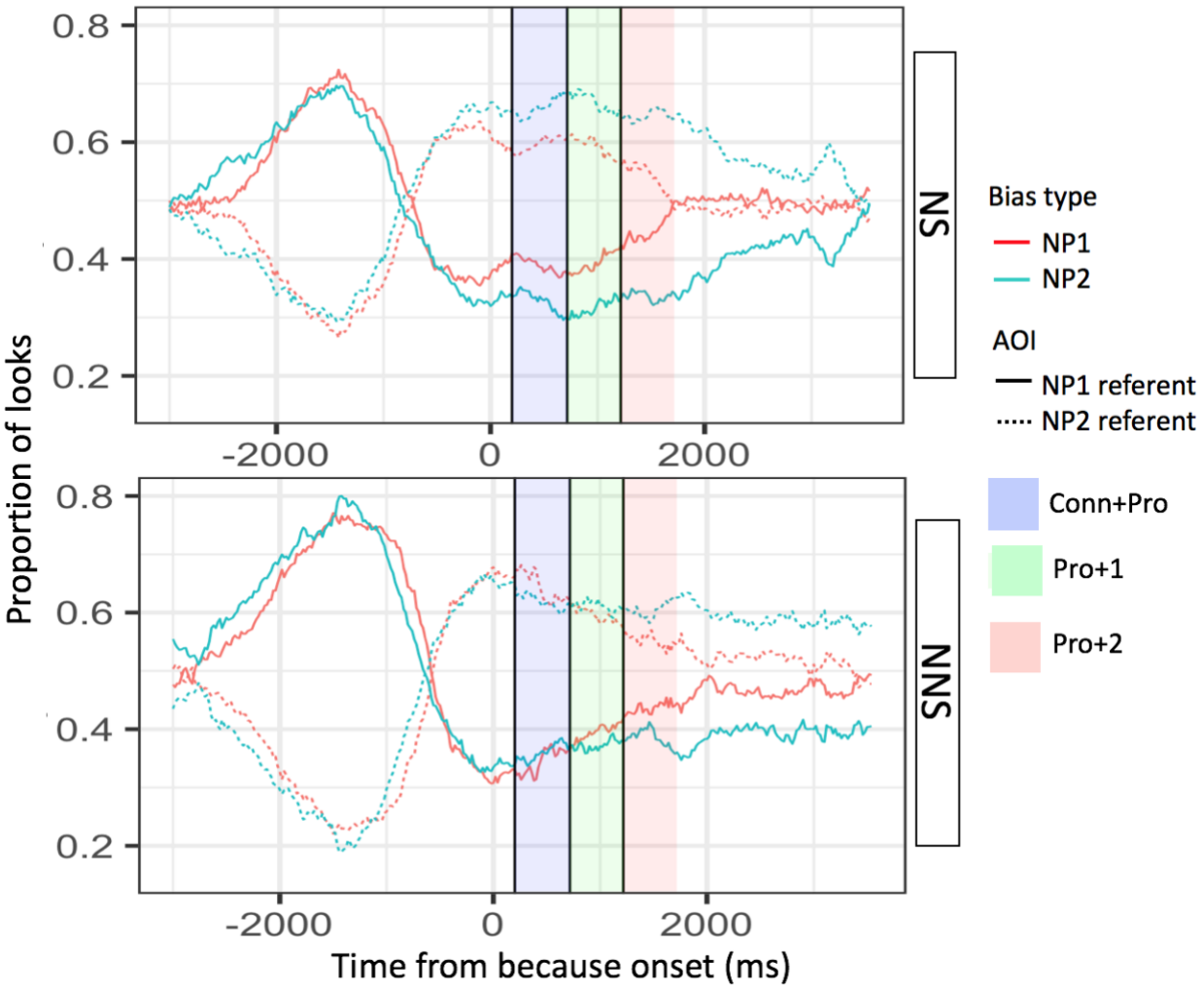


Figure 6.2. Overview of the time course of fixations across the trial for the critical sentences by verb bias (NP1, NP2) and AOI (NP1, NP2 referent) starting from *because* onset (0ms) for NS (upper panel) and NNS (lower panel) groups in Experiment 4. Proportions of looks to each AOI are calculated out of all fixations to NP1 or NP2. Means represent means over trials.

In order to assess when effects of verb bias emerge in each group, linear mixed-effects regression (lmer) was conducted on participants' preference for fixating on NP2 versus NP1 images for each window. Table 6.3 presents a summary of the output from these models. In the first window (Because+Pro), there were no main effects of *Group* ($b$ = –0.605, $SE$ = 1.334, $p$ = .652) or *Verb bias* ($b$ = 1.108, $SE$ = 0.863, $p$ = .206), but an interaction emerged between *Group* and *Verb bias* ($b$ = –3.394, $SE$ = 1.397 $p$ = .020). Separate analyses examining this interaction were performed for each group, with the model including *Verb bias* as a fixed effect.

The results showed a main effect of *Verb bias* in the NS ($b$ = 2.897, $SE$ = 1.002, $p$ = .008), but not in the NNS group ($b$ = –0.446, $SE$ = 1.192, $p$ = .710), indicating that only the NS group showed sensitivity to the remention bias information in this window.

In the Pro+1 region, there was an effect of *Verb bias* ($b$ = 1.865, $SE$ = 0.903, $p$ = .045), with a greater preference for fixating on NP2 versus NP1 in the NP2 than in the NP1 condition. There was no effect of *Group* ($b$ = –1.686, $SE$ = 1.225, $p$ = .173), but a trend towards an interaction ($b$ = –2.192, $SE$ = 1.288, $p$ = .095). Follow-up analyses conducted within each group revealed an effect of *Verb type* in the NS ($b$ = 2.989, $SE$ = 1.085, $p$ = .010), but not in the NNS group ($b$ = 0.835, $SE$ = 1.164, $p$ = .477). Consistent with the results from the previous window, the NSs' looking patterns, but not the NNSs', were influenced by remention bias in this region.

In the third window (Pro+2), a main effect of *Verb bias* ($b$ = 3.171, $SE$ = 0.975, $p$ = .002) emerged without a significant effect of *Group* ($b$ = –1.104, $SE$ = 1.230, $p$ = .372) or an interaction ($b$ = –2.025, $SE$ = 1.434, $p$ = .166), suggesting that *Verb bias* influenced both groups. Follow-up analyses conducted within each group demonstrated a robust effect of *Verb bias* for the NS ($b$ = 4.241, $SE$ = 1.138, $p$ < .001) and a weaker effect in the same direction for the NNS group ($b$ = 2.216, $SE$ = 1.271, $p$ = .089). These results suggest an emerging role of remention bias in the NNS group in this region.

Table 6.3. Results of the mixed-effects logistic regression in Experiment 4: Use of remention bias information

| | *b* | *SE* | *p* |
|---|---|---|---|
| Because+Pro | | | |
| (Intercept) | 5.113 | 0.711 | < .001 |
| Group | –0.605 | 1.334 | .652 |
| Verb bias | 1.108 | 0.863 | .206 |
| **Group × Verb bias** | **–3.394** | **1.397** | **.020** |
| Pro+1 (0ms-500ms after pronoun offset) | | | |
| (Intercept) | 4.407 | 0.687 | < .001 |
| Group | –1.686 | 1.225 | .173 |
| **Verb bias** | **1.865** | **0.903** | **.045** |
| Group × Verb bias | –2.192 | 1.288 | .095 |
| Pro+2 (500ms-1000ms after pronoun offset) | | | |
| (Intercept) | 3.411 | 0.696 | < .001 |
| Group | –1.104 | 1.230 | .372 |
| **Verb bias** | **3.171** | **0.975** | **.002** |
| Group × Verb bias | –2.025 | 1.434 | .166 |

*Note*. Formula for each model: lmer(type ~ verb.bias*group + (1 + verb.bias|participant) + (1 + group|item)

In order to further examine this late emerging effect in the NNS group, I decided to conduct exploratory analyses in the following 500ms region: from 1000ms to 1500ms after pronoun offset. A main effect of *Verb bias*, not modulated by *Group*, emerged in this window ($b$ = 4.330, $SE$ = 1.233, $p$ < .001). Separate analyses for each group showed a continually robust effect of *Verb bias* in the NS group ($b$ = 5.430, $SE$ = 1.684, $p$ = .002), and a weaker but increased (compared to the previous window) effect in the NNS group ($b$ = 3.331, $SE$ = 1.273, $p$ = .012). Given that this window roughly corresponds to the region including one or two words after the verb in the *because* clause of the critical sentence, the increased effect of verb bias in the NNS

group indicates that the learners were able to use remention bias information around the time when they processed the verb in the *because* clause.

Finally, to further explore whether the native and nonnative speakers' use of remention bias increases over the course of the experiment, potentially reflecting learning or adaptation effects, I conducted additional analyses for each time window including Verb bias (NP1-biased, NP2-biased) and *Trial Number* as fixed effects for each group. Results from the NS group showed that the main effect of Verb bias emerging in the three windows did not interact with Trial Number. Likewise, results from the NNS group showed that Trial Number did not interact with the weak effect of Verb bias in Pro+2 or the main effect of Verb bias in the exploratory window. These results indicate that the effect of remention bias is consistent over the course of the experiment, and thus unlikely to reflect learning or adaptation within the experiment itself.

In sum, the NS group showed sustained evidence of using remention bias information, starting from (at least) the onset of *because* and stretching up to 1500ms after pronoun offset, whereas similar effects did not start to emerge until substantially later in the NNS group, with a weak effect of verb bias emerging in the Pro+2 region, and becoming somewhat stronger in the 1000-1500ms window. These findings provide some indication that remention bias information affects L2 learners' referential processing during online comprehension, yet the timing of using this information was delayed compared to native speakers.

To probe for effects of proficiency, each measure of LexTALE scores, cloze-test scores, self-ratings (averaged across the scores in the four domains), and a combined score averaged over the z-scores of the three measures was added to the model of the NNS data. In separate exploratory models, the proficiency measure was added either as a continuous or a categorical variable. On analogy to Experiment 3, a proficiency measure added as a continuous variable was transformed to z-scores. When it was added as a categorical variable, L2 participants were divided into two proficiency groups based on a median score: NNS-H group with scores the same as or higher than the median and NNS-L group with scores below the median. Table 6.4 shows results of grouping by the median score of each proficiency measure.

Table 6.4. Results of median-split for each proficiency measure in Experiment 4

| | NNS-H | | NNS-L | |
|---|---|---|---|---|
| | Number of participants | Mean score (SD) | Number of participants | Mean score (SD) |
| LexTALE (median = 68.75) | 28 | 80.4 (8.3) | 24 | 59.2 (5.1) |
| Cloze test (median = 26) | 28 | 34.3 (5.6) | 24 | 19.1 (5.7) |
| Self-ratings (median = 6.8) | 26 | 7.9 (0.7) | 26 | 5.6 (0.9) |
| Combined z-score (median = –0.11) | 26 | 0.7 (0.5) | 26 | –0.7 (0.4) |

Figure 6.3 illustrates the time-course of fixations during the critical sentences, collapsing over participants and items, as a function of verb bias type for each proficiency group (NNS-H, NNS-L), created based on the mean split of the combined z-scores. Visual comparison of looking patterns in the NNS-H versus NNS-L groups suggest that the late effect of verb bias is greater for higher-proficiency learners than lower-proficiency learners.

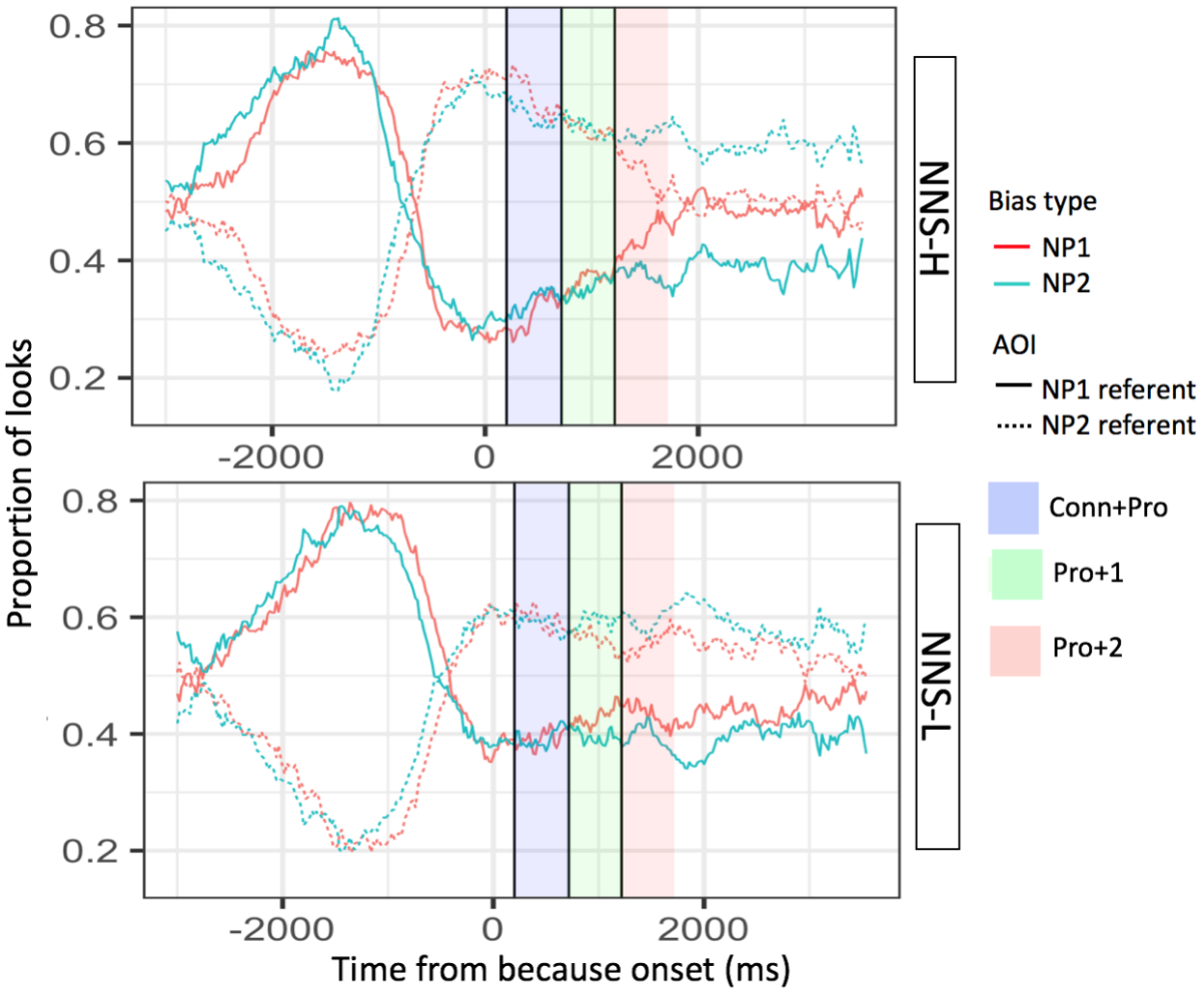


Figure 6.3. Overview of the time course of fixations across the trial for the critical sentences by verb bias (NP1, NP2) and AOI (NP1, NP2 referent) starting from *because* onset (0ms) for NNS-H (upper panel) and NNS-L (lower panel) groups in Experiment 4.

Modulating effects of proficiency on L2 use of remention bias information were modelled in each of the three original time windows of Because+Pro, Pro+1 and Pro+2, as well as in the additional exploratory window between 1000ms-1500ms after pronoun offset where the effect of verb bias appears strongest in the NNS group. Each model included *Verb bias* (NP1, NP2) and proficiency measure (either continuous or categorical) as fixed effects along with the maximal random effects structure allowed by the design. Results of the analyses showed that none of the proficiency measures significantly interacted with Verb bias in any of the time windows,

regardless of whether the measure was added to the model as a continuous or categorical variable. In model comparison using the anova() function in R, adding an interaction between *Predicate type* and *Proficiency* to the model containing the two fixed effects did not improve overall model fit, no matter what proficiency measure was used, indicating no interacting role of proficiency in the effect of predicate type.

In order to further probe what in Figure 6.3 appears to be a trend towards a stronger effect of Verb bias in the higher proficiency group later in the sentence, I inspected the effect of Verb bias in NNS-H and NNS-L groups divided by each of the proficiency measures summarized in Table 6.4 for each time window. The results of the separate analyses for each proficiency group showed no effect of *Verb bias* for any of the proficiency groups in Because+Pro, Pro+1, and Pro+2. In the window between 1000ms and 1500ms after pronoun offset, when the L2 learners were divided by their self-ratings and combined scores, the effect of *Verb bias* emerged somewhat more strongly for the NNS-H group (self-ratings: $b = 3.578$, $SE = 1.715$, $p = .043$; combined scores: $b = 3.918$, $SE = 1.632$, $p = .021$), than for the NNS-L group (self-ratings: $b = 2.868$, $SE = 1.549$, $p = .070$; combined scores: $b = 2.418$, $SE = 1.526$, $p = .119$). Yet overall, these findings indicate that the effect of predicate type was not robustly modulated by proficiency.

#### *6.3.2.2 Effects of cross-linguistic activation*

To investigate the role of cross-linguistic activation, participants' preference for fixating on NP2 vs. NP1 was analysed in the three analysis windows (offset by 200ms, Because+Pro, Pro+1, and Pro+2), focusing only on the items with NP1-biased verbs. On analogy to Experiments 2 and 3, the data were analysed in three different ways: (1) analysis of total data, (2) analysis of translation-consistent data, and (3) analysis of data by participant-driven category. Each analysis included *Predicate type* (SC, non-SC) and *Group* (NS, NNS) as binary fixed effects. In addition, proficiency measures (LexTALE scores, cloze test scores, self-ratings, combined scores) were added to the model to inspect a modulating role of proficiency in cross-linguistic influence.

##### *6.3.2.2.1 Analysis 1: Total data*

Figure 6.4 illustrates looks to NP1 and NP2 images across predicate types (SC, non-SC) in the total data without taking into consideration participants' translations of individual items in

the translation task. Visual inspection of the graphs suggests that the preference for fixating on NP2 vs. NP1 images was greater for SC than non-SC conditions in the NS group in the Pro+1 and Pro+2 regions. The NNS group showed a pattern similar to that in the NS group but in almost all regions after the onset of *because*.

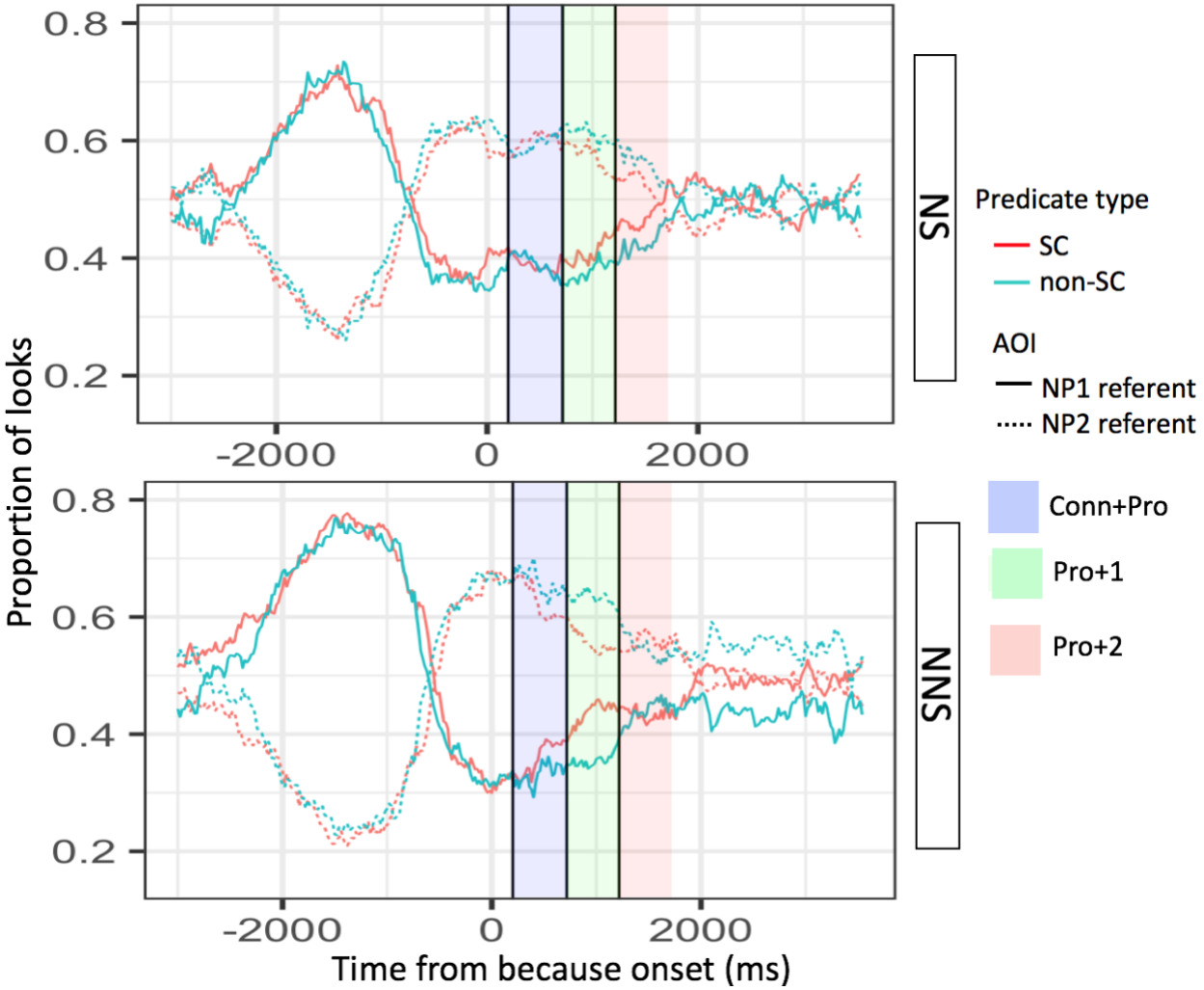


Figure 6.4. Overview of the time course of fixations across the trial for the critical sentences by predicate type (SC, non-SC) and AOI (NP1, NP2 referent) starting from *because* onset (0ms) for NS (upper panel) and NNS (lower panel) groups (Analysis 1: Total data) in Experiment 4.

A mixed-effects logistic regression (lmer) was conducted separately for each of the three time windows of interest, as well as for the exploratory 1000ms-1500ms window. As shown in Table 6.5, there were no clear effects of *Predicate type*, nor interactions with *Group*, in any of the regions. Separate analyses within each group and region also did not show any effect of *Predicate type*. Thus despite what in Figure 6.4 appeared as temporary trends in the expected direction, there is no statistical evidence for an effect of cross-linguistic activation in the analysis of the total data.

Table 6.5. Results of the mixed-effects logistic regression in Experiment 4: Effects of cross-linguistic activation (Analysis 1: Total data)

| | *b* | *SE* | *p* ($\alpha$ = .017) |
|---|---|---|---|
| Because+Pro | | | |
| (Intercept) | 4.568 | 0.879 | < .001 |
| Group | 1.052 | 1.595 | .512 |
| Predicate type | 0.403 | 1.286 | .757 |
| Group × Predicate type | 0.800 | 2.090 | .705 |
| Pro+1 (0ms-500ms after pronoun offset) | | | |
| (Intercept) | 3.527 | 0.866 | < .001 |
| Group | –0.550 | 1.483 | .713 |
| Predicate type | 2.146 | 1.392 | .138 |
| Group × Predicate type | 1.239 | 2.125 | .566 |
| Pro+2 (500ms-1000ms after pronoun offset) | | | |
| (Intercept) | 1.849 | 0.930 | .054 |
| Group | –0.038 | 1.327 | .977 |
| Predicate type | 0.657 | 1.568 | .679 |
| Group × Predicate type | –2.081 | 1.733 | .233 |
| Exploratory window (1000ms-1500ms after pronoun offset) | | | |
| (Intercept) | 0.332 | 0.938 | .726 |
| Group | 1.367 | 1.234 | .272 |
| Predicate type | 1.533 | 1.754 | .391 |
| Group × Predicate type | –0.703 | 2.071 | .736 |

*Note*. Formula for the models in Because+Pro, Pro+1, and the exploratory window: lmer(type ~ predicate.type*group + (1 + predicate.type|participant) + (1 + group|item), Formula for the model in Pro+2: lmer(type ~ predicate.type*group + (1 + predicate.type|participant) + (1|item)

When each measure of proficiency scores (LexTALE scores, cloze test scores, self-ratings, and combined scores) were added as a fixed factor to the models of the NNS data, there were no

significant interactions between any of the measures and Predicate type in any of the time windows. Adding an interaction of *Proficiency* and *Predicate type* to the model containing the two fixed effects did not improve overall model fit. Also, separate analyses by proficiency subgroup divided by median scores did not show any effects of *Predicate type* in the higher- or lower-proficiency groups. There is thus no evidence for a modulating role of proficiency in cross-linguistic activation in the analysis of the total data.

*6.3.2.2.2 Analysis 2: Translation-consistent data*

In this analysis, items with a translation that was inconsistent with the predetermined predicate type were removed. Table 6.6 presents the distribution of consistent and inconsistent translations for predicate type in the L2 group.

Table 6.6. Number (percentage) of translation-consistent and translation-inconsistent items in Experiment 4

| Original membership of predicate type | Consistent items | Inconsistent items | Total |
|---|---|---|---|
| Non-SC | 487 (80.2) | 120 (19.8) | 607 |
| SC | 420 (71.3) | 169 (28.7) | 589 |

Figure 6.5 illustrates participants' looking patterns after removing translation-inconsistent items. Visual inspection of the graphs shows that the NS group appears more likely to fixate on the bias-consistent referent in the SC condition than the non-SC condition in the Pro+1 and Pro+2 regions. Also, the NNS group had more fixations on the bias-consistent referent in the SC condition than the non-SC condition in the Because+Pro and Pro+1 regions and the regions following 2000ms after pronoun offset. However, the effect of predicate type appears greater for the NNS than the NS group, particularly in the Pro+1 region.

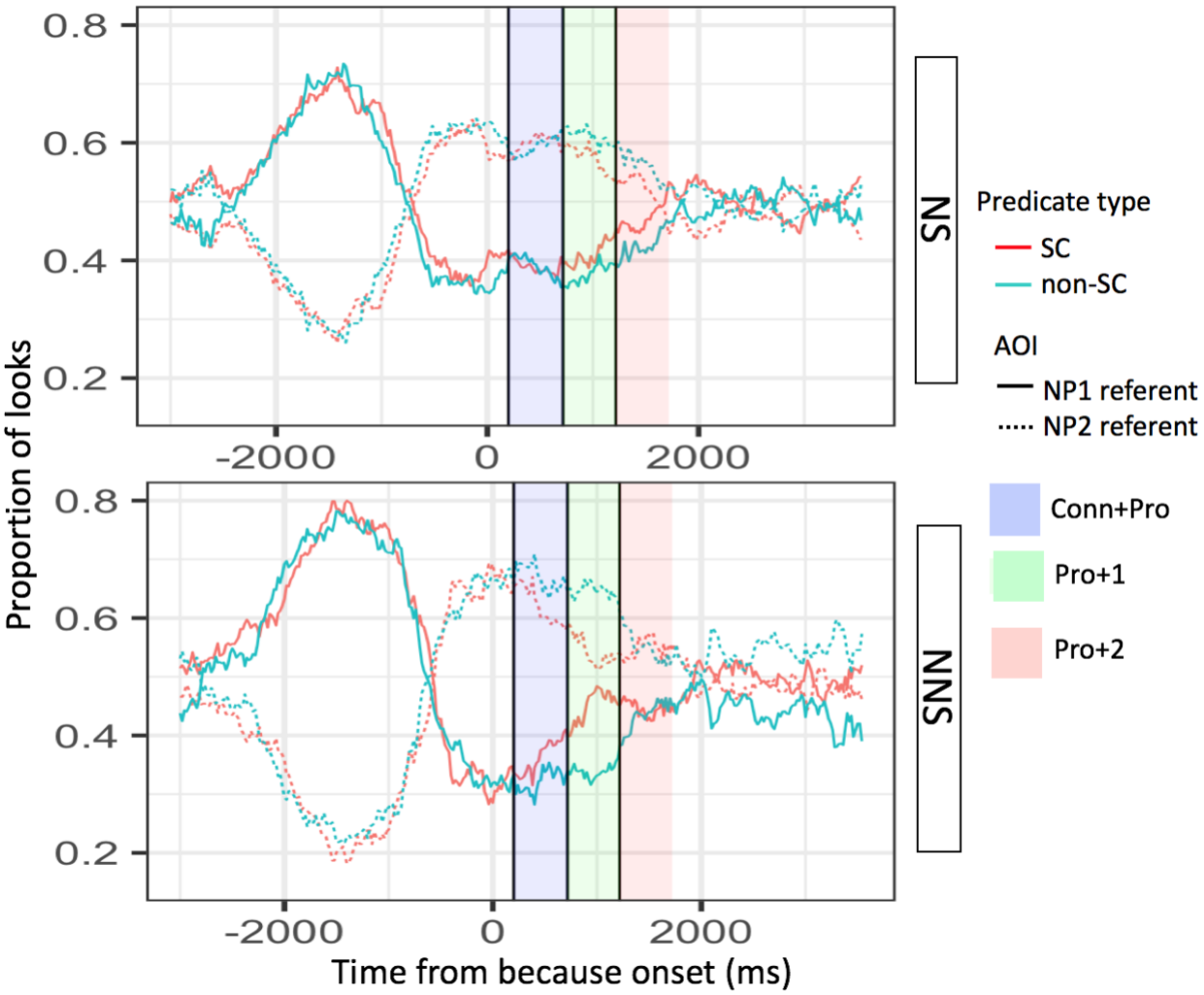


Figure 6.5. Overview of the time course of fixations across the trial for the critical sentences by predicate type (SC, non-SC) and AOI (NP1, NP2 referent) starting from *because* onset (0ms) for NS (upper panel) and NNS (lower panel) groups (Analysis 2: translation-consistent data) in Experiment 4.

Mixed-effects logistic regression conducted separately for each time window showed no main effect of *Predicate type* or its interaction with *Group* in any of the regions. Also, despite a trend toward a predicate type effect in the NNS group in the Pro+1 window, no main effect of *Predicate type* was found in any of the regions, again showing no evidence for the effect of cross-linguistic activation in the second analysis (see Table 6.7).

Additional analyses including proficiency measures demonstrated no significant interaction between proficiency and *Predicate type* in any of the windows. Adding an interaction of *Proficiency* and *Predicate type* to the model containing the two fixed effects did not improve overall model fit. Also, by-group analyses (median split) did not show a main effect of *Predicate type* both for the higher and lower groups, providing no evidence that proficiency affected cross-linguistic activation in this analysis.

Table 6.7. Results of the mixed-effects logistic regression in Experiment 4: Effects of cross-linguistic activation (Analysis 2: Translation-consistent data)

| | *b* | *SE* | *p* ($\alpha$ = .017) |
|---|---|---|---|
| Because+Pro | | | |
| (Intercept) | 4.507 | 0.883 | < .001 |
| Group | 1.073 | 1.611 | .508 |
| Predicate type | 0.702 | 1.233 | .575 |
| Group × Predicate type | 1.577 | 2.050 | .451 |
| Pro+1 (0ms-500ms after pronoun offset) | | | |
| (Intercept) | 3.687 | 0.905 | < .001 |
| Group | –0.330 | 1.504 | .827 |
| Predicate type | 2.483 | 1.493 | .111 |
| Group × Predicate type | 2.193 | 2.237 | .337 |
| Pro+2 (500ms-1000ms after pronoun offset) | | | |
| (Intercept) | 1.876 | 0.929 | .051 |
| Group | –0.043 | 1.384 | .975 |
| Predicate type | 1.004 | 1.584 | .533 |
| Group × Predicate type | –1.625 | 1.999 | .425 |
| Exploratory window (1000ms-1500ms after pronoun offset) | | | |
| (Intercept) | 0.062 | 0.991 | .951 |
| Group | 0.912 | 1.256 | .472 |
| Predicate type | 1.406 | 1.872 | .460 |
| Group × Predicate type | –1.118 | 2.164 | .610 |

*Note*. Formula for each model: lmer(type ~ predicate.type*group + (1 + predicate.type |participant) + (1 + group|item)

*6.3.2.2.3 Analysis 3: Participant-driven analysis*

In this analysis, the L2 data were re-categorized into SC and non-SC depending on whether or not participants' translations contained -*key ha*, regardless of the original membership of the

predicate type. Table 6.8 presents the distribution of the recategorized data.

Table 6.8. Distribution of data after the reorganization process in Experiment 4

| Original membership of predicate type | Number (%) of items categorized as Non-SC | Number (%) of items categorized as SC | Total |
|---|---|---|---|
| Non-SC | 487 (80.2) | 120 (19.8) | 607 |
| SC | 169 (28.7) | 420 (71.3) | 589 |

Figure 6.6 presents results from the L2 data with participant-driven categories. Visual inspection of the graphs suggest a pattern similar to that in Analysis 2: more fixations on the bias-consistent referent in the SC condition than the non-SC condition for the NS group in the Pro+1 and Pro+2 regions and a greater effect in the same direction for the NNS group in the Because+Pro and Pro+1 regions.

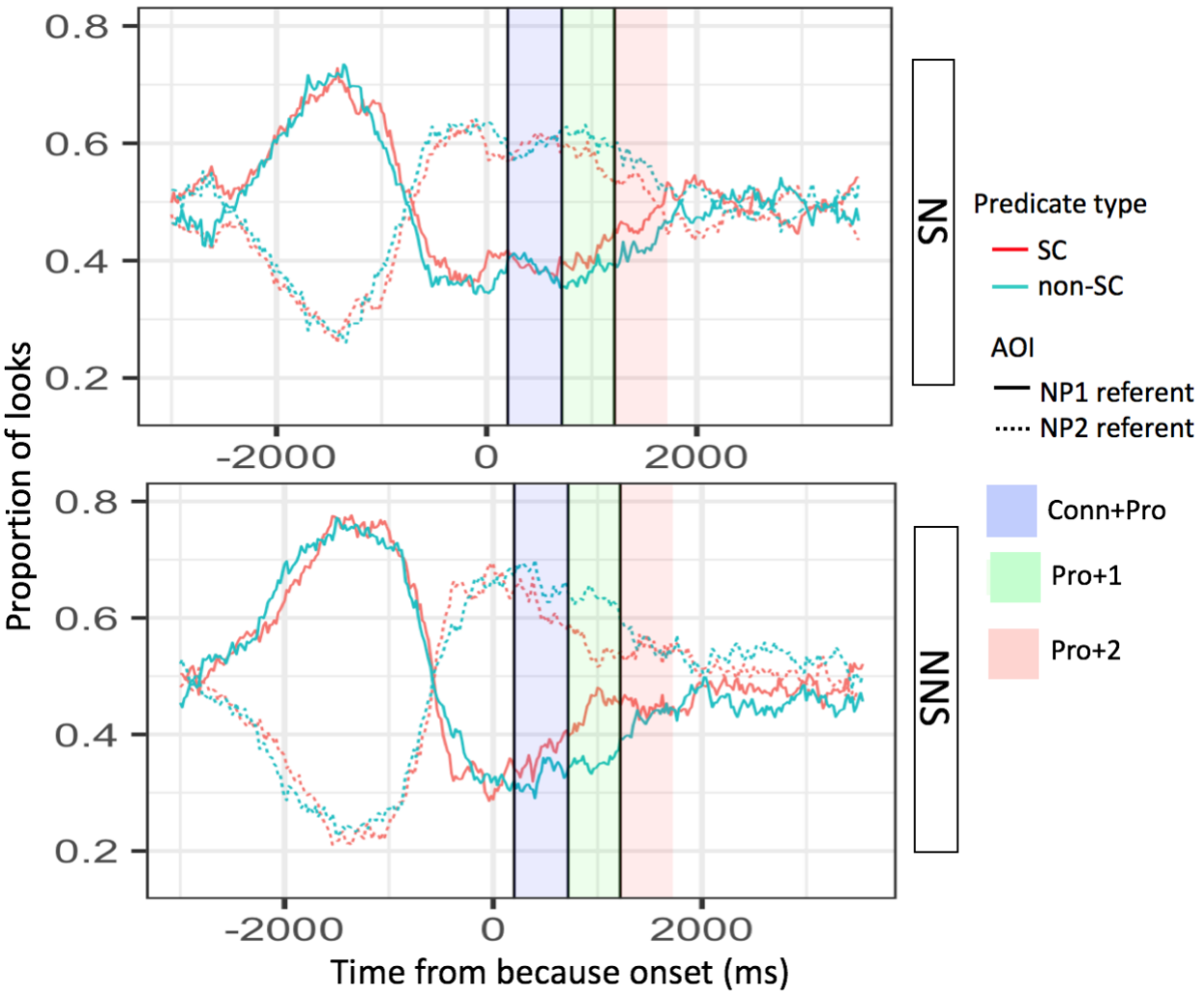


Figure 6.6. Overview of the time course of fixations across the trial for the critical sentences by predicate type (SC, non-SC) and AOI (NP1, NP2 referent) starting from *because* onset (0ms) for NS (upper panel) and NNS (lower panel) groups (Analysis 3: data with participant-driven categories) in Experiment 4.

Consistent with the results of the previous two analyses, there were no robust effects of *Predicate type* or *Group*, nor any interaction in any of the regions (see Table 6.9). Notably, an effect of *Predicate type* trending in the predicted direction emerged in the Pro+1 condition, albeit not modulated by an interaction with *Group*. Analyses within each group indicated that there was a weak effect of *Predicate type* for the NNS ($b$ = 2.828, $SE$ = 1.369, $p$ = .040), showing a trend toward more fixations on the bias-consistent referent in the SC condition than in the non-SC

condition, but no such effect was found for the NS group ($b$ = 1.525, $SE$ = 1.628, $p$ = .359). However, this trend in the NNS group was only short-lived: there was no effect of *Predicate type* for the NNS group in the following window of Pro+2 ($b$ = 1.078, $SE$ = 1.390, $p$ = .439) or in the window from 1000ms to 1500ms from pronoun offset ($b$ = –0.050, $SE$ = 1.354, $p$ = .971). In sum, although visual inspection of the graphs suggests an effect of predicate type trend in the predicted direction in the Pro+1 region, these results do not present evidence of any robust effects of cross-language activation in the L2 processing of remention bias.

Table 6.9. Results of the mixed-effects logistic regression in Experiment 4: Effects of cross-linguistic activation (Analysis 3: Data with participant-driven categories)

| | *b* | *SE* | *p* ($\alpha$ = .017) |
|---|---|---|---|
| Because+Pro | | | |
| (Intercept) | 4.560 | 0.876 | < .001 |
| Group | 1.038 | 1.580 | .514 |
| Predicate type | 0.745 | 1.093 | .499 |
| Group × Predicate type | 1.481 | 1.973 | .458 |
| Pro+1 (0ms-500ms after pronoun offset) | | | |
| (Intercept) | 3.503 | 0.862 | < .001 |
| Group | –0.643 | 1.481 | .666 |
| Predicate type | 2.089 | 1.106 | .063 |
| Group × Predicate type | 1.667 | 1.991 | .407 |
| Pro+2 (500ms-1000ms after pronoun offset) | | | |
| (Intercept) | 1.856 | 0.926 | .052 |
| Group | –0.124 | 1.415 | .930 |
| Predicate type | 1.532 | 1.171 | .195 |
| Group × Predicate type | –1.069 | 1.955 | .587 |
| Exploratory window (1000ms-1500ms after pronoun offset) | | | |
| (Intercept) | 0.360 | 0.936 | .704 |
| Group | 1.376 | 1.234 | .269 |
| Predicate type | 0.337 | 1.295 | .796 |
| Group × Predicate type | –0.941 | 2.001 | .640 |

*Note*. Formula for each model: lmer(type ~ predicate.type*group + (1 + predicate.type |participant) + (1 + group|item)

When proficiency measures were added as a fixed factor, there was no significant interaction between proficiency and *Predicate type* in any of the time windows, including the Pro+1 region where the L2 learners showed the small effect of *Predicate type*. Adding an interaction of

*Proficiency* and *Predicate type* to the model containing the two fixed effects did not improve overall model fit. Also, separate analyses for each proficiency group (median split) did not show a main effect of *Predicate type* both for the higher and lower groups. These results indicate that proficiency did not affect cross-linguistic activation in this analysis.

Overall, the current results showed no evidence of robust effects of cross-linguistic activation in L2 learners' online processing of remention bias information. Although visual inspection of fixation patterns suggests a trend in the expected direction in the Pro+1 region, little support for this pattern was found in the statistical models adopted here. No modulating effects of proficiency were observed.

### 6.4 Discussion

The primary goal of Experiment 4 was to investigate the extent to which cross-linguistic differences between Korean and English predicates in terms of remention bias strength affect L2 learners' referential processing. To this end, I used the visual world paradigm to test (1) whether L2 learners can use verb-induced remention bias information and (2) whether cross-linguistic activation influences their use of remention bias information in three successive temporal windows after the onset of *because* in the causal dependent clause.

With regard to the use of remention bias information, the native and nonnative groups showed different results in terms of the timing of using remention bias information. The analysis of the native speakers' preference for fixating on NP2 vs. NP1 conditioned on verb bias type (NP1, NP2) indicated an early and robust effect of verb bias, which persisted from the onset of *because* up to 1500ms after pronoun offset. The native speakers' sustained preference for fixating the bias-consistent referent from *because* onset suggests their early detection and recruitment of remention bias information, consistent with previous findings on L1 processing of remention bias (e.g., Cozijn et al., 2011; Itzhak & Baum, 2015). Like the current outcomes from the native speakers, Cozijn et al. (2011) found an early effect of remention bias among native speakers of Dutch, who showed a stronger preference for fixating on the bias-consistent referent than the bias-inconsistent referent in the time windows including the causal conjunction *omdat* ('because') and the subject pronoun. Similarly, native English speakers in Itzhak and Baum (2015) showed stronger fixation preferences for bias-consistent versus bias-inconsistent referents

from the onset of *because* up to 400ms after pronoun onset. Together with these previous findings, the early detection of remention bias information in the native speakers in this study suggests that remention bias information is readily available to native speakers well before they encounter disambiguating information, aligning well with the focusing account of referential processing (Greene & McKoon, 1995; Koornneef & Van Berkum, 2006; Long & De Ley, 2000), which predicts an early activation of remention bias information before disambiguating information.

Turning to the results of the L2 group, I found that L2 learners drew on remention bias information, but not as early as native speakers. There was an emerging effect of verb bias in the L2 group in the time window between 500ms and 1000ms after pronoun offset. As indicated by exploratory analyses, this effect became stronger in the subsequent time window (1000ms-1500ms after pronoun offset). The eye-movement patterns of the L2 learners contrasted with those of the native speakers, who showed an early effect of remention bias. Although the L2 group showed a delayed effect of remention bias, their processing patterns are not aligned with the integration account (Garnham et al., 1996; Stewart et al., 2000), which claims that the effect of remention bias manifests only after disambiguating information is encountered. Instead, the results from the L2 group are also compatible with the focusing account (Greene & McKoon, 1995; Koornneef & Van Berkum, 2006; Long & De Ley, 2000), which posits that remention bias can be activated prior to disambiguating information. Considering that the time windows from 500ms-1000ms and from 1000ms-1500ms after pronoun offset roughly correspond to one or two words after the verb in the *because*-clause, and these regions are far from providing disambiguating information, the emerging effect of remention bias in these regions for the L2 group suggests that the learners did use remention bias before they encountered disambiguating information. The focusing and the integration accounts were initially formulated for native speakers, and previous studies have tested these accounts exclusively based on native speaker processing. However, the effect of remention bias in the L2 learners, albeit delayed until a few words into the second clause, suggests that the focusing account may be extended to account for the processing patterns of L2 learners. Further studies testing the focusing and integration accounts with diverse L2 populations should advance our understanding of this issue.

Among theoretical accounts proposed for L2 learners, the delayed effect of remention bias in the L2 learners is compatible with the RAGE hypothesis (Grüter et al., 2017) in that the learners showed a reduced ability to use the verb's remention bias information to generate expectations about who will be mentioned in the causal dependant clause. Supporting evidence of L2 learners' reduced ability to use remention bias for predictive processing is provided by Contemori and Dussias (2018), who observed that highly proficient bilinguals successfully integrated remention bias information during discourse processing, but their use of the information was significantly delayed compared to native speakers. Unlike their study, the current experiment involved less proficient, sequential L2 learners, and more direct comparison between the L1 and L2 groups. The present study provides novel evidence that less proficient L2 learners can use remention bias information for referential processing during online comprehension, yet their abilities to predict an upcoming referent using this information is more limited compared to the native speakers, as predicted by the RAGE hypothesis.

The delayed use of remention bias information in L2 learners may stem from learners' difficulties with accessing and retrieving lexical representations for the verbs, and integrating the information to incrementally update their discourse models during real-time processing. Several studies have provided empirical evidence that L2 learners are more restricted compared to native speakers in terms of their ability to consistently access and retrieve lexical information due to lower-quality lexical representations (Dijkgraaf, et al., 2017; Kaan, 2014), and to integrate multiple sources of information (Hopp, 2010; Roberts et al., 2008; Sorace & Filiaci, 2006) during online processing. In the current eye-tracking experiment, successful recruitment of remention bias information for discourse processing requires participants to access and retrieve properties of remention bias verbs, integrate this information with the connector *because* to establish the discourse relations between the clauses, and create mental models of the events to create expectations about who or what will be mentioned in the following clause. While these processes appear to be accomplished relatively easily by native speakers, as shown in the early effect of remention bias observed in the NS group in this study, they may be more taxing for L2 learners, which may have led to the delayed effect of verb bias in the L2 group in this study.

Although using remention bias information in real-time comprehension presents challenges for L2 learners due to their more restricted abilities to access and retrieve lexical information in

real time and integrate the information for discourse processing, one may ask whether increased proficiency can improve learners' use of remention biases. It has been argued that L2 learners with high proficiency can have "more high-quality lexical representations" that are needed for predictive processing (Kaan, 2014, p. 268). Other studies show that L2 learners can successfully integrate information from multiple sources in discourse-level processing (e.g., Pan et al., 2015; Trenkic et al., 2014). In the current study, I found no statistical evidence of a role of proficiency in the L2 processing of remention bias.

A key motivation for this experiment, as addressed in RQ4, was not only to examine whether L2 learner can use remention bias information, but whether cross-linguistic activation influences the learners' processing of remention biases. Regarding this research question, despite some weak trend toward a stronger subject bias in the SC than non-SC condition in a specific time window for the L2 group (when the L2 data were analysed based on participant-driven categories), I found no robust evidence of effects of cross-linguistic activation, nor evidence of an interacting role of proficiency on cross-linguistic activation. These results indicate that cross-linguistic differences in bias strength do not appear to affect learners' referential biases during real-time processing, regardless of L2 proficiency.

The absence of effects of cross-linguistic activation may be related to the fact that the L2 group showed a delayed effect of remention bias. The delayed effect of remention bias in the L2 group may stem from L2 learners' difficulties with accessing and retrieving remention bias information and integrating it with the discourse information during online processing. It is possible that these difficulties may have dampened the chance that cross-linguistic activation influenced the learners' processing of remention bias. In other words, it may be possible that even if the learners activated both English verbs and their Korean counterparts in parallel, the cross-language activation may not have been able to make an impact on learners' reference expectations, since the bias information did not influence referential choices until later in the discourse, as indicated by the delayed effect observed in the first analysis. I will discuss this point in more detail in the next chapter by comparing the current results with those from the offline sentence-completion tasks in Experiments 2 and 3.

Overall, the results of Experiment 4 show that L2 listeners can use remention bias information during discourse processing, although the effect was delayed compared to what was

observed in native-speaker processing. No robust evidence was found of cross-language influence or of proficiency in the online L2 processing of remention biases. This is the first study that showed emerging, albeit delayed, effects of remention bias on L2 learners' eye gaze patterns in a homogeneous group of sequential bilinguals.

# CHAPTER VII
# GENERAL DISCUSSION AND CONCLUSION

The goal of this dissertation was to investigate if and how the strength of referential biases associated with remention bias predicates in Korean affects Korean-speaking learners' reference choices and referential processing in English. To this aim, I conducted offline written sentence-completion and online eye-tracking studies across four experiments. In what follows, I will report a summary of the outcomes from each experiment and discuss potential implications of these findings for the field of psycholinguistics and second language acquisition.

Experiment 1 tested Hartshorne et al.'s (2013) prediction that predicates containing explicit marking of causality will lead to stronger remention biases than predicates that do not encode such marking. Results from written sentence-completion tasks in Korean and English provided evidence in support of this prediction: Native Korean speakers showed stronger subject bias with syntactic-causative predicates, which contain the explicit causative marker *-key ha*, than with (subject-biased) non-syntactic-causative predicates in Korean; native English speakers showed similar biases for the English translation equivalents of these predicates. These results not only offered the first empirical support for Hartshorne et al.'s prediction, but also allowed for clear predictions regarding cross-language influence in Korean-speaking learners' referential processing in English. These predictions were tested in Experiment 2, where Korean-speaking learners of English completed a written sentence completion task in English. The analyses of participants' continuations that considered the individual translations provided by the L2 participants in the translation task indicated that the strength of a verb's referential bias in Korean affected Korean-speaking L2 learners' reference choices in English. This effect emerged regardless of the presence or absence of translation priming. When the L2 data were re-grouped into SC and non-SC depending on participants' translations (Analysis 3), a stronger subject bias was found in the SC condition than the non-SC condition both for learners who completed the translation task proceeding the sentence-completion task and for those who completed the translation task following the sentence-completion task. Experiment 3 replicated the findings from Experiment 2, using a more controlled set of remention bias verbs selected from consistent VerbNet classes (Kipper et al., 2008). Experiment 3 also tested to what extent L2 proficiency and

learning experience (immersed vs. instructed) influence effects of cross-linguistic activation in referential processing, as they have been shown to do in lexical processing (e.g., Basnight-Brown & Altarriba, 2007; Prior et al., 2017). The results showed no robust evidence that L2 English proficiency or learners' English learning experience modulates cross-linguistic activation in L2 referential choices. The effect of cross-linguistic activation in L2 learners' (offline) referential choices observed in Experiments 2 and 3 led to the question as to whether L2 learners' co-activation of English verbs and their Korean translation counterparts would also affect their use of remention biases during real-time language comprehension. To this end, Experiment 4 employed the visual-world paradigm to test L2 learners' use of remention bias information during online processing, and the potential influence of cross-linguistic activation in this process. L2 learners' eye-gaze data indicated that the learners made use of remention bias information in their incremental construction of meaning during real-time listening, but the timing of learners' use of remention bias was delayed compared to native speakers. No robust effects of proficiency or cross-linguistic activation were found.

The findings that the effect of cross-linguistic activation emerged for the L2 learners during their referential choices in sentence-completion tasks indicate that the learners had accessed English verbs and their Korean translation counterparts, which in turn affected their referential choices. These results are consistent with previous work that shows L2 learners' co-activation of words in both L1 and L2 (e.g., Altarriba, 1992; De Groot & Nas, 1991; Dijkstra & van Heuven, 2002; Dijkstra et al., 2018; Gollan et al., 997; Kroll, & Stewart, 1994; Marian & Spivey, 2003; Prior et al., 2017; Spivey & Marian, 1999). Unlike these studies, however, the present investigation goes further in demonstrating that the effects of cross-linguistic activation have an impact beyond the word and construction levels and can affect referential processing at the sentence- and discourse-level. These findings suggest that the integrated system of mental representations in L2 learners not only allows for lexical co-activation in parallel, but also influences learners' referential choices based on remention biases at the discourse level.

It is notable that the effect of cross-linguistic activation on L2 referential choices in offline sentence completion emerged irrespective of the presence of translation priming, L2 proficiency, or learners' L2 learning experience. Regarding the effect of translation priming, previous studies have shown that prior exposure to words in an L1 can facilitate the activation of their translation

counterparts in an L2 (e.g., Canseco-Gonzalez et al., 2010; De Groot & Nas, 1991; Elston-Güttler et al., 2005; Gollan et al., 1997; Kroll & Stewart, 1994). The analysis of participants' continuations that only included items with translations consistent with expected predicate type (Analysis 2) in Experiment 2 showed somewhat more robust effects of cross-linguistic activation on learners' referential choices when learners completed the translation task preceding the continuation task. However, the effect of cross-linguistic activation of remention bias strength occurred even without translation priming when the data were analysed taking participants' individual cross-linguistic associations into consideration (Analysis 3). On the assumption that the analysis respecting individual translations affords the most accurate picture of the individual differences among learners with regard to the specific cross-linguistic associations in their mental lexicons, these results indicate that cross-linguistic activation at word and construction levels can impact L2 referential choices at the discourse level both with and without translation priming.

I also found little evidence of a modulating role of L2 proficiency or learners' English learning experience in Experiment 3. Kaan (2014) argued that (predictive) processing in a non-native language may be modulated by factors such as L2 proficiency and the amount/type of exposure to the target language. Yet previous research has provided inconsistent findings regarding potential effects of proficiency or L2 learning experience in cross-language lexical co-activation. Some studies report reduced effects of cross-linguistic activation as L2 proficiency increases (e.g., Libben & Titone, 2009), while others demonstrate that lexical co-activation is observed only when learners' L2 proficiency is sufficiently high (e.g., Brenders et al., 2011). Still other studies show evidence of cross-language activation regardless of L2 proficiency (e.g., Chambers & Cooke, 2009; Duyck et al., 2004; Haigh & Jared, 2007; Jared & Kroll, 2001; Jared & Szucs, 2002; van Hell & Tanner, 2012; Zhou et al., 2010). In addition, several studies report consistent effects of cross-linguistic activation of words for learners who had only classroom instruction (e.g., Brenders et al., 2011; Jacob et al., 2017; Kantola & van Gompel, 2011; Lemhöfer et al., 2004; Weber & Cutler, 2004) as well as for those extensively immersed in an L2 (e.g., Costa et al., 2000; Haigh & Jared, 2007; Poarch & Van Hell, 2012; Wang & Foster, 2015; Zhao et al., 2011). This dissertation found consistent effects of cross-linguistic influence in offline sentence completion independent of L2 proficiency or language learning experience. It is possible that despite my effort to include L2 learners with varied proficiency levels, there may

not have been sufficient variance to detect an interacting role of proficiency with effects of cross-linguistic activation. It remains an open question whether recruiting L2 learners with more varied proficiency levels would allow us to see interactions between proficiency or language experience and cross-linguistic activation in L2 referential choices. I leave such further exploration for future research.

Unlike the robust effect of cross-linguistic activation in L2 learners' referential choices in Experiments 2 and 3, there was no clear evidence of cross-linguistic activation in the online task in Experiment 4. The contrasting results between the offline and online tasks are unlikely to come from differences in learner characteristics: L2 learners in Experiments 3 and 4 were closely matched in their English proficiency and experience. It also cannot be the case that differences in the experimental stimuli led to the divergent results across the experiments, since the same set of remention bias verbs was used in Experiments 3 and 4. Rather, the discrepancies may be related to the nature of the tasks employed in each study. Unlike in the written sentence-completion task in Experiment 3, in which participants had unlimited time to derive a sentence meaning and construe causal relations in the discourse when they completed the sentences, participants in the visual-world task in Experiment 4 were under significant time constraints. During the task, they needed to access and retrieve semantic properties of remention bias verbs, construe discourse coherence, and track reference across clauses during auditory language comprehension in real time. Thus, the divergent results obtained from the offline and online experiments may be related to the amount of processing burden imposed by the task, with increased processing load potentially reducing the effect of cross-linguistic activation. Previous studies have shown that L2 processing may differ from monolingual processing in global performance due to several processing-related problems (e.g., Cook, 1997; Cunnings, Fotiadou & Tsimpli, 2017; Hahne, 2001; Weber & Broersma, 2012). In particular, L2 learners may experience processing difficulties when they retrieve lexical information for predictive processing (e.g., Kaan, 2014; Hopp, 2013) or integrate multiple sources of information during online processing (Hopp, 2010; Roberts et al., 2008; Sorace & Filiaci, 2006). Consistent with this line of research, the results of the eye-tracking study in Experiment 4 showed that L2 learners were restricted in retrieving remention bias information and integrating this information with the discourse-level information for predictive processing, as indicated by the delayed effect of remention bias in L2 learners. In

addition, this study also provides suggestive evidence that the processing burden that presumably led to this delay may also dampen the chances that cross-linguistic activation will occur. Relevant to the effect of processing burden in cross-linguistic activation, some studies have shown that effects of cross-linguistic activation of words occur only when L2 learners had sufficiently high proficiency (e.g., Brenders et al., 2011), suggesting that freed-up resources as a result of high proficiency may increase chances for cross-linguistic activation. Since there are few studies that investigated the role of processing demands in cross-linguistic activation, it remains less clear whether or not it is the increased processing burden that led to the delayed use of remention bias information and the lack of effects of cross-linguistic activation for the L2 learners in this study. Thus, further research is required to systematically examine how the effect of cross-linguistic activation during L2 processing of remention bias will be influenced by learners' increased proficiency and/or the quality of lexical representations, as these qualities are argued to modulate L2 processing (Kaan, 2014).

An alternative account for the strong effect of cross-linguistic activation in the offline tasks but no clear evidence of such an effect in online processing in this study is that the L2 participants in the offline tasks might have had conscious access to both English verbs and their Korean translation counterparts by explicitly translating the sentences into Korean. Although no strong effect of translation priming was found in Experiment 2, one cannot rule out the possibility that even in the absence of an additional translation task, learners may have consciously translated the sentences before they provided continuations for the sentence fragment during the sentence-completion task. Such explicit translation could have increased the learners' awareness of and reliance on their L1 knowledge when they engaged in the completion task, which may magnify the effect of cross-linguistic influence. In the online task, on the other hand, the effect of cross-linguistic activation might have been attenuated as the learners had insufficient time for consciously translating the stimuli. To the best of my knowledge, there is no study that directly tested the role of conscious translation in the effect of cross-linguistic activation, and thus the findings of this dissertation provide a promising framework to investigate this issue in further research.

The overall findings of this dissertation contribute simultaneously to the psycholinguistic literature on remention bias, which has been dominated by studies with monolingual native

speakers, and to our understanding of cross-linguistic activation in bilinguals, where previous research has focused predominantly on lexical (word-level) processing. This dissertation thus adds to the growing body of psycholinguistic research on the role of remention bias in reference processing by investigating the non-native processing of remention bias by less proficient L2 learners (cf. Contemori & & Dussias, 2018; Liu & Nicol, 2010). The finding that the effect of remention bias emerged in the less proficient, sequential L2 learners suggests that the learners actively recruited remention bias information to make inferences about inter-clausal relations, consistent with previous findings showing that L2 learners can rely on discourse cues during sentence processing (e.g., Foucart et al., 2016). Moreover, this dissertation provides novel evidence that the focusing account of referential processing (Greene & McKoon, 1995; Koornneef & Van Berkum, 2006; Long & De Ley, 2000) applies to L2 processing of remention bias as well as to L1 processing by showing that the L2 learners used remention bias information to resolve reference in the causal dependent clause without the aid of disambiguating information. The finding that both L1 speakers and L2 learners demonstrated similar processing patterns supports the claim that L2 learners may not be fundamentally different from L1 speakers in their discourse-level processing. The only difference observed between the L1 and L2 groups was in the timing of the remention bias effect: Compared to the early effect in the L1 speakers, the effect emerged substantially later in the L2 learners, not modulated by L2 proficiency. These results suggest that L2 learners are guided by qualitatively the same predictive processing mechanisms that operate in L1 processing (Kaan, 2014), but their ability to generate expectations is more limited, as postulated by the RAGE hypothesis (Grüter et al., 2017), presumably due to L2 learners' slower access and retrieval of lexical information. These findings are expected to contribute to the SLA field by providing additional evidence of discourse-level expectations in L2 processing, which may help reach a more in-depth understanding of the extent to which L1 and L2 processing is similar or different.

At the same time, this dissertation pushes the realm of inquiry in research on cross-language influence in L2 learners to a level that goes beyond associations at lexical and constructional levels alone by investigating how cross-linguistic activation at word and construction levels may influence L2 learners' pragmatic inferences at a discourse level. While there has been prolific research on cross-language activation within diverse levels of representations, there has been

relatively little inquiry into transfer between two different levels. This dissertation provides the novel evidence that cross-language activation can occur between two differently sized units – a word (i.e., a lexical remention-bias verb in English) and a phrase-level construction (i.e., the Korean SC construction). Most notably, this dissertation showed that cross-linguistic activation extends beyond the word and construction levels and can influence L2 learners' reference choices in English sentences including remention bias verbs. This suggests that cross-language activation may exert a stronger and more prolonged influence on L2 processing than what has been reported by previous research. Although the effect of cross-linguistic activation emerged only in the offline tasks, but not in the online eye-tracking task, the results provide compelling evidence of effects of cross-linguistic activation on L2 discourse processing. Further research is required to examine whether such effects can emerge during L2 learners' real-time comprehension by recruiting learners with higher L2 proficiency and/or by designing experimental stimuli that impose fewer processing demands.

In addition, this dissertation contributes to broadening the empirical scope of research on remention biases cross-linguistically (a) by including Korean as a key language of interest and (b) by including L2 learners whose languages are typologically very distinct, which stands in contrast to the majority of previous research on cross-linguistic activation, which has largely focused on speakers of two Indo-European languages. Finally, this project contributes to our understanding of how L2 speakers derive meaning in context, and in particular to our understanding of the role of the native language in this process. This study suggests that L2 learners activate information from both their L1 and L2 in parallel, which impacts the referential processing of remention bias, yet the effect of cross-linguistic activation is more evident in offline tasks than online processing. To reach an in-depth understanding, future research should further investigate specific roles of potentially influencing factors that may influence the effect of cross-linguistic activation in L2 processing of remention biases, such as learners' cognitive capacity and/or conscious efforts to translate sentences.

APPENDIX A

Predicates used for Experiments 1a and 1b and their subject bias score in that experiment (=the percentage of subject reference out of all responses with either subject or object reference).

| Predicate type | Korean predicate in Experiment 1a | English translation from the NAVER dictionary | English predicate in Experiment 1b |
|---|---|---|---|
| Non-SC type | 사과하다 (sakwahata) – AP† – 91.7% | apologize to | apologize to – AP – 96.7% |
| | 접근하다 (cepkunhata) – AP – 62.1% | approach to | approach – AP – 70.0% |
| | 간청하다 (kanchenghata) – AP – 75.0% | plead | beg – AP – 87.5% |
| | 애원하다 (aywenhata) – AP – 51.5% | beg | beg – AP – 89.7% |
| | 전화하다 (cenhwahata) – AP – 82.9% | telephone | call – AP – 93.3% |
| | 부르다 (pwuluta) – AP – 71.0% | call | call – AP – 93.5% |
| | 사기치다 (sakichita) – AP – 94.1% | swindle | cheat – AP – 75.9% |
| | 자백하다 (capaykhata) – AP – 70.0% | confess to | confess to – AP – 89.3% |
| | 거역하다 (keyekhata) – AP – 36.7% | disobey | disobey – AP – 96.9% |
| | 방해하다 (panghayhata) –AP/SE – 85.7% | interfere | distract – AP/SE – 97.0% |

| | | | |
|---|---|---|---|
| | 이기다 (ikita) – AP – 72.2% | win | do better than – AP – 95.8% |
| | 아부하다 (apwuhata) – AP – 21.9% | flatter | flatter – AP/SE – 58.1% |
| | 굴욕주다 (kwulyokcwuta) – AP/SE – 84.8% | humiliate | humiliate – AP/SE – 65.5% |
| | 상처주다 (sangchecwuta) – AP/SE – 94.3% | hurt | hurt – AP/SE – 86.2% |
| | 훼방놓다 (hweypangnohta) – AP – 62.5% | interrupt | interrupt – AP – 56.0% |
| | 초대하다 (chotayhata) – AP – 81.0% | invite | invite – AP – 65.2% |
| | 거짓말하다 (kecismalhata) – AP – 96.8% | lie to | lie to – AP – 97.0% |
| | 지다 (cita) – AP – 68.8 | lose to | lose to – AP – 61.8% |
| | 협박하다 (hyeppakhata) – AP – 37.1% | threaten | threaten – AP/SE – 40.0% |
| | 위협하다 (wihyephata) – AP/SE – 70.6% | intimidate | threaten – AP/SE – 62.1% |
| SC type | 신경질나게하다 (sinkyengcilnakeyhata) – AP/SE – 100% | aggravate | aggravate – AP/SE – 90.6% |
| | 즐겁게하다 (culkepkeyhata) – SE – 63.3% | amuse | amuse – SE – 69.6% |
| | 분노하게하다 (pwunnohakeyhata) – | infuriate | anger – AP/SE – 81.8% |

| | | |
|---|---|---|
| AP/SE – 100% | | |
| 짜증나게하다 (ccacungnakeyhata) – AP/SE – 96.7% | annoy | annoy – AP/SE – 84.4% |
| 지루하게하다 (cilwuhakeyhata) – AP/SE – 100% | bore | bore – SE – 88.5% |
| 실망하게하다 (silmanghakeyhata) – SE – 100% | disappoint | disappoint – SE – 80.0% |
| 낙담하게하다 (naktamhakeyhata) – SE – 96.7% | discourage | discourage – AP/SE – 40.7% |
| 당황하게하다 (tanghwanghakeyhata) – AP/SE – 100% | bewilder | embarrass – AP/SE – 72.7% |
| 곤란하게하다 (konlanhakeyhata) – AP/SE – 93.3% | trouble | embarrass – AP/SE – 80.6% |
| 매혹시키다 (mayhoksikhita)* – AP/SE – 100% | fascinate | enchant – SE – 87.5% |
| 무섭게하다 (mwusepkeyhata) – AP/SE – 70.6% | scare | frighten – AP/SE – 90.3% |
| 화나게하다 (hwanakeyhata) – AP/SE – 97.1% | anger | make angry – AP/SE – 87.1% |
| 불안하게하다 (pwulanhakeyhata) – SE – 100% | disturb | make uneasy – SE – 83.3% |
| 불쾌하게하다 (pwulkhwayhakeyhata) – AP/SE – | offend | offend – AP/SE – 93.5% |

| | | | |
|---|---|---|---|
| | 90.9% | | |
| | 기쁘게하다 (kippukeyhata) – SE – 90.3% | please | please – SE – 90.3% |
| | 기분나쁘게하다 (kipwunnappukeyhata) – AP/SE – 96.9% | irritate | put in a bad mood – AP/SE – 100% |
| | 겁먹게하다 (kepmekkeyhata) – AP/SE – 82.4% | frighten | scare – AP/SE – 71.4% |
| | 놀라게하다 (nollakeyhata) – AP/SE – 93.3% | surprise | surprise – AP/SE – 80.0% |
| | 근심하게하다 (kunsimhakeyhata) – SE – 95.2% | concern | worry – SE – 87.9% |
| | 걱정하게하다 (kekcenghakeyhata) – SE – 96.4% | worry | worry – SE – 97.0% |
| Object-biased predicate | 고발하다 (kopalhata) – 17.1% | sue | accuse – 44.4% |
| | 부끄러워하다 (pwukkulewehata) – 5.9% | be ashamed of | be ashamed of – 3.0% |
| | 무서워하다 (mwusewehata) – 9.7% | fear | be scared of – 10.7% |
| | 의심하다 (uysimhata) – 9.4% | doubt | be suspicious of – 10.7% |
| | 수상히여기다 (swusanghiyekita) – 12.9% | suspect | be suspicious of – 13.3% |
| | 믿다 (mitta) – 4.3% | believe in | believe – 3.7% |
| | 얕보다 (yathpota) – 14.3% | look down on | belittle – 80.0% |

| | | |
|---|---|---|
| 불평하다 (pwulphyenghata) – 11.8% | grumble | complain – 60.0% |
| 비난하다 (pinanhata) – 12.5% | criticize | criticize – 14.3% |
| 싫어하다 (silhehata) – 5.9% | dislike | dislike – 3.3% |
| 탈락시키다 (thallaksikhita) – 10.7% | disqualify | eliminate – 7.7% |
| 시기하다 (sikihata) – 14.8% | be jealous of | envy – 0% |
| 질투하다 (cilthwuhata) – 3.4% | envy | envy – 10.7% |
| 부러워하다 (pwulewehata) – 5.0% | envy | envy – 3.6% |
| 두려워하다 (twulyewehata) – 20.6% | fear | fear – 6.3% |
| 해고하다 (haykohata) – 0% | fire | fire – 6.3% |
| 증오하다 (cungohata) – 11.1% | detest | hate – 12.9% |
| 혐오하다 (hyemohata) – 5.6% | loathe | hate – 19.4% |
| 도와주다 (towacwuta) – 16.1% | assist | help – 45.5% |
| 때리다 (ttaylita) – 8.3% | beat | hit – 28.6% |
| 죽이다 (cwukita) – 44.1% | kill | kill – 53.3% |
| 경멸하다 (kyengmyelhata) – 0% | despise | look down on – 6.9% |
| 불신하다 (pwulsinhata) – 0% | distrust | mistrust – 10.0% |
| 과대평가하다 (kwatayphyengkahata) – 9.7% | overestimate | overestimate – 33.3% |
| 가여워하다 (kayewehata) – 15.6% | pity | pity – 12.5% |
| 벌주다 (pelcwuta) – 0% | punish | punish – 0% |

| | | |
|---|---|---|
| 교체하다 (kyocheyhata) – 14.3% | substitute | replace – 6.1% |
| 책망하다 (chaykmanghata) – 8.8% | condemn | reproach – 35.5% |
| 존경하다 (conkyenghata) – 3.2% | respect | respect – 10.3% |
| 비웃다 (piwusta) – 0% | laugh at | ridicule – 24.1% |
| 야단치다 (yatanchita) – 8.3% | scold | scold – 3.0% |
| 꾸중하다 (kkwucwunghata) – 2.8% | rebuke | scold – 6.1% |
| 꾸짖다 (kkwucicta) – 3.1% | reproach | scold – 6.5% |
| 말리다 (mallita) – 0% | prevent | stop – 13.8% |
| 중단하다 (cwungtanhata) – 13.8% | stop | stop – 50.0% |
| 고소하다 (kosohata) – 5.7% | accuse | sue – 16.7% |
| 신뢰하다 (sinloyhata) – 16.0% | trust | trust – 10.7% |
| 주의주다 (cwuuycwuta) – 0% | note | warn – 35.0% |
| 경고하다 (kyengkohata) – 6.1% | warn | warn – 45.8% |
| 걱정하다 (kekcenghata) – 10.0% | worry about | worry about – 11.8% |

†AP = Agent-Patient verb; SE = Stimulus-Experiencer verb

* 매혹시키다 (mayhoksikhita) is the only SC item not including *-key ha*. It instead contains a lexical causative verb *–shiki* which means ‘to cause/force

APPENDIX B

Predicates used for Experiment 2

| Predicate type | | Predicate |
|---|---|---|
| Subject-biased | SC ($k$ = 12) | amuse, anger, annoy, bore, disappoint, embarrass, enchant, frighten, offend, please, scare, surprise |
| | non-SC ($k$ = 12) | apologize, approach, beg, call, cheat, confess, distract, invite, lie, lose, threaten, hurt |
| Object-biased (OB) ($k$ = 12) | | criticize, envy, fire, hate, hit, punish, replace, ridicule, scold, stop, sue, trust |
| Distractor ($k$ = 12) | | chat with, interview, know, listen to, live next to, resemble, see, study with, talk to, watch, work with, smile at |

APPENDIX C

Predicates used for Experiment 3 (verbs that appeared in Experiment 2 in bold face)

| Predicate type | | Predicate (frequency score) |
|---|---|---|
| Subject-biased | SC ($k$ = 12) | **anger** (5975), **annoy** (2218), bother (21866), comfort (5940), depress (1485), discourage (8071), distress (962), **embarrass** (3665), exhaust (7149), **frighten** (5322), **surprise** (14486), upset (11963) |
| | non-SC ($k$ = 12) | convince (23345), disturb (7247), encourage (46632), hurt (48181), impress (13611), inspire (23040), insult (3362), **offend** (6435), overwhelm (6874), provoke (6787), tease (6203), wound (8065) |
| Object-biased (OB) ($k$ = 24) | | admire (12437), applaud (5067), blame (28164), compliment (1392), condemn (8805), congratulate (2729), **criticize** (18207), despise (2519), dislike (3935), **envy** (2441), favor (17330), fear (29393), like (263201), love (159950), miss (72667), pity (1257), praise (10582), **punish** (9511), respect (15741), reward (7815), ridicule (1874), **scold** (2069), thank (130554), **trust** (28941) |

APPENDIX D

Items used for Experiment 4

| type | verb class | bias | context | critical | Question |
|---|---|---|---|---|---|
| non-SC | 31.1 | NP1 | Nicolas and Dean had a debate on politics. | Nicolas convinced Dean during the debate because he was easily persuaded. | Who was easily persuaded? |
| non-SC | 31.1 | NP1 | Nathan and Owen used to study together at the library. | Nathan disturbed Owen all the time because he needed help with his homework. | Who needed help with his homework? |
| non-SC | 31.1 | NP1 | Cecelia and Tracy were studying together for an exam. | Cecelia encouraged Tracy at that time because she was not confident in herself. | Who was not confident in herself? |
| non-SC | 31.1 | NP1 | Derek and George went to the same high school. | Derek hurt George at school because he was a sensitive and timid person. | Who was a sensitive and timid person? |
| non-SC | 31.1 | NP1 | Rachel and Elizabeth joined a dance class recently. | Rachel impressed Elizabeth in the first practice because she couldn't even do the moonwalk. | Who couldn't even do the moonwalk. |
| non-SC | 31.1 | NP1 | Samuel and Dylan shared a dream to learn how to play the drums. | Samuel inspired Dylan later on because he became a famous drummer. | Who became a famous drummer? |
| non-SC | 31.1 | NP1 | Valerie and Ella were classmates at school. | Valerie insulted Ella in class because she was wearing a funny-looking hat. | Who was wearing a funny-looking hat? |

| | | | | | |
|---|---|---|---|---|---|
| non-SC | 31.1 | NP1 | Peter and Jack shared a room in a dormitory. | Peter offended Jack all the time because he was sensitive and easily offended. | Who was sensitive and easily offended? |
| non-SC | 31.1 | NP1 | Mark and Tom decided to study together for an exam. | Mark overwhelmed Tom during that time because he asked too many questions. | Who asked too many questions? |
| non-SC | 31.1 | NP1 | Brenda and Nancy had hated each other since childhood. | Brenda provoked Nancy last Friday because she wanted to start a fight. | Who wanted to start a fight? |
| non-SC | 31.1 | NP1 | Lisa and Mara went to the same school. | Lisa teased Mara last Monday because she wanted the class to laugh. | Who wanted the class to laugh? |
| non-SC | 31.1 | NP1 | Malcolm and Brett were invited to a party last Saturday. | Malcolm wounded Brett at the party because he said something rude to him. | Who said something rude to him? |
| SC | 31.1 | NP1 | Amelia and Evelyn used to study together at a cafe. | Amelia angered Evelyn last Sunday because she didn't like the constant gossiping. | Who didn't like the constant gossiping? |
| SC | 31.1 | NP1 | Larry and Gavin started planning a welcome party for the freshmen. | Larry annoyed Gavin during the meeting because he had no patience for repetitions. | Who had no patience for repetitions? |
| SC | 31.1 | NP1 | Patrick and Curtis were solving math problems in class. | Patrick bothered Curtis every few minutes because he was the smartest kid in class. | Who was the smartest kid in class? |
| SC | 31.1 | NP1 | Ethel and Jasmine | Ethel comforted Jasmine | Who wanted to |

| | | | | | |
|---|---|---|---|---|---|
| | | | were very good friends. | last week because she wanted to be supportive. | be supportive? |
| SC | 31.1 | NP1 | Katherine and Barbara had been best friends for a long time. | Katherine depressed Barbara last night because she shared some bad news about their friend. | Who shared some bad news about their friend? |
| SC | 31.1 | NP1 | Grace and Hannah were both selected to do an internship at a well-known company. | Grace discouraged Hannah last night because she had heard bad things about the company. | Who had heard bad things about the company? |
| SC | 31.1 | NP1 | Joseph and Andrew were having dinner at home. | Joseph distressed Andrew during the dinner because he was anxious and got stressed out easily. | Who was anxious and got stressed out easily? |
| SC | 31.1 | NP1 | Lydia and Tania knew all of each other's secrets. | Lydia embarrassed Tania yesterday because she accidentally shared one of their secrets with someone else. | Who accidentally shared one of their secrets with someone else? |
| SC | 31.1 | NP1 | Erina and Nana meet for coffee every Wednesday. | Erina exhausted Nana last week because she kept telling the same story over and over. | Who kept telling the same story over and over? |
| SC | 31.1 | NP1 | Justin and Steve met each other at a Halloween party last year. | Justin frightened Steve at first sight because he was wearing a ghost costume. | Who was wearing a ghost costume? |
| SC | 31.1 | NP1 | Eliza and Natalie were supposed to meet for dinner at a restaurant. | Eliza surprised Natalie at the restaurant because she didn't expect to have a | Who didn't expect to have a party thrown for |

| | | | | | |
|---|---|---|---|---|---|
| | | | | party thrown for her. | her? |
| SC | 31.1 | NP1 | John and Chris shared a room in an apartment building. | John upset Chris every day because he couldn't bear the loud music. | Who couldn't bear the loud music? |
| OB | 31.2 | NP2 | Jackson and Logan liked to train in boxing when they were in college. | Jackson admired Logan back in those days because he knew how to take a punch. | Who knew how to take a punch? |
| OB | 31.2 | NP2 | Benjamin and Matthew went to the same high school. | Benjamin despised Matthew at school because he used to make fun of people. | Who used to make fun of people? |
| OB | 31.2 | NP2 | Molly and Dorothy went to a summer camp last August. | Molly disliked Dorothy at that time because she was rude and arrogant. | Who was rude and arrogant? |
| OB | 31.2 | NP2 | Bethany and Naomi both cared a lot about money. | Bethany envied Naomi all the time because she came from a poor family. | Who came from a poor family? |
| OB | 31.2 | NP2 | Gianna and Lauren took a history class last year. | Gianna favored Lauren at that time because she enjoyed being with smart students. | Who enjoyed being with smart students? |
| OB | 31.2 | NP2 | Harry and Fred got lost in the woods last summer. | Harry feared Fred at the time because he was known to have a short temper. | Who was known to have a short temper? |
| OB | 31.2 | NP2 | Austin and Burt met at a cocktail party last week. | Austin liked Burt right away because he really enjoyed hearing jokes. | Who really enjoyed hearing jokes? |
| OB | 31.2 | NP2 | Emily and Amanda got | Emily loved Amanda | Who was so |

| | | | | | |
|---|---|---|---|---|---|
| | | | to know each other when they were roommates. | whole-heartedly because she was so sweet and lovable. | sweet and lovable? |
| OB | 31.2 | NP2 | Olivia and Isabella were good friends at school. | Olivia missed Isabella at the class reunion because she was the only person who did not come. | Who was the only person who did not come? |
| OB | 31.2 | NP2 | Ethan and Michael both really wanted to learn to play guitar. | Ethan pitied Michael for some time because he had an easier time learning it. | Who had an easier time learning it? |
| OB | 31.2 | NP2 | Kara and Claire worked together at a trading company. | Kara respected Claire very much because she had never seen such a hard-working person. | Who had never seen such a hard-working person? |
| OB | 31.2 | NP2 | Ann and Gloria used to lend each other books when they lived in Toronto. | Ann trusted Gloria with the books because she knew whom to trust. | Who knew whom to trust? |
| OB | 33 | NP2 | David and Paul were both firefighters. | David applauded Paul at the party because he wished to express the station's gratitude. | Who wished to express the station's gratitude? |
| OB | 33 | NP2 | Rebecca and Janet had a fire in their kitchen the other day. | Rebecca blamed Janet after the fire because she didn't want to take responsibility herself. | Who didn't want to take responsibility herself? |
| OB | 33 | NP2 | Sarah and Bella saw each other at a family reunion. | Sarah complimented Bella in front of everyone because she was wearing a beautiful and elegant dress. | Who was wearing a beautiful and elegant dress? |

| | | | | | |
|---|---|---|---|---|---|
| OB | 33 | NP2 | Mason and William started a new business last year. | Mason condemned William at work because he didn't like mean people. | Who didn't like mean people? |
| OB | 33 | NP2 | Ronald and Bruce got their black belts in Taekwondo last week. | Ronald congratulated Bruce at the event because he had worked really hard for it. | Who had worked really hard for it? |
| OB | 33 | NP2 | Betty and Meilani were both on the environmental committee this semester. | Betty criticized Meilani at the last meeting because she always arrived very late. | Who always arrived very late? |
| OB | 33 | NP2 | Annabelle and Nora tried out for the cheerleading team. | Annabelle praised Nora in front of the girls because she enjoyed giving nice compliments. | Who enjoyed giving nice compliments? |
| OB | 33 | NP2 | Ken and Anthony are brothers and share a room. | Ken punished Anthony last week because he broke his favorite toy. | Who broke his favorite toy? |
| OB | 33 | NP2 | Beth and Diana worked on a team project at school. | Beth rewarded Diana after the project because she had promised to do so. | Who had promised to do so? |
| OB | 33 | NP2 | Kevin and Jonathan were taking a dance lesson at school. | Kevin ridiculed Jonathan during class because he had never seen such bad dance moves. | Who had never seen such bad dance moves? |
| OB | 33 | NP2 | Jeremy and Tucker were each supposed to clean a part of the house yesterday. | Jeremy scolded Tucker this morning because he hadn't cleaned the living room. | Who hadn't cleaned the living room? |

| | | | | | |
|---|---|---|---|---|---|
| OB | 33 | NP2 | Aaron and Jordan celebrated Christmas together. | Aaron thanked Jordan at the party because he had put up some beautiful decorations. | Who had put up some beautiful decorations? |
| Filler | - | - | Caroline and Julia were watching a horror movie last night. | Caroline covered her face with her hands, and Julia screamed with fear the whole time. | Who screamed with fear the whole time? |
| Filler | - | - | Adam and Matt made a scene at the airport. | Adam yelled with excitement, and Matt shed tears of joy. | Who yelled with excitement? |
| Filler | - | - | Hunter and Luis were planning their trip to Paris environmental committee this semester. | Hunter made a list of places to visit, and Luis booked airplane tickets. | Who booked airplane tickets? |
| Filler | - | - | Alice and Ariana were both at the beach last Saturday. | Alice was swimming in the sea, and Ariana was putting on her sunscreen. | Who was swimming in the sea? |
| Filler | - | - | Max and Charlie travelled to Spain last winter. | Max packed the suitcases, and Charlie took care of the train and airplane tickets. | Who took care of the train and airplane tickets? |
| Filler | - | - | Emma and Ava worked side by side in their office last year. | Emma took care of paperwork, and Ava was in charge of computer maintenance. | Who took care of paperwork? |
| Filler | - | - | Natalia and Lily were at a welcome party for new students. | Natalia tried to start a conversation but Lily didn't seem to notice. | Who didn't seem to notice? |

| | | | | | |
|---|---|---|---|---|---|
| Filler | - | - | Tyler and Parker played basketball in the park yesterday. | Tyler scored many goals, but Parker couldn't even throw a pass properly. | Who scored many goals? |
| Filler | - | - | Jessica and Savannah were taking math class together last semester. | Jessica mastered all math equations, but Savannah could not solve a single problem. | Who could not solve a single problem? |
| Filler | - | - | Lucas and Oliver worked at the same summer camp this summer. | Lucas wanted to put up the tent by the river, but Oliver put up the tent on the hillside. | Who wanted to put up the tent by the river? |
| Filler | - | - | Hudson and Thomas accidentally stumbled upon the pile of presents in the closet on Tuesday. | Hudson asked what they were, but Thomas pretended not to know about them. | Who pretended not to know about them? |
| Filler | - | - | Landon and Jacob got drunk yesterday. | Landon kept talking about his girlfriend, but Jacob got bored and went outside. | Who kept talking about his girlfriend? |
| Filler | - | - | Carter and Daniel decided to stop smoking last year. | Carter did not keep the promise to Daniel to quit smoking. | Who did not keep the promise to quit smoking? |
| Filler | - | - | Linda and Jennifer arrived in court early last Monday. | Linda talked to Jennifer about the change of plans before the trial began. | Who talked about the change of plans? |
| Filler | - | - | Martin and Eric always looked out for each other at school. | Martin sent a text message to Eric late last night to warn him about the new bully in their class. | Who sent a text message to warn about the new bully in their |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | class? |
| Filler | - | - | Lillian and Lucy are ballet instructors at the same dance school. | Lillian gave some advice to Lucy about how to train their students without using harsh methods. | Who gave advice about how to train their students? |
| Filler | - | - | Albert and Sean were throwing a birthday party for their friend. | Albert smiled at Sean to give a signal to set off the firecrackers. | Who gave a signal to set off the firecrackers? |
| Filler | - | - | Violet and Ashley had a big fight last week. | Violet wrote to Ashley last night to apologize about the fight. | Who wrote to apologize about the fight? |
| Filler | - | - | Lily and Mila got into a fight with some guys in a parking lot. | Lily fought the guys alone, so Mila was able to escape from the fight safely. | Who was able to escape from the fight safely? |
| Filler | - | - | Charles and Ivan studied at the library last Saturday. | Charles said he wanted to sit alone, so Ivan moved to another table. | Who wanted to sit alone? |
| Filler | - | - | Jaxon and Ryan spent a lot of time together last year. | Jaxon broke up with his girlfriend, so Ryan came to give words of consolation. | Who came to give words of consolation? |
| Filler | - | - | Nina and Margaret worked at the same shoe shop a couple of years ago. | Nina didn't know how to talk to customers, so Margaret took care of serving customers. | Who took care of serving customers? |
| Filler | - | - | Roger and Luke were taking an exam at school. | Roger was tapping his foot the whole time, so Luke couldn't pay attention to the exam. | Who was tapping his foot? |
| Filler | - | - | Sofia and Chloe went to the market to buy | Sofia forgot to bring her wallet, so Chloe paid for | Who paid for all the things they |

| | | | | | |
|---|---|---|---|---|---|
| | | | some food. | all the things they bought | bought? |
| Filler | - | - | Nora and Camila played soccer on Tuesday. | Nora dribbled the ball as Camila was running toward the goal. | Who was running toward the goal? |
| Filler | - | - | Robert and Miles lived next to each other. | Robert was mowing the lawn in the morning as Miles went out to take a walk. | Who was mowing the lawn? |
| Filler | - | - | Bill and Noah finally went skydiving this week. | Bill was frozen with fear when Noah opened the airplane door to jump. | Who was frozen with fear? |
| Filler | - | - | Easton and Jeremiah were about to cross the street. | Easton shouted a warning before Jeremiah was almost hit by a truck | Who was almost hit by a truck? |
| Filler | - | - | Nolan and Jake both worked as computer technicians. | Nolan took notes, while Jake explained how to remove computer viruses. | Who took notes? |
| Filler | - | - | Luna and Victoria participated in the race two days ago. | Luna took the lead in the race when Victoria slowed down from fatigue. | Who slowed down from fatigue? |
| Filler | - | - | Liam and Jayden used to be friends until their big fight. | As Liam asked for help with a problem in math class, Jayden pretended not to listen. | Who pretended not to listen? |
| Filler | - | - | Sophia and Mia were both at the engagement party on Friday. | As Sophia was talking to other people, Mia was eating cake. | Who was eating cake? |
| Filler | - | - | Eva and Elena hadn't seen each other in a long time. | When Eva showed up at a school reunion, Elena shouted with joy. | Who shouted with joy? |

| | | | | | |
|---|---|---|---|---|---|
| Filler | - | - | Helen and Lidia visited Disneyland last Saturday. | When Helen was talking to Mickey Mouse, Lidia took a picture of Donald Duck. | Who took a picutre of Donald Duck? |
| Filler | - | - | Edward and Connor were at hockey practice last night. | When Edward was injured and fell to the ground, Connor shouted out for help | Who shouted out for help? |
| Filler | - | - | Aria and Scarlett argued with each other about money yesterday. | When Aria first apologized later that night, Scarlett accepted the apology with a smile. | Who accepted the apology with a smile? |
| Filler | - | - | Sara and Mary woke up early in the morning. | While Sara went out for a walk, Mary stayed at home and ate breakfast. | Who stayed at home and ate breakfast? |
| Filler | - | - | Emilia and Maria went over their chemistry homework yesterday. | While Maria struggled with the first question Emilia finished the homework in an hour. | Who struggled with the first question? |
| Filler | - | - | Henry and Brian visited the new exhibition at the botanical gardens. | While Henry was looking at the tulips, Brian went to see the other flowers. | Who went to see the other flowers? |
| Filler | - | - | Ivy and Athena decided to buy a birthday present for their friend. | While Ivy decorated a room for the party, Athena went to a department store to buy a ring. | Who decorated a room for the party? |
| Filler | - | - | Ian and Alex were both suspected of stealing the trophy. | While Ian looked anxious and worried, Alex made an innocent face. | Who made an innocent face? |
| Filler | - | - | Asher and Richard | While Asher was preparing | Who was |

| | | | | | |
|---|---|---|---|---|---|
| | | | participated in a summer camp last week. | a meal, Richard set up a tent in the backyard. | preparing a meal? |
| Filler | - | - | Karen and Sandra had a fun day on Christmas. | Karen baked cookies and cakes, while Sandra decorated the Christmas tree. | Who baked cookies and cakes? |
| Filler | - | - | Anna and Leah went to a fastfood restaurant for lunch yesterday. | Anna was looking over the menu, while Leah went to the counter to order. | Who went to the counter to order? |
| Filler | - | - | Phillip and Leo were left alone at home on Sunday. | Phillip watched TV all day long, while Leo studied for the upcoming exam. | Who watched TV all day long? |
| Filler | - | - | Jeff and Clark decided to clean their house on Sunday. | Jeff cleaned up the kitchen and living room, while Clark raked the backyard. | Who raked the backyard? |
| Filler | - | - | Evan and Donald visited the art museum yesterday. | Evan looked at classic paintings, while Donald was looking at statues. | Who looked at classic paintings? |
| Filler | - | - | Brooklyn and Maya both tried out for the volleyball team on Saturday. | Brooklyn practiced with other players, while Maya practiced alone in the gym. | Who practiced with other players? |

## REFERENCES


Altarriba, J. (1992). The representation of translation equivalents in bilingual memory. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 157–174). Amsterdam: Elsevier.

Altarriba, J., & Basnight-Brown, D. M. (2007). Methodological considerations in performing semantic- and translation-priming experiments across languages. *Behavior Research Methods, 39*, 1–18.

Ariel, M. (1990). *Accessing NP antecedents*. London: Routledge.

Arnold, J. E. (1998). *Reference form and discourse patterns* (Unpublished doctoral dissertation). Stanford University, CA.

Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes, 31*, 137–162.

Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass, 4*, 187–203.

Au, T. (1986). A verb is worth a thousand words: The causes and consequences of interpersonal events implicit in language. *Journal of Memory and Language, 25*, 104–122.

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*, 457–474.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Basnight-Brown, D. M., & Altarriba, J. (2007). Differences in semantic and translation priming across languages: The role of language direction and language dominance. *Memory & Cognition, 35*, 953–965.

Bernolet, S., Hartsuiker, R. J., & Pickering, M. J. (2013). From language-specific to shared syntactic representations: The influence of second language proficiency on syntactic sharing in bilinguals. *Cognition, 124*, 287–306.

Bott, O., & Solstad, T. (2014). From verb to discourse: A novel account of implicit causality. In B. Hemforth, B. Schmiedtová, & C. Fabricius-Hansen (Eds.), *Psycholinguistic approaches to meaning and understanding across languages* (pp. 213–251), Cham, Switzerland: International Publishing.

Brenders, P., Van Hell, J. G., & Dijkstra, A. (2011). Word recognition in child second language learners: Evidence from cognates and false friends. *Journal of Experimental Child Psychology, 109*, 383–396.

Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal, 64*, 311–317.

Brown, R., & Fish, D. (1983). The psychological causality implicit in language. *Cognition, 14*, 237–273.

Brysbaert, M., van Dyck, G., & van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology. Human Perception and Performance, 25*, 137–148.

Cai, Z. G., Pickering, M. J., Yan, H., & Branigan, H. P. (2011). Lexical and syntactic representations in closely related languages: Evidence from Cantonese–Mandarin bilinguals. *Journal of Memory and Language, 65*, 431–445.

Canseco-Gonzalez, E., Brehm, L., Brick, C. A., Brown-Schmidt, S., Fischer, K., & Wagner, K. (2010). Carpet or carcel: The effect of age of acquisition and language mode on bilingual lexical access. *Language and Cognitive Processes, 25*, 669–705.

Caramazza, A., Grober, E., Garvey, C., & Yates, J. (1977). Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior, 16*, 601–609.

Chae, H-R. (1997). Verbal nouns and light verbs in Korean. *Language Research, 33*, 581–600.

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1029–1040.

Cheng, W., & Almor, A. (2017). The effect of implicit causality and consequentiality on nonnative pronoun resolution. *Applied Psycholinguistics, 38*, 1–26.

Cheng, W., & Almor, A. (2018). A Bayesian approach to establishing coreference in second language discourse: Evidence from implicit causality and consequentiality verbs. *Bilingualism: Language and Cognition*, doi:10.1017/S136672891800055X. Published online by Cambridge University Press, May 16, 2018.

Clahsen, H., & Felser, C. (2006). Continuity and shallow structures in language processing. *Applied Psycholinguistics, 27*, 107–26.

Commandeur, E. (2010). *Implicit causality and implicit consequentiality in language comprehension (*Unpublished doctoral dissertation). Tilburg University, Tilburg.

Contemori, C., & Dussias, E. P. (2018). Prediction at the discourse level in L2 English speakers: an eye-tracking study. In A. B. Bertolini & M. J. Kaplan (Eds.), *Proceedings of the 42nd annual Boston University conference on language development* (pp. 159–171). Somerville, MA: Cascadilla.

Cook, V. (1997). The consequences of bilingualism for cognitive processing. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism* (pp. 279–299). Mahwah, NJ: Erlbaum.

Corrigan, R. (2001). Implicit causality in language: Event participants and their interactions. *Journal of Language and Social Psychology, 20*, 285–320.

Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1283–1296.

Cozijn, R., Commandeur, E., Vonk, W., & Noordman, L. G. (2011). The time course of the use of implicit causality information in the processing of pronouns: A visual world paradigm study. *Journal of Memory and Language, 64*, 381–403.

Crawley, R. A., Stevenson, R. J., & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research, 19*, 245–264.

Crinean, M., & Garnham, A. (2006). Implicit causality, implicit consequentiality and semantic roles. *Language and Cognitive Processes, 21*, 636–648.

Cunnings, I., Fotiadou, G., & Tsimpli, I. (2017). Anaphora resolution and reanalysis during L2 sentence processing: Evidence from the visual world paradigm. *Studies in Second Language Acquisition, 39*, 621-652.

Davies, M. (2009). The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics, 14*, 159–190.

De Groot, A.M.B., & Nas, G. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30*, 90–123.

Dery, J. E., & Bittner, D. (2016). Time and causation in Discourse: temporal proximity, implicit causality, and remention biases. *Journal of Psycholinguistic Research, 45*, 883–899.

Dijkgraaf, A., Hartsuiker, R. J., & Duyck, W. (2017). Predicting upcoming information in native-language and non-native-language auditory word recognition. *Bilingualism: Language and Cognition*, 20, 917–930.

Dijkstra, T. (2005). Bilingual visual word recognition and lexical access. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 179–201). New York: Oxford University Press.

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*, 175–197.

Dijkstra, A., Wahl, A., Buytenhuijs, F., van Halem, N., Al- jibouri, Z., de Korte, M., & Rekké, S. (2018). Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, doi:10.1017/S1366728918000287. Published online by Cambridge University Press, July 6, 2018.

Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011a). Phonology by itself: Masked phonological priming effects with and without orthographic overlap. *Journal of Cognitive Psychology, 23*, 185–203.

Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011b). Two words, one meaning: Evidence for automatic co-activation of translation equivalents. *Frontiers in Psychology, 2*, 1–20.

Dussias, P. E. (2003). Syntactic ambiguity resolution in L2 learners. *Studies in Second Language Acquisition, 25*, 529–557.

Dussias, P. E., & Piñar, P. (2010). Effects of reading span and plausibility in the reanalysis of wh-gaps by Chinese-English L2 speakers. *Second Language Research, 26*, 443–472.

Dussias, P. E., & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish–English bilinguals. *Bilingualism: Language and Cognition,10*, 101–116.

Duyck, W. (2005). Translation and associative priming with cross-lingual pseudohomophones: evidence for nonselective phonological activation in bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1340–1359.

Duyck, W., Diependaele, K., Drieghe, D., & Brysbaert, M. (2004). The size of the cross-lingual masked phonological priming effect does not depend on second language proficiency. *Experimental Psychology, 51*, 1–9.

Duyck, W., Van Assche, E., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 663–679.

Ellis, N. C., Römer, U., & O'Donnell, M. B. (2016). *Usage-based approaches to language acquisition and processing: Cognitive and corpus investigations of construction grammar*. Language Learning Monograph Series. West Sussex: Wiley-Blackwell.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research, 25*, 57–70.

Featherstone, C. R., & Sturt, P. (2010). Because there was a cause for concern: An investigation into a word-specific prediction account of the implicit-causality effect. *The Quarterly Journal of Experimental Psychology, 63*, 3–15.

Ferstl, E. C., Garnham, A., & Manouilidou, C. (2011). Implicit causality bias in English: A corpus of 300 verbs. *Behavior Research Methods, 43*, 124–135.

Fillmore, C. J., Kay, P., & O'Connor, M. C. (1988). Regularity and idiomaticity in grammatical constructions: the case of let alone. *Language, 64*, 501–38.

Foucart, A., Romero-Rivas, C., Gort, B. L., & Costa, A. (2016). Discourse comprehension in L2: Making sense of what is not explicitly said. *Brain and Language, 163*, 32–41.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In J. Altarriba & R. Herridia (Eds.), *Bilingual sentence processing* (pp. 218–236). Amsterdam: Elsevier.

Fukumura, K., & Van Gompel, R. P. (2010). Choosing anaphoric expressions: Do people take into account likelihood of reference? *Journal of Memory and Language, 62*, 52–66.

Garnham, A., Traxler, M., Oakhill, J., & Gernsabacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory and Language, 35*, 517–543.

Garvey, C., & Caramazza, A. (1974). Implicit causality in verbs. *Linguistic inquiry, 5*, 459–464.

Garvey, C., Caramazza, A., & Yates, J. (1974). Factors influencing assignment of pronoun antecedents. *Cognition, 3*, 227–243.

Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Lawrence Erlbaum.

Gernsbacher, M. A., Hargreaves, D.J., & Beeman, M. (1989). Building and accessing clausal representations: The advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language, 28*, 735–755.

Givón, T. (1983). Topic continuity in discourse: an introduction. In T. Givón (Ed.), *Topic continuity in discourse: a quantitative cross-language study*, (pp. 1–42). Amsterdam: John Benjamins Publishing.

Goikoetxea, E., Pascual, G., & Acha, J. (2008). Normative study of the implicit causality of 100 interpersonal verbs in Spanish. *Behavior Research Methods, 40*, 760–772.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Goldberg, A. E. (2013). Argument structure constructions versus lexical rules or derivational verb templates. *Mind & Language, 28*, 435–465.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 1122–1139.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science, 17*, 311–347.

Greene, S. B., & McKoon, G. (1995). Telling something we can't know: Experimental approaches to verbs exhibiting implicit causality. *Psychological science, 6*, 262–270.

Grüter. T, Rohde, H., & Schafer. A. J. (2017). Coreference and discourse coherence in L2: The roles of grammatical aspect and referential form. *Linguistic Approaches to Bilingualism, 7*, 199–229.

Grüter, T., Takeda, A., Rohde, H., & Schafer, A. J. (2018). Intersentential coreference expectations reflect mental models of events. *Cognition, 177*, 172–176.

Gundel, J. K., Hedberg, N., & Zacharaski, R. (1993). Cognitive status and the form of referring expressions. *Language, 69*, 274–307.

Hahn, H. (2011). Korean EFL learners' underpassivization and overpassivization: Psych verbs and unaccusatives. *Foreign Language Education Research, 14*, 54–76.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research, 30*, 251–266.

Haigh, C. A., & Jared, D. (2007). The activation of phonological representations by bilinguals while reading silently: Evidence from interlingual homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 623–644.

Han, N. R. (2006). *Korean zero pronouns: analysis and resolution* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia.

Hartshorne, J. K. (2014). What is implicit causality? *Language, Cognition and Neuroscience, 29*, 804–824.

Hartshorne, J. K., O'Donnell, T. J., & Tenenbaum, J. B. (2015). The causes and consequences explicit in verbs. *Language, Cognition and Neuroscience, 30*, 716–734.

Hartshorne, J. K., & Snedeker, J. (2013). Verb argument structure predicts implicit causality: The advantages of finer-grained semantics. *Language and Cognitive Processes, 28*, 1474–1508.

Hartshorne, J. K., Sudo, Y., & Uruwashi, M. (2013). Are implicit causality pronoun resolution biases consistent across languages and cultures? *Experimental Psychology, 60*, 179–196.

Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish-English bilinguals. *Psychological Science, 15*, 409–414.

Hawkins, R., & Chan C. (1997). The partial availability of Universal Grammar in second language acquisition: the 'failed functional features hypothesis'. *Second Language Research, 13*, 187–226.

Helms-Park, R. (2001). Evidence of lexical transfer in learner syntax: The acquisition of English causatives by speakers of Hindi-Urdu and Vietnamese. *Studies in Second Language Acquisition, 23*, 71–102.

Hobbs, J. R. (1979). Coherence and coreference. *Cognitive Science, 3*, 67–90.

Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua, 120*, 901–931.

Hopp, H. (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research, 29*, 33–56.

Hopp, H. (2016). The timing of lexical and syntactic processes in second language sentence comprehension. *Applied Psycholinguistics, 37*, 1253–1280.

Hopp, H. (2017). Cross-linguistic lexical and syntactic co-activation in L2 sentence processing. *Linguistic Approaches to Bilingualism, 7*, 96–130.

Hunston, S., & Francis, G. (2000). *Pattern grammar: A corpus driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.

Itzhak, I., & Baum, S. R. (2015). Misleading Bias-Driven Expectations in Referential Processing and the Facilitative Role of Contrastive Accent. *Journal of Psycholinguistic Research, 44*, 623–650.

Jacob, G., Katsika, K., Family, N., & Allen, S. E. (2017). The role of constituent order and level of embedding in cross-linguistic structural priming. *Bilingualism: Language and Cognition, 20*, 269–282.

Jaeger, T. F. (2008). Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language, 59*, 434–446.

Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language, 44*, 2–31.

Jared, D., & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism: Language and Cognition, 5*, 225–239.

Jarvis, S. (2009). Lexical transfer. In A. Pavlenko (Ed.), *The bilingual mental lexicon: Interdisciplinary approaches* (pp. 99–124). Clevedon: Multilingual Matters.

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism, 4*, 257–282.

Kantola, L., & van Gompel, R. P. (2011). Between-and within-language priming is the same: Evidence for shared bilingual syntactic representations. *Memory & Cognition, 39*, 276–290.

Kasof, J., & Lee, J. Y. (1993). Implicit causality as implicit salience. *Journal of Personality and Social Psychology, 65*, 877–891.

Kehler, A., Kertz, L., Rohde, H., & Elman, J. L. (2008). Coherence and coreference revisited. *Journal of Semantics*, *25*, 1–44.

Kehler, A., & Rohde, H. (2015). Pronominal reference and pragmatic enrichment: A Bayesian account. *Proceedings of the 37th Annual Conference of the Cognitive Science Society (CogSci2015)* (pp. 1063–1068). Pasadena, CA, USA.

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition, 14*, 185–209.

Kim, H., & Grüter, T. (2018). Cross-linguistic activation of implicit causality biases in Korean learners of English. *Bilingualism: Language and Cognition*, doi: https://doi.org/10.1017/S1366728918000561. Published online by Cambridge University Press, June 4, 2018.

Kim, J., & Davis, C. (2003). Task effect in masked cross-script translation and priming. *Journal of Memory and Language, 49*, 484–499.

Kim, M. K. (1999). Conditions on Deletion in Korean based on Information Packaging, *Discourse and Cognition, 1*, 61–88

Kipper, K., Korhonen, A., Ryant, N., & Palmer, M. (2008). A large-scale classification of English verbs. *Language Resources and Evaluation Journal, 42*, 21–40.

Koornneef, A., Dotlačil, J., van den Broek, P., & Sanders, T. (2016). The influence of linguistic and cognitive factors on the time course of verb-based implicit causality. *The Quarterly Journal of Experimental Psychology*, *69*, 455–481.

Koornneef, A. W., & Van Berkum, J.J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *Journal of Memory and Language*, *54*, 445–465.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149–174.

Kroll, J. F., van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism: Language and Cognition, 13*, 373–381.

Kweon, S. O. (2008). A corpus-based analysis of the ordering of main and adverbial clauses in Korean learners of English. *Korean Journal of Applied Linguistics 24*, 159–176.

Kweon, S. O. (2011). Processing null and overt pronoun subject in ambiguous sentences in Korean. *International Journal of Linguistics, 3*, 1–12.

Lakoff, G. (1987). *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*. Chicago, IL: University of Chicago Press.

Lee, H-S. (2017). 'X-key hata'satongkwusengkwa kwusengmwunpepcek cepkun [A study on a causative construction of 'X-key ha-ta' from the viewpoint of Construction Grammar]. *Studies in Linguistics, 42*, 71–92.

Lee, I., Lee, S., & Chae, W. (1997). *Korean Language*. Shingu Publishing.

Lee, K. (1993). *A Korean grammar on semantic pragmatic principles*. Seoul: Hankwuk Munwhasa.

Lee, K. S. (1996). Syntactic causatives in Korean: clause union or not? *Language, Information and Computation* (PACLIC 11), 83–92.

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods, 44*, 325–343.

Lemhöfer, K., Dijkstra, T., & Michel, M. C. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes, 19*, 585–611.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago, IL: University of Chicago Press.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 35, 381– 390.

Liu, R. & Nicol. J. (2010). Online processing of anaphora by advanced English learners. In M. T. Prior, Y. Watanabe & S.-K. Lee (Eds.), *Proceedings of the 2008 Second Language Research Forum* (pp. 150–165). Somerville, MA: Cascadilla Proceedings Project.

Long, D.L., & De Ley, L. (2000). Implicit causality and discourse focus: The interaction of text and reader characteristics in pronoun resolution. *Journal of Memory and Language, 42*, 545–570.

Majid, A., Sanford, A. J., & Pickering, M. J. (2007). The linguistic description of minimal social scenarios affects the extent of causal inference making. *Journal of Experimental Social Psychology, 43*, 918–932.

Mannetti, L., & de Grada, E. (1991). Interpersonal verbs: Implicit causality of action verbs and contextual factors. *European Journal of Social Psychology, 21*, 429–443.

Marian, V., & Spivey, M. (2003). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition, 6*, 97–115.

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception, & Psychophysics, 53*, 372–380.

McDonald, J. L., & MacWhinney, B. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. *Journal of Memory and Language, 34*, 543–543.

Meriläinen, L. (2006) *Lexical transfer errors in the written English of Finnish upper secondary school students* (Unpublished doctoral dissertation). University of Joensuu, Finland.

Michael, E., & Gollan, T. (2005) Being and becoming bilingual: Individual differences and consequences for language production. In J. Kroll & A.M.B. De Groot (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches* (pp. 389–407). Oxford: Oxford University Press.

Michaelis, L. A., & Lambrecht, K. (1996). Toward a Construction-based model of language function: the case of nominal extraposition. *Language, 72*, 215–47.

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009). Masked repetition and translation priming in second language learners: A window on the time-course of form and meaning activation using ERPs. *Psychophysiology, 46*, 551–565.

Miltsakaki, E. (2007). A rethink of the relationship between salience and anaphora resolution. In A. Branco (Ed.), *Proceedings of the 6th discourse anaphora and anaphor resolution colloquium* (p. 91–96). Lago, Portugal.

Miyao, M. (2017). *The processing of referential expressions in discourse by Chinese, English, and Japanese native speakers and by Chinese and Japanese learners of English* (Unpublished doctoral dissertation). University of Hawai'i at Manoa.

Nakayama, M., Verdonschot, R. G., Sears, C. R., & Lupker, S. J. (2014). The masked cognate translation priming effect for different-script bilinguals is modulated by the phonological similarity of cognate words: Further support for the phonological account. *Journal of Cognitive Psychology, 26*, 714–724.

O'Grady, W. (1991). *Categories and case: The sentence structure of Korean*. John Benjamins Publishing.

Omaki, A., & Schulz, B. (2011). Filler-gap dependencies and island constraints in second-language sentence processing. *Studies in Second Language Acquisition, 33*, 563–588.

Pan, H.-Y., Schimke, S., & Felser, C. (2015). Referential context effects in non-native relative clause ambiguity resolution. *International Journal of Bilingualism, 19*, 298–313.

Park, J. W. (1994). *Morphological causatives in Korean: problems in grammatical polysemy and constructional relations* (Unpublished doctoral dissertation). University of California at Berkeley.

Park, K. (2009). Effects of implicit causality: Evidence from Korean interpersonal verbs. *The Korean Journal of Cognitive and Biological Psychology. 21*, 191–213.

Park, Y. (1999). The Korean connective nuntey in conversational discourse. *Journal of Pragmatics, 31*, 191–218.

Pickering, M. J., & Majid, A. (2007). What are implicit causality and consequentiality?. *Language and Cognitive Processes, 22*, 780–788.

Pliatsikas, C., & Marinis, T. (2013). Processing empty categories in a second language: When naturalistic exposure fills the (intermediate) gap. *Bilingualism: Language and Cognition, 16*, 167–182.

Poarch, G. J., & Van Hell, J. G. (2012). Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology, 111*, 419–438.

Prior, A., Degani, T., Awawdy, S., Yassin, R., & Korem, N. (2017). Is susceptibility to cross-language interference domain specific? *Cognition, 165*, 10–25.

Pyykkönen, P., & Järvikivi, J. (2010). Activation and persistence of implicit causality information in spoken language comprehension. *Experimental Psychology, 57*, 5–16.

Quirk, R., Greenbaum, S., Leech, G., Svartvik, J., & Crystal, D. (1985). *A comprehensive grammar of the English language*. London: Longman.

R Core Team (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ringbom, H. (2001). Lexical transfer in L3 production. In J. Cenoz, B. Hufeisen & U. Jessner (Eds.), *Cross-linguistic Influence in Third Language Acquisition: Psycholinguistic Perspectives* (pp. 59–68). Clevedon: Multilingual Matters.

Ringbom, H. (2007). *Cross-linguistic Similarity in Foreign Language Learning*. Clevedon: Multilingual Matters.

Roberts, L., Gullberg, M., & Indefrey, P. (2008). Online pronoun resolution in L2 discourse: L1 influence and general learner effects. *Studies in Second Language Acquisition, 30,* 333–357.

Roelofs, A., Meyer, A. & Levelt, W. (1998) A case for the lemma/lexeme distinction in models of speaking: Comment on Caramazza and Miozzo (1997). *Cognition, 69,* 219–230.

Roh, J. E., & Lee, J. H. (2003). An empirical study for generating zero pronoun in Korean based on cost-based centering model. In C. Bow & B. Hughes (Eds.), *Proceedings of Australasian Language Technology Association* (pp. 90–97). Melbourne: University of Melbourne.

Rohde, H., & Ettlinger, M. (2011). Integration of pragmatic and phonetic cues in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 967–983.

Rohde, H., & Kehler, A. (2014). Grammatical and information-structural influences on pronoun production. *Language, Cognition and Neuroscience*, *29*, 912–927.

Rohde, H., Kehler, A., & Elman, J. L. (2006). Event structure and discourse coherence biases in pronoun interpretation. In R. Sun (Ed.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 617–622). Mahwah, NJ: Lawrence Erlbaum Associates.

Rohde, H., Levy, R., & Kehler, A. (2011) Anticipating explanations in relative clause processing. *Cognition, 118,* 339–358.

Saito, K. (2017). Effects of sound, vocabulary, and grammar learning aptitude on adult second language speech attainment in foreign language classrooms. *Language Learning, 67,* 665–693.

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,* 548– 560.

Schoonbaert, S., Hartsuiker, R. J., & Pickering, M. J. (2007). The representation of lexical and syntactic information in bilinguals: Evidence from syntactic priming. *Journal of Memory and Language, 56,* 153–171.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55,* 197–212.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics, 10,* 209–241.

Sheldon, A. (1974). Role of parallel function in acquisition of relative clauses in English. *Journal of Verbal Learning and Verbal Behavior, 13*, 272–281.

Shibatani, M., & Chung, S. Y. (2001). Japanese and Korean Causatives Revisited. In A. Noriko & S. Strauss (Eds.), *Japanese/Korean Linguistics Vol 10,* (pp. 32–49). Stanford Linguistic Association, Stanford, CA: Center for the Study of Language and Information (CSLI) Publications.

Sohn, H. M. (2001). *The Korean Language*. Cambridge: Cambridge University Press.

Solstad, T., & Bott, O. (2017). Causality and causal reasoning in natural language. In M. R. Waldmann (Ed.), *The Oxford Handbook of Causal Reasoning* (pp. 619–644). Oxford: Oxford University Press.

Sorace, A., & Filiaci, F. (2006). Anaphora resolution in near-native speakers of Italian. *Second Language Research, 22,* 339–368.

Spivey, M. J. & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10,* 281–284.

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes, 9,* 473–592.

Stewart, A. J., Pickering, M. J., & Sanford, A. J. (1998). Implicit consequentiality. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society* (pp. 1031–1036). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Stewart, A. J., Pickering, M. J., & Sanford, A. J. (2000). The time course of the influence of implicit causality information: Focusing versus integration accounts. *Journal of Memory and Language, 42*, 423–443.

Trenkic, D., Mirkovic, J., & Altmann, G. (2014). Real-time grammar processing by native and non-native speakers: Constructions unique to the second language. *Bilingualism: Language and Cognition, 17*, 237–257.

Trousdale, G., & Hoffmann, T. (2013). *Oxford handbook of construction grammar*. Oxford, UK: Oxford University Press.

Tulving, E., & Colotla, V. A. (1970). Free recall of trilingual lists. *Cognitive Psychology, 1*, 86–98.

Ueno, M., & Kehler, A. (2016). Grammatical and pragmatic factors in the interpretation of Japanese null and overt pronouns. *Linguistics, 54*, 1165–1221.

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language, 64*, 88–107.

Van Assche, E., Duyck, W., & Brysbaert, M. (2013). Verb processing by bilinguals in sentence contexts: The effect of cognate status and verb tense. *Studies in Second Language Acquisition, 35*, 237–259.

van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9*, 780–789.

van Hell, J. G., & Tanner, D. (2012). Second language proficiency and cross-language activation. *Language Learning, 62*, 148–171.

Walker, M., Cote, S., & Iida, M. (1994). Japanese discourse and the process of centering. *Computational linguistics, 20*, 193–232.

Wang, X., & Forster, K. (2015). Is translation priming asymmetry due to partial awareness of the prime? *Bilingualism: Language and Cognition, 18*, 657–669.

Wang, X., Wang, J., & Malins, J. G. (2017). Do you hear ‘feather’ when listening to ‘rain’? Lexical tone activation during unconscious translation: Evidence from Mandarin-English bilinguals. *Cognition, 169*, 15–24.

Warren, T., McConnell, K., & Rayner, K. (2008). Effects of context on eye movements when reading about possible and impossible events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34,* 1001–1010.

Weber, A., & Broersma, M. (2012). Spoken word recognition in second language acquisition. In C. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 5368–5375). Hoboken, NJ, USA: John Wiley & Sons, Inc.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*, 1–25.

Witzel, J., Witzel, N., & Nicol, J. (2012). Deeper than shallow: Evidence for structure-based parsing biases in second-language sentence processing. *Applied Psycholinguistics, 33*, 419–456.

Yang, H.–K. (2010). Linguistic systems of target language as organizational foundations of foreign language teaching. *The Education of Korean Language, 133*, 63–81.

Zhao, X., Li, P., Liu, Y., Fang, X., & Shu, H. (2011). Cross-language priming in Chinese-English bilinguals with different second language proficiency levels. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 801–806). Austin, TX: Cognitive Science Society.

Zhou, H., Chen, B., Yang, M., & Dunlap, S. (2010). Language nonselective access to phonological representations: Evidence from Chinese-English bilinguals. *Quarterly Journal of Experimental Psychology*, *63*, 2051–2056.