CrossMark

Applied Psycholinguistics **36** (2015), 1283–1315
doi:10.1017/S0142716414000290

Second language sensitivity to agreement errors: Evidence from eye movements during comprehension and translation

JUNG HYUN LIM
Konkuk University

KIEL CHRISTIANSON
University of Illinois at Urbana–Champaign

Received: May 26, 2012 Accepted for publication: March 22, 2014

ADDRESS FOR CORRESPONDENCE
Jung Hyun Lim, Konkuk University, 1 Hwayang-dong, Gwangjin-gu, Seoul, Korea. E-mail: junghlim@konkuk.ac.kr

ABSTRACT
The present study addresses the questions of (a) whether Korean learners of English show sensitivity to subject–verb agreement violations in an eye-tracking paradigm, and (b) how reading goals (reading for comprehension vs. translation) and second language (L2) proficiency modulate depth of morphological agreement processing. Thirty-six Korean speakers of L2 English and 32 native English speakers read 40 stimulus sentences, half of which contained subject–verb agreement violations in English. The factors were whether a head and a local intervening noun matched in number and whether a sentence was grammatical or not. In linear mixed models analyses, both agreement violations and noun phrase match/mismatch were found to be disruptive in processing for native speakers at the critical regions (verb and following word), and locally distracting number-marked nouns yielded an asymmetric pattern depending on grammaticality. When L2 speakers were asked to produce offline oral translations of the English sentences into Korean, they became more sensitive to agreement violations. In addition, higher L2 proficiency predicted greater sensitivity to morphological violations. The results indicate that L2 speakers are not necessarily insensitive to morphological violations and that L2 proficiency and task modulate the depth of L2 morphological processing.

Adult second language (L2) learners' difficulty with inflectional morphemes, such as tense, agreement, and number, has been reported in a number of studies in the L2 acquisition literature (e.g., Ellis, 1988; Johnson, Shenkman, Newport, & Medin, 1996; Lardiere, 1998; Wei, 2001). Although there have been several approaches taken to explain the source(s) of the difficulty in L2 acquisition, a consensus has not yet been reached. In recent years, L2 researchers have questioned whether the difficulty stems from differences in grammatical processing between

© Cambridge University Press 2014 0142-7164/14 $15.00

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

the first language (L1) and L2 (Hopp, 2006; Jiang, 2004; Keating, 2009; Sato & Felser, 2007). This line of research also sheds light on mechanisms underlying L2 processing, exploring the overarching question in L2 sentence processing research of whether nonnative speakers are implicitly sensitive to morphosyntactic information. A growing body of research has started to examine whether adult L2 learners can attain a nativelike ability to detect morphological violations during on-line comprehension. Some studies have shown that L2 learners are less sensitive to agreement errors compared with native speakers (e.g., Jiang, 2004, 2007; Keating, 2009; Sato & Felser, 2007), whereas others show that nonnative processing is not qualitatively different from native processing (e.g., Foote, 2011; Hopp, 2006; Jackson, 2008; Sagarra & Herschensohn, 2010; Tokowicz & Warren, 2010).

While there are debates over whether L2 speakers can process their nonnative language in a nativelike fashion, recent studies indicate that L2 proficiency (Foote, 2011; Frenck-Mestre, 2002; Hopp, 2006; Jackson, 2008) and task type (Jackson & Bobb, 2009; Leeser, Brandl, & Whiteglass, 2001; Lim & Christianson, 2013b; Williams, 2006) influence L2 processing patterns. The present study aims to add to this line of research, examining whether adult L2 Korean learners of English with varying L2 proficiency are sensitive to subject–verb number agreement in English in two different tasks. The sentences under investigation here include both *seeming* and *real agreement violations*, such as "The teachers who instructed the student were very strict" (seeming agreement violation*), or* *"The teacher who instructed the student were very strict" (real agreement violation). Another sentence type involves the so-called *agreement attraction error*, in which a local or intervening noun interferes with agreement between the head noun and the main verb: *"The teacher who instructed the students were very strict." In this paper, we argue that L2 learners are able to show sensitivity to both seeming and real agreement violations and that the sensitivity differs as a function of task type and L2 proficiency. For sentences involving agreement attraction errors, we predict that L2 speakers should be relatively insensitive to agreement errors, given that native speakers often produce errors like these (e.g., Bock & Eberhard, 1993; Bock & Miller, 1991; Bock, Nicol, & Cutting, 1999), and fail to notice them in comprehension (Pearlmutter, Garnsey, & Bock, 1999; Staub, 2009; Wagers, Lau, & Phillips, 2009). Two different tasks were also compared here, namely reading for comprehension versus reading for translation, while readers' eye movements were recorded. In the reading for translation task, participants generated Korean translations of each English sentence they had just read. As several previous studies have suggested (Jackson & Bobb, 2009; Leeser et al., 2001; Lim & Christianson, 2013b; Williams, 2006), L2 learners' sensitivity to morphological features in the L2 should vary depending on the task they undertake. The translation task is predicted to focus attention on morphology, and this heightened attention should be apparent both in the on-line eye movement patterns and in off-line translations.

MORPHOLOGICAL PROCESSING IN L1 AND L2 STUDIES

L1 studies

Studies on L2 morphological processing are largely informed by research on L1 speakers. A well-documented area within L1 production is subject–verb number

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

agreement in English (e.g., Bock & Eberhard,1993; Bock & Miller, 1991; Bock et al., 1999). These studies have shown that native speakers are more likely to produce subject–verb agreement errors when the subject (first) noun is singular and the intervening (second) noun is plural. This phenomenon is termed the *broken agreement effect* or *agreement attraction* (Bock & Miller, 1991). The plurality of the intervening noun "attracts" the verb agreement erroneously. For example, completion of a sentence fragment like (1b) produced more subject–verb agreement errors than in (1a) (Bock & Miller, 1991).

(1) a. The key to the cabinet *are* . . .
b. The key to the cabinets *are* . . .

This production effect has also been observed in reading comprehension (Nicol, Forster, & Veres, 1997; Pearlmutter et al., 1999; Staub, 2009; Wagers et al., 2009). Pearlmutter et al. (1999) examined the extent to which the systems underlying production and comprehension are related. Word-by-word self-paced reading and eye tracking were used to explore the time course of agreement processing, measuring reading times on the critical regions, the verb, and the word following the verb, as in sentences like (2).

(2) a. The key to the cabinet was rusty from many years of disuse.
b. The key to the cabinets was rusty from many years of disuse.
c. *The key to the cabinet were rusty from many years of disuse.
d. *The key to the cabinets were rusty from many years of disuse.

Grammaticality and the number mismatch between the head noun and local noun were manipulated in the stimuli. The results were consistent with previous production studies: comprehenders took longer to read (2b) than (2a), showing evidence of the broken agreement effect in normal reading comprehension. This finding suggested automatic activation of morphological knowledge during online language processing by native speakers without subsidiary tasks, such as grammaticality judgment or lexical decision, given that the agreement marking in English is meager and not critical to understanding sentences. An interaction between the grammaticality and head/local noun number match was also found, indicating that the distracting local noun created more difficulty in processing grammatical sentences, such as (2b), while it eased processing difficulty in ungrammatical sentences, such as (2d). In addition, the finding that the same sensitivity to the agreement violations was found in normal comprehension was taken as evidence that mechanisms for processing agreement information involved in comprehension and production are closely related. It should be noted, however, that Wagers et al. (2009) showed that the attraction effects were observed only in ungrammatical sentences in a self-paced reading task. It therefore seems that there are still conflicting findings in the L1 agreement comprehension studies, and thus there remain questions regarding the nature of the attraction effects in L1 comprehension and production.

L2 studies

Based on the L1 findings, L2 researchers have started to investigate whether morphological knowledge is an integrated part of linguistic competence in bilinguals

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

and language learners. Using time-sensitive methodologies, such as self-paced reading and eye tracking, studies in L2 morphological processing have started to examine the cause of the consistent morphological difficulty reported in L2 acquisition. In particular, the inquiry has focused on whether this difficulty is attributable to an incomplete underlying grammar or to L2 processing mechanisms. Thus far, studies have provided mixed results regarding L2 learner sensitivity to L1 morphosyntactic information during online processing.

Jiang (2004) investigated whether inflectional morphology is an integral part of L2 competence, specifically testing whether the broken agreement effect found in native speakers is also present in L2 speakers. Using a word-by-word self-paced reading task, Chinese learners of English were asked to read sentences as in (3) to examine whether L2 learners are sensitive to number agreement morphology.

(3) a. The key to the cabinet was rusty from many years of disuse.
b. The key to the cabinets was rusty from many years of disuse.

L2 learners were not found to show any sensitivity to the number agreement morphology, unlike native speakers, who showed longer reading times for sentences like (3b), which had number disagreement between the local noun and the main verb (*be*). The study further tested L2 learners' sensitivity to different types of morphology as in the sentences below, namely subject–verb number agreement (4a,b), pronoun-verb disagreement (4c,d), and subcategorization violation (4e,f), in both grammatical and ungrammatical conditions.

(4) a. The bridges to the island were about ten miles away.
b. *The bridge to the island were about ten miles away.
c. I told you I am a professor of psychology.
d. *I told you she am a professor of psychology.
e. The teacher encouraged the children to mail the letter to the president.
f. *The teacher insisted the children to mail the letter to the president.

Results showed that while native speakers responded with significantly different reading times in all three types of violations, nonnative speakers were sensitive to pronoun–verb disagreement and subcategorization violations but not to subject–verb number agreement violations. The study concluded that L2 learners are less likely to integrate morphological knowledge while reading for comprehension.

Similar results were found by Sato and Felser (2007), who demonstrated that nonnative speakers showed reduced sensitivity to agreement as compared with case errors, unlike the native control group. The study investigated the extent to which L2 learners of English (L1 German, Japanese, and Chinese) were affected by different types of morphosyntactic violations, using both an online speeded grammaticality judgment task and an offline sentence completion task. Results revealed that nonnative speakers were significantly less sensitive to agreement violations than case errors regardless of L1 background, unlike native speakers, who were sensitive in detecting both case and agreement errors. The authors interpreted the results as showing that learners' reduced sensitivity to agreement errors may reflect their inability to use full-fledged syntactic structure during online L2 processing, given that subject–verb agreement requires computing a functional head (IP) with nonadjacent constituents, whereas case morphology is

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

local within the verb phrase. (See also Keating, 2009, for data on L2 Spanish processing of grammatical gender marking.)

Other research contradicts the above conclusions and indicates that L2 learners are sensitive to morphosyntactic violations and suggests possible other factors that may modulate this sensitivity (Coughlin & Tremblay, 2012; Foote, 2011; Hopp, 2006; Jackson, 2008; Sagarra & Herschensohn 2010; Tokowicz & Warren, 2010). For example, Hopp (2006) examined whether advanced and near-native L2 speakers of German would be able to use syntactic disambiguation information (case-marking vs. verb agreement) when processing subject–object ambiguities in the same way as do native German speakers. The study found that near-native L2 speakers displayed nativelike online reading time patterns, consistently showing local slowdowns in the critical region of object–subject (OS) sentences in both types of syntactic disambiguation conditions. The advanced L2 speakers, however, were generally slow in reading the OS order only at the end of the sentence, regardless of syntactic disambiguation type. These findings revealed the role of proficiency in L2 learner morphosyntactic processing, suggesting that sensitivity to syntactic information is not necessarily restricted in nonnative processing (see also Jackson, 2008).

Foote (2011) recently investigated whether early and late Spanish–English bilinguals possess integrated knowledge of agreement in Spanish using a self-paced reading task. Three types of agreement were tested: subject–verb, person and number agreement, and noun–adjective gender agreement. Results showed that both early and late bilinguals displayed equal sensitivity to all types of agreement in Spanish, although there was diminished sensitivity in subject–verb number agreement when intervening material existed between subjects and verbs (cf. the intervening material effect observed by Keating, 2009, in L2 Spanish gender agreement).

Sagarra and Herchensohn (2010) also examined gender and number agreement in Spanish with beginning and intermediate English-speaking learners of Spanish and found that L2 proficiency and working memory play a role in learners' processing. In a self-paced reading task and a grammaticality judgment task, only intermediate learners, not beginning learners, exhibited sensitivity to both gender and number agreement errors equivalent to native Spanish speakers. In addition, learners with higher working memory displayed higher sensitivity to gender agreement violations, indicating that gender agreement requires more cognitive resources. The results suggested that L2 learners can develop nativelike online computation of grammatical knowledge as proficiency increases, and differences between native and nonnative processing are quantitative, not qualitative (see also Coughlin & Tremblay, 2012, for working memory and proficiency effects in L2 French number agreement).

In summary, the extent to which L2 learners deploy knowledge of agreement during sentence comprehension is not yet clear. Some studies suggest that L2 learners have reduced sensitivity to agreement errors compared to native speakers (Jiang, 2004, 2007; Keating, 2009; Sato & Felser, 2007), but others provide evidence that even late learners show implicit knowledge of L2 agreement (Foote, 2011; Hopp, 2006; Jackson, 2008), perhaps modulated by L2 proficiency (Coughlin & Tremblay, 2012; Hopp, 2006; Jackson, 2008; Sagarra & Herchensohn, 2010). It may then be possible that L2 speakers can eventually become nativelike

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

in on-line morphological processing as L2 proficiency develops. Furthermore, secondary tasks might play a role in affecting the depth of L2 morphological processing (Jackson & Bobb, 2009; Leeser et al., 2001; Lim & Christianson, 2013b; Williams, 2006), as task has been shown to affect depth of L1 processing (Swets, Desmet, Clifton, & Ferreira, 2008).

TASK EFFECTS IN L2 PROCESSING AND GOOD-ENOUGH LANGUAGE PROCESSING

Recently, some L2 processing studies have reported that the relative depth of L2 speakers' on-line processing may be influenced by different tasks, such as the use of explicit grammaticality judgments versus comprehension questions (Jackson & Bobb, 2009; Leeser et al., 2011), the use of plausibility judgment tasks versus memory tasks (Williams, 2006), the type or sheer number of comprehension questions (Havik, Robert, van Hout, Schreuder, & Haverkort, 2009), or reading for comprehension versus reading for translation (Lim & Christianson, 2013b; Macizo & Bajo, 2004, 2006). These studies suggest that, if the task is designed to highlight a particular grammatical structure, L2 learners are able to allocate greater attention to the relevant grammatical structures, accordingly yielding longer reading times and deeper processing.

Similarly, the L1 literature also reports that reading goals and task demands can affect the depth of language processing. Swets et al. (2008; see also Christianson & Luke, 2011) showed that the depth of processing and final interpretation in syntactic ambiguity resolution can be influenced by the frequency and type of comprehension questions that are asked. The study found that when people were specifically asked about the attachment of globally ambiguous modifier phrases, reading times were longer than when they read the sentences with superficial questions. Swets et al. concluded that language processing can sometimes be good enough (Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira, Bailey, & Ferraro, 2002; Ferreira & Patson, 2007), such that comprehenders compute linguistic input using both syntactic and semantic/heuristic processing, but the outputs of these routes are not always reconciled if one output appears to be good enough for comprehension. In other words, even native speakers appear not to automatically process all available morphosyntactic information if present task demands do not require doing so (Ferreira & Patson, 2007).

Given these observations, the present study examines task effects on L2 morphological processing, using a translation paradigm, to be specific, whether a task that draws the attention of even lower proficiency L2 learners to morphosyntax will induce online behaviors and/or offline comprehension that more closely approximate native speaker behaviors and/or comprehension. Translation, a process that L2 speakers regularly engage in (either explicitly or implicitly), is assumed to require additional attention to linguistic input and thus increases cognitive load, as one needs to establish semantic and syntactic matches between two languages during reading text (Danks & Griffin, 1997; Gerver, 1976). In addition, the end result of translation, the translation itself, contains explicit evidence of how morphosyntax has been understood and how successfully people decode one language and encode into the other (Campbell, 1998). In other words, if comprehension of

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

the given L2 input morphosyntax fails, its translation into the L1 counterpart will not be accurate. This is especially true in cases where a given feature of the source (input) language is either optional or infrequent in the target (output) language. In such cases, if the given L2 morphology input is accurately translated into L1 output, it is reasonable to infer that the morphology in the L2 input was recognized, processed, and comprehended, as the chances are small that it was accidentally included in the output, free of influence from the L2 input (Lim, 2011; Lim & Christianson, 2013a, 2013b).

Such is precisely the case in the experiment reported here, in which plural subject marking is obligatory in the English input but optional in the Korean output. Because Korean does not have subject–verb number agreement, subject–verb number agreement in English may not be as important for Korean learners of English as it is for native English speakers to get the meaning of a sentence. Therefore, if L2 learners notice plurality in English while they are reading and produce plural markers in Korean translations of the input where these markers occurred in English, it would indicate that Korean-speaking learners of English are sensitive enough to respond to subject–verb agreement morphology. Below is the possible Korean translation (with and without plural marker *dul* in Korean) for an ungrammatical sentences.

(5) *haksayng-(dul) ul kaluchi-essten sensayng-un maywu emkeykha-yess-ta.*
student-PL-ACC instruct-RC-PST teacher-NOM very strict-PST-DE
(*The teacher who instructed the students were very strict.)

One could argue that translation is a purely off-line, strategic task. Several experimental studies present evidence contradicting this view, having shown that people initiate and carry out translation prior to the completion of processing the entire input, and thus the translation output reflects at least some processes that likely take place during on-line comprehension (Lim & Christianson, 2013a, 2013b; Macizo & Bajo, 2004, 2006). Macizo and Bajo (2004, 2006) reported that L2 learners' reading times in a translation task were much longer than in a repetition task, suggesting that translation requires a reformulation process between the two languages, which is the main idea underlying the *horizontal approach* in cognitive theories of translation. The horizontal approach (Danks & Griffin, 1997; Gerver, 1976) proposes that people who are engaged in translation experience increased cognitive load and allocate additional attention because they reformulate from one language to the other in a parallel manner rather than waiting until the input has been fully processed before beginning to translate. Even if readers produce the translation based on a "memory trace" of the input, readers must still process that input during reading, and L1 translations will contain the structure or morphology of the sentence. We do not propose that translation is the best measure of L2 processing; rather, we simply suggest that it can offer unique insights into task effects and the salience of the structural and semantic content of the translation input in L2 processing and comprehension. Translation is simply another way to probe the specific aspects of a given sentence under investigation, serving as an alternative to comprehension questions that may draw more explicit attention to the critical morphology (e.g., "Was there just one X?") or miss the critical aspect of the input sentence altogether (as in the case of Jiang's, 2004, study).

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

In the current study, by examining data from both reading for comprehension and reading for translation, we examine how eye-movement patterns differ based on reading goals in L2 sentence comprehension; to be more specific, whether Korean learners of English are able to detect violations of subject–verb agreement in English as a function of reading goals and L2 proficiency. This is an important issue in the ongoing debate on L2 morphological processing compared to native processing. Good enough theory, generalized for both L1 and L2 processing, posits that depth of processing can vary for a number of reasons, including cognitive load demands, processing tasks or goals, and in the case of L2 processing, proficiency. Accordingly, good enough theory predicts that changing tasks could shift less proficient L2 learners' attention towards morphosyntax, resulting in on-line and off-line indications of processing that more closely pattern with native speakers.

Based on previous translation studies (Lim, 2011; Lim & Christianson, 2013a; Macizo & Bajo, 2004, 2006; Danks & Griffin, 1997), the translation task used here was predicted to require participants to focus more on L2 linguistic input than the comprehension task. Together with other studies that report effects of a secondary task in L2 processing (Jackson & Bobb, 2009; Leeser et al., 2001; Lim & Christianson, 2013b; Williams, 2006), the increased attention to linguistic input was predicted to result in more sensitivity to morphosyntactic features of the L2 input and longer reading times in the translation session than in the comprehension session. Moreover, according to previous studies that report a crucial role of proficiency in L2 reading (Hopp, 2006; Jackson, 2008; Sagarra & Herchensohn, 2010), a proficiency effect was also expected to occur: L2 learners with lower proficiency should be less sensitive to morphological violations overall, but this insensitivity should be moderated in the reading for translation task compared to the reading for comprehension task.

METHOD

Participants

Participants were 37 Korean-speaking learners of English and 35 native speakers of English. All participants received $15 or a course credit in compensation for participation. English proficiency of the L2 speakers was measured using a cloze test (G. Dussias, Pennsylvania State University, with permission), which had a maximum possible score of 40.

Apparatus

An Eyelink 1000 eye tracker (SR Research Ltd., Toronto, Ontario, Canada), with spatial resolution of less than 0.05 degrees of visual arc and a sampling rate of 1000 Hz, was used to monitor participants' right eyes while reading sentences on a computer screen. Stimuli were presented in a 14-pt monotype font. Participants viewed the stimuli binocularly on a monitor 66 cm from their eyes; at this distance, approximately 2.5 characters equaled a 1 degree visual angle. The tracker was a tower mount, such that participants rested their foreheads on a forehead rest and their chins in a chin rest.

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Materials

There were 40 items, half of which on each of the four counterbalanced lists contained subject–verb agreement violations in English, to investigate sensitivity to subject–verb number agreement in native English speakers and L2 learners. Factors were whether the sentence was grammatical (G) or ungrammatical (UG), and whether a head and a local intervening noun matched in number (M) or not (No M), resulting in a 2 × 2 design. An example sentence in each condition is provided in (4).

(6) a. The *teachers* who instructed the *students were* very strict. (G, M)
b. The *teachers* who instructed the *student were* very strict. (G, No M)
c. *The *teacher* who instructed the *student were* very strict. (UG, M)
d. *The *teacher* who instructed the *students were* very strict. (UG, No M)

The critical regions were the matrix verb *were* and the following word *very*, following the design and analysis precedents established by Pearlmutter et al. (1999). The differences in materials between Pearlmutter et al. and the current study are as follows. First, grammatical conditions all contained plural head nouns in this experiment in order to keep the verb *were* in all conditions, whereas singular head nouns were used by Pearlmutter et al. Accordingly, the seeming agreement violation in this study is quite different from the sentence type used in the previous studies. The seeming agreement violation here includes a plural head noun and singular local noun as in (6b), which has not normally elicited large agreement attraction effects in the L1 production and comprehension literature (Bock & Miller, 1991; Eberhard, Cutting, & Bock, 2005). Second, head and local nouns appeared in subject relative clauses in this study, whereas head and local nouns were embedded in subject NPs consisting of NP1–preposition–NP2, such as "the key to the cabinet was rusty from . . . " in Pearlmutter et al.

As there are several differences in test sentences from previous processing studies, we predict some of the results to be somewhat different from the previous findings and to be novel in several aspects. First, we predict that native speakers should show effects for both real and seeming ungrammaticality consistent with previous studies (Pearlmutter et al., 1999; Wagers et al., 2009); however, the seeming agreement effect might not be strong because a singular local noun has been reported to elicit a less attraction effect. Second, native speakers are predicted to yield a so-called broken agreement effect to show processing disruptions in response to grammaticality and noun phrase (NP) match/mismatch. Third, L2 speakers are expected to display different sensitivity to grammaticality and NP-match/mismatch as a function of task and L2 proficiency. We predict more sensitivity to grammaticality in both reading sessions as proficiency increases. In addition, we expect the L2 speakers to show more nativelike processing patterns in detecting real and seeming agreement violations in the translation session compared to the comprehension session. We also note that the seeming agreement violation test sentence used in the present study deviates from stimuli used in the previous literature. As such, finding an effect in response to this type of sentence would be a novel finding regardless of the results, whether L2 speakers show sensitivity to this type of sentence.

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Relative clauses were chosen rather than prepositional phrase constructions for two reasons. First, it has been reported that native speakers make fewer agreement attraction errors with intervening nouns in relative clauses than in prepositional phrases (Bock & Cutting, 1992; Solomon & Pearlmutter, 2004). As such, this construction would be a conservative test of L2 processing: if L2 speakers display sensitivity to agreement violations in relative clauses, it would be strong evidence that even relatively less noticeable morphology was being processed. Second, the off-line dependent variable here is an L2-English to L1-Korean translation. Korean contains postpositions, which, unlike English prepositions, are bound morphemes with certain cliticlike properties that are substantially different from anything in English. Furthermore, English preposition use can be somewhat idiosyncratic and regionally determined. (e.g., Is it "the key *to* the cabinet" or "the key *for* the cabinet"?) We were therefore concerned that translating English prepositions into Korean postpositional clitics might be relatively more difficult than translating relative clauses. In contrast, a previous study using Korean–English translation (Lim & Christianson, 2013b) found reliably consistent translation of relative clauses (and passives, cf. Lim & Christianson, 2013a), despite word order differences between the two languages.

Sentences were distributed across four lists in a Latin square design such that each participant saw every item in only one of the four experimental conditions. Each list contained 84 filler sentences with various structures that were not relevant to the target structures, such as active/passive or locative structures. Two-thirds of the fillers were grammatical and the rest were ungrammatical sentences. The number of ungrammatical fillers was less than that of grammatical fillers due to the concern that participants might notice the ungrammaticality of the sentences, resulting in conscious attention to grammaticality during the experiment. The ungrammatical filler sentences included past tense errors and incorrect use of colloquial phrases. None of these ungrammatical filler sentences generated any problems in translations for L2 participants. We looked at the translations of ungrammatical fillers to find out whether asking people to translate ungrammatical sentences would be strange and thus affect their eye movements and reading strategies. Participants translated most of the ungrammatical fillers with appropriate meanings and concepts; L2 participants did not seem to be bothered by translating the ungrammatical sentences.

Procedure

The procedures for native speakers and L2 learners were slightly different. Native speakers read a total of 124 sentences, including 40 stimuli and 84 fillers, at their own pace while their eye movements were being monitored. Following each sentence, participants were asked to answer a comprehension question via a button press. The comprehension question followed both grammatical sentences and ungrammatical sentences, asking about the general content of the relative clause such as "Did the teacher instruct the students?" The rationale for asking a comprehension question after every sentence, regardless of grammaticality, was to distract participants from focusing their attention on grammaticality, instead

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

focusing more on normal comprehension. Participants were given a short break after the first 62 trials.

For the L2 learners, there were two reading sessions, comprehension and translation. Each list contained a total of 124 sentences and was divided into two blocks of 62 sentences (including 20 items and 42 fillers) for each session. The comprehension session was the same as for the native speakers: participants read each sentence on a computer screen at their own pace while their eyes were being monitored. After reading each sentence, they were asked to answer a comprehension question. In the translation session, participants were asked to read each sentence (including filler items) just as they did in the comprehension session. After they finished reading the sentence, they pressed a button to the next screen, which said "Translate!" instead of a comprehension question. Then they translated it into Korean verbally. Participants were not told to reflect any particular aspect of grammaticality in the L2; rather, they were asked to translate the English sentences as naturally and normally as possible in Korean. Because the tracker was a tower mount, the L2 participants were asked to say the translations, lifting their chins from the chin rest as minimally as possible, while their foreheads remained on the forehead rest. Most of the participants were successful at speaking in this way without moving their heads noticeably, but in instances when eye calibration was disturbed by translating, participants were recalibrated before continuing. The translation task usually took about 15–30 s for each sentence, and most of the participants finished the translation session in about 30–35 min, without any outlier participants. Translations were recorded and transcribed later for data analysis. The order of comprehension and translation sessions was counterbalanced such that half of the participants did the comprehension task first and the other half did the translation task first.[1] For both the native and L2 group, a short practice session preceded the main task to allow participants to familiarize themselves with the task.

L2 speakers were also given a cloze test after the eye-tracking session (provided with consent by P. Dussias, Pennsylvania State University; cf. Kim & Christianson, 2012; Lim & Christianson, 2013a, 2013b). The test consisted of a short narrative story with every 10th word replaced by a blank. Participants were asked to fill in the blanks as they read the story. The blanks could be filled with any part of speech. The maximum possible score was 40.

Data analysis

Dependent variables were reading times at two critical regions: the verb (*were*) and the following word (*very*). All reading times were rounded to the nearest millisecond in the data tables for easier reading. The reading time data were analyzed using linear mixed effects models (Baayen, Davidson, & Bates, 2008) instead of analysis of variance, which allow for simultaneous consideration of all factors in structuring experimental data, such as grammaticality, NP-match/mismatch, and proficiency. Linear mixed effects models also allow subject and item to be included as random variables without having to run separate by-participants and by-items analyses (Baayen et al., 2008). The translation data were coded into two categories to see whether plural markers were used in the Korean

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

translations. These categorical data were analyzed using a logit mixed model (Jaeger, 2008). For all models, random effect structure was fitted using likelihood ratio tests, while the fixed effect structure was fitted using a stepwise model selection procedure, with only predictors and interactions that were significant ($p < .05$) or marginal ($p < .10$) retained in the model and reported below. All final models reported below had random intercepts and slopes for participants and items. The p values for fixed effects were obtained using Markov-chain Monte Carlo sampling of random-intercept-only models, as removal of random slopes did not change the significance or pattern of any results.

RESULTS

Reading time data

Various reading time measures at the critical regions where participants should show sensitivity to grammaticality and NP match/mismatch were examined (Pearlmutter et al., 1999; Warren, McConnell, & Rayner, 2008). Five related measures of reading times are reported here: *first fixation duration* (the duration of the first fixation that is within an interest area during first pass reading), *gaze duration* (the summation of the time spent fixating an interest region from the first entrance into the region until leaving it in either direction during first pass reading), *go-past time* (also called *regression path duration*; the summed fixation duration from when the current interest area is first fixated until the eyes enter a forward interest area, including reading times on previous material that is regressed into), *total reading time* (the summation of the duration across all fixations on an interest region), and *regression out* (the proportion of trials on which there was a regression made from the current interest area to earlier regions prior to leaving that interest area in a forward direction, considering first-pass regressions only). The first two measures indicate early processing, and go-past is a combination of early and later processing (Warren et al., 2008). Regressions out were also examined because this measure is considered to reflect processing difficulty (Pearlmutter et al., 1999).

Native speakers. Comprehension accuracy was used as a cutoff to exclude data from participants who did not properly engage in comprehension. Three participants whose accuracy performance was below 85% were excluded from the reading time data analysis. For each of the reading time variables separately, fixations shorter than 80 ms and greater than 3 *SD* from the mean were eliminated, excluding about 3% of the data. All reading times were log-transformed and centered. Subject and Item were included as random variables. Regression out was analyzed using a logit mixed model.

VERB (WERE). Table 1 presents the mean fixation times for the five reading measures (first fixation, gaze duration, go-past, total time, and regression out) in the four conditions. The results of the linear mixed model for each reading time variable and the results of regression out in the logit mixed model are presented in Table 2.

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Table 1. *Mean (standard deviation) fixation times at "were" for native speakers*

Conditions	First Fix (ms)	Gaze (ms)	Go-Past (ms)	Total (ms)	Reg. Out (%)
G M	217 (71)	242 (112)	329 (440)	346 (217)	9
G No M	225 (76)	247 (110)	294 (230)	365 (230)	8
UG M	228 (85)	265 (138)	398 (572)	423 (263)	11
UG No M	234 (77)	265 (128)	421 (573)	400 (243)	13

Note: First Fix, First fixation duration; Reg. Out, regression out; G, grammatical; M, noun phrase (NP) match; No M, NP mismatch; UG, ungrammatical.

Table 2. *Fixed effects in the linear mixed model of reading times and logit mixed model of regression out at "were" for native speakers*

Linear Mixed Model of Reading Times				
Predictor	Coefficient	*SE*	*t*	*p*
Gaze duration ($R^2 = .32$)				
Intercept	2.35	0.02	124.80	<.01
G	0.01	0.01	1.92	.05†
Go-past ($R^2 = .30$)				
Intercept	2.41	0.02	89.03	<.01
G	0.02	0.01	2.96	<.01*
Total time ($R^2 = .28$)				
Intercept	2.49	0.02	112.94	<.01
G	0.02	0.01	4.19	<.01*
G × NP	−0.01	0.01	−2.02	.04*
Logit Mixed Model of Regression Out				
Predictor	Coefficient	*SE*	Wald *Z*	*p*
Regression out ($C = 1.40$)				
Intercept	−2.59	0.22	−11.73	<.01
G	0.23	0.12	1.88	.06†

Note: G, Grammaticality; NP, noun phrase match/mismatch.
†$p < .10$. *$p < .05$.

At the verb, first fixation durations showed numerically longer reading times in the ungrammatical conditions, but it did not approach significance. Gaze durations showed an effect of grammaticality ($t = 1.92$, $p = .05$), indicating longer reading times in the ungrammatical conditions. In go-past, the ungrammatical conditions also took longer to process than the grammatical conditions, resulting in a main effect of grammaticality ($t = 2.96$, $p < .01$). Total time also showed an effect of grammaticality ($t = 4.19$, $p < .01$). In addition, total time showed a significant interaction ($t = –2.02$, $p = .04$) of grammaticality and NP-match, indicating that ungrammaticality increased processing difficulty when two NPs

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Table 3. *Mean (standard deviation) fixation times at "very" for native speakers*

Conditions	First Fix (ms)	Gaze (ms)	Go-Past (ms)	Total (ms)	Reg. Out (%)
G M	244 (86)	275 (130)	327 (153)	366 (205)	13
G No M	233 (88)	277 (147)	371 (299)	380 (259)	14
UG M	232 (75)	271 (129)	443 (270)	408 (233)	30
UG No M	240 (84)	266 (117)	408 (356)	397 (223)	21

Note: First Fix, First fixation duration; Reg. Out, regression out; G, grammatical; M, noun phrase (NP) match; No M, NP mismatch; UG, ungrammatical.

matched in number compared to when they did not match, while grammaticality decreased processing load when two NPs matched compared to when they did not match. In addition, a marginal grammaticality effect ($z = 1.88$, $p = .06$) was observed in regressions out, indicating that the native speakers regressed more in the ungrammatical sentences than the grammatical ones as shown in Table 1.

WORD FOLLOWING THE VERB (VERY). The average fixation times for the same five reading measures are presented in Table 3. The results of the mixed-effects model for the word following the verb (*very*) and the results of logit mixed model are shown in Table 4.

First fixation duration showed a significant interaction ($t = 2.15$, $p = .03$), indicating that the NP-match/mismatch effect observed in the grammatical conditions was not there in the ungrammatical conditions. As seen in Table 3, reading times were longer in the NP-match condition than NP-mismatch in the grammatical sentences, but this difference was smaller and in the opposite direction in the ungrammatical conditions. Nothing was significant in gaze durations at the word *very*. In the analysis of go-past times, there were trials where no fixations were made past *very*, with 18.2% of the total data points in trials on 24 items in which *very* was the penultimate word. These trials were removed from the analyses for go-past times and regressions out because these trials do not reflect true go-past times or a first-pass regression proportion. After removing these trials, only grammaticality ($t = 2.68$, $p = .01$) reached significance. Total time yielded a reliable grammaticality effect ($t = 3.03$, $p < .01$) but no interaction. Analyses of regressions out yielded a reliable main effect of grammaticality ($z = 4.28$, $p < .01$), showing that regressions occurred more in ungrammatical conditions.

SUMMARY. In sum, native speakers showed sensitivity to ungrammaticality very early in processing, demonstrating that ungrammaticality was more disruptive to processing than the NP number (mis)match. The ungrammatical conditions took longer overall to read than the grammatical conditions and resulted in more regressions. Although it seems that native speakers detected ungrammaticality faster than the NP-factor, they were still influenced by whether the NPs matched or not, as signaled by a seeming agreement violation effect and broken agreement effect on total time at the verb. One aspect of the results that appears somewhat

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Table 4. *Fixed effects in the linear mixed model of reading times and logit mixed model of regression out at "very" for native speakers*

Linear Mixed Model of Reading Times				
Predictor	Coefficient	*SE*	*t*	*p*
First fixation duration ($R^2 = .23$)				
Intercept	2.35	0.01	207.72	<.01
G × NP	0.00	0.00	2.15	.03*
Gaze duration ($R^2 = .25$)				
Intercept	2.39	0.02	159.03	<.01
Go-past ($R^2 = .48$)				
Intercept	2.65	0.05	53.35	<.01
G	0.03	0.01	2.68	.01*
Total time ($R^2 = .19$)				
Intercept	2.52	0.02	152.90	.00
G	0.02	0.01	3.03	<.01*
Logit Mixed Model of Regression Out				
Predictor	Coefficient	*SE*	Wald *Z*	*p*
Regression out ($C = 2.05$)				
Intercept	−1.70	0.21	−8.07	<.01
G	0.46	0.11	4.28	<.01*

Note: G, Grammaticality; NP, noun phrase match/mismatch.
†$p < .10$. *$p < .05$.

inconsistent with some previous research (Wagers et al., 2009) was a tendency for the native speakers in this experiment to show an effect of grammaticality in both NP-match and NP-mismatch conditions. At this juncture, we can speculate that this inconsistency stems from a difference in methods: Wagers et al. (2009) used self-paced reading. Note that the match/mismatch effect observed here was in the later measure of total time. In self-paced reading, material cannot be reread, so any measure of rereading would naturally be lacking from their results. Another reason that we speculate is that our seeming ungrammatical sentences include singular local nouns, which have not normally attracted agreement errors in the previous L1 literature. Accordingly, our native speaker participants might have been able to show an effect of grammaticality regardless of NP-match/mismatch. Similarly, the interaction of grammaticality and NP-match at *very* in the very early measure of first fixation duration observed here could easily have been swamped in self-paced reading, given that gaze duration showed no effects here, and first-fixation and gaze duration cannot be dissociated in self-paced reading. Note, however, that the interaction in first fixation durations is in an unexpected direction if this measure on the word following the critical verb is interpreted as reflecting processing of the verb (i.e., as a true "spillover" region). To the extent that this interaction can be considered reliable, it appears to be more informative about the degree of uncertainty about previously encountered material than about processing of

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

either *very* or the preceding verb. Recall that first fixation durations generally reflect visual word identification processes (Rayner, 1998); however, in these sentences, the word *very* is extremely frequent and semantically uninformative, requiring little effort for identification. Levy, Bicknell, Slattery, and Rayner (2009) propose that readers "maintain uncertain beliefs" about the identity of previously read material and update these beliefs based on the grammatical consistency of incoming material. These changing beliefs then guide saccadic behavior, both forwards and backwards. The ungrammatical-match condition and grammatical-mismatch condition are precisely where grammaticality and NP number alignment conflict, and thus they would be expected to be the conditions in which the highest level of lingering uncertainty about what had been read would be experienced. The shorter first fixation durations in these conditions, then, appear to result from readers quickly launching leftward saccades and rereading previous material to try to alleviate this uncertainty. Note also that the two conditions yielded longer go-past times, an index of rereading, consistent with this explanation.

Otherwise, the pattern of results was largely consistent with the study of Pearlmutter et al. (1999), supporting the argument that subject–verb agreement is an integral component of comprehension for native speakers. The present data did differ from the data in Pearlmutter et al. in one way, however. Specifically, Pearlmutter et al.'s participants showed sensitivity at one word after the verb, whereas our participants displayed sensitivity at the verb itself. Potential reasons for this difference are addressed in the discussion.

L2 speakers. The same five reading time measures used for native speakers were analyzed for L2 speakers. All participants except one displayed comprehension accuracy above 90% correct. The data from this one participant, whose accuracy was 72%, was excluded from the data analyses. The same analyses were performed on the data for the remaining 36 L2 speakers as were performed for the native speaker data. Linear mixed-effects models were used to analyze the reading time data in each reading session, comprehension and translation. Logit mixed models were used to analyze regression out data.

For each of the reading time measures separately, fixations shorter than 80 ms and longer than 5 *SD* from the mean were removed, instead of 3 *SD* from the mean, resulting in less than 3% data loss. Using 3 *SD* caused too much data loss (more than 10%), due to the increased variability of L2 reading times compared to native speakers. All reading times were log-transformed and centered, as was done for native speakers. Cloze scores were centered as well. Subject and Item were included as random effects.

Comprehension session.

VERB (WERE). Table 5 presents the mean fixation times in the four conditions, and Table 6 presents the results of the mixed model for each reading time variable and the result of the logit mixed model for regression out.

At the verb in the comprehension session, the effects of grammaticality and NP-match did not approach significance in any measure. However, proficiency yielded marginally significant effects ($ps < .10$) in most of the reading times, suggesting

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Table 5. *Mean (standard deviation) fixation times at "were" for comprehension session for second language speakers*

Conditions	First Fix (ms)	Gaze (ms)	Go-Past (ms)	Total (ms)	Reg. Out (%)
G M	242 (72)	278 (101)	345 (262)	537 (349)	10
G No M	252 (77)	284 (110)	382 (379)	532 (360)	9
UG M	245 (77)	285 (119)	363 (315)	561 (354)	7
UG No M	234 (71)	271 (101)	347 (330)	494 (307)	8

Note: First Fix, First fixation duration; Reg. Out, regression out; G, grammatical; M, noun phrase (NP) match; No M, NP mismatch; UG, ungrammatical.

Table 6. *Fixed effects in the linear mixed model of reading times and logit mixed model of regression out at "were" for comprehension session for second language speakers*

Linear Mixed Model of Reading Times				
Predictor	Coefficient	*SE*	*t*	*p*
First fixation duration ($R^2 = .19$)				
Intercept	2.36	0.01	238.59	<.01
Gaze duration ($R^2 = .23$)				
Intercept	2.41	0.02	187.42	<.01
Prof.	−0.01	0.00	−1.81	.07†
Go-past ($R^2 = .30$)				
Intercept	2.47	0.02	120.55	<.01
Prof.	−0.02	0.01	−2.38	.02*
NP × Prof.	−0.01	0.00	−1.74	.08†
Total time ($R^2 = .34$)				
Intercept	2.63	0.03	89.17	<.01
Prof.	−0.01	0.01	−1.71	.09†
Logit Mixed Model of Regression Out				
Predictor	Coefficient	*SE*	Wald *Z*	*p*
Intercept	−2.93	0.26	−11.41	<.01
Prof.	−0.19	0.09	−2.27	.02*

Note: Prof., Proficiency; NP, noun phrase match/mismatch.
†$p < .10$. *$p < .05$.

that the lower proficiency learners tended to read more slowly and regress more than the higher proficiency learners.

WORD FOLLOWING THE VERB (VERY). Mean fixation times are presented in Table 7. Table 8 presents the results of the mixed-model for each reading time variables at the word following the verb *very*, in the comprehension session.

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Table 7. *Mean (standard deviation) fixation times at "very" for comprehension session for second language speakers*

Conditions	First Fix (ms)	Gaze (ms)	Go-Past (ms)	Total (ms)	Reg. Out (%)
G M	228 (70)	234 (75)	352 (481)	410 (266)	11
G No M	229 (66)	256 (98)	331 (391)	404 (263)	10
UG M	225 (74)	241 (79)	398 (485)	439 (286)	17
UG No M	225 (70)	236 (85)	345 (369)	383 (240)	16

Note: First Fix, First fixation duration; Reg. Out, regression out; G, grammatical; M, noun phrase (NP) match; No M, NP mismatch; UG, ungrammatical.

Table 8. *Fixed effects in the linear mixed model of reading times and logit mixed model of regression out at "very" for comprehension session for second language speakers*

Linear Mixed Model of Reading Times				
Predictor	Coefficient	*SE*	*t*	*p*
First fixation duration ($R^2 = .18$)				
Intercept	2.34	0.01	244.77	<.01
Gaze duration ($R^2 = .20$)				
Intercept	2.36	0.01	218.99	<.01
G × NP	−0.01	0.01	−2.35	.02*
Go-past ($R^2 = .29$)				
Intercept	<.01	<.01	87.46	<.01
Prof.	<.01	<.01	1.91	.06†
Total time ($R^2 = .21$)				
Intercept	2.53	0.02	122.68	<.01
NP	−0.02	0.01	−1.70	.09†
G × Prof.	0.01	0.00	2.02	.04*
Logit Mixed Model of Regression Out				
Predictor	Coefficient	*SE*	Wald *Z*	*p*
Regression out ($C = 1.52$)				
Intercept	−2.66	0.35	−7.60	<.01
G	0.44	0.15	2.89	<.01*
Prof.	0.19	0.08	2.47	.01*

Note: G, Grammaticality; NP, noun phrase match/mismatch; Prof., proficiency.
†$p < .10$. *$p < .05$.

First fixation durations were similar in the four conditions; however, gaze duration showed different reading times depending on grammaticality and NP-match/mismatch, resulting in an interaction of grammaticality and NP-match/mismatch ($t = –2.35$, $p = .02$). The interaction revealed that the singular local noun induced processing difficulty in grammatical sentences but not in

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

the ungrammatical sentences, yielding longer reading times in the grammatical, NP-mismatch condition. In addition, this interaction also indicated that L2 learners showed an attraction effect for ungrammatical sentences, showing attenuated processing difficulty in the UG-No-M condition, mirroring the interaction found in total times at *were* in native speakers. In total time, proficiency interacted with grammaticality ($t = 2.02$, $p = .04$). This indicates that high-proficiency learners read the ungrammatical conditions more slowly than low-proficiency learners, suggesting that high-proficiency learners were more sensitive to ungrammaticality than were low-proficiency learners. L2 speakers regressed more in ungrammatical conditions ($z = 2.89$, $p < .01$), indicating that L2 speakers noticed ungrammaticalities. In addition, high-proficiency participants were more likely to regress ($z = 2.47$, $p = .01$).

SUMMARY. In the comprehension session, L2 learners showed sensitivity to grammaticality and NP-match/mismatch at the word following the verb in both an earlier measure (gaze duration) and a later measure (regressions out). In addition, the sensitivity varied according to proficiency: higher proficiency speakers showed more sensitivity to ungrammaticality in total time. However, it appears that the sensitivity to ungrammaticality in L2 learners was not as clear and consistent as in native speakers, since the effects were only shown in gaze duration, total time (only for higher proficiency group), and regressions out, whereas the effects of grammaticality were consistently found in most of the reading time measures in native speakers. Another difference between L1 and L2 speakers is that L2 learners did not show any significant reaction to either subject–verb agreement violations or NP-match/mismatch at the verb, as opposed to native speakers, who responded to ungrammaticality at the verb. This indicates that L2 processing may be slower than native speaker processing, that is, effects arose one word later in the sentence. Yet the finding that the higher proficiency L2 learners showed reliably longer reading times in the ungrammatical conditions at the word following the verb seems to suggest that they were also able to detect the morphological violations, similar to native speakers but not as quickly as the native speakers, and proficiency played a crucial role. This delayed sensitivity to the agreement morphology may reflect a quantitative difference between L1 and L2 speakers in processing speed or capacity (cf. McDonald, 2000, 2006). (Note, however, that in some previous research, e.g., Pearlmutter et al., 1999, even native speakers failed to show sensitivity to agreement errors until the word following the verb.)

Translation session.

VERB (WERE). Table 9 shows mean fixation times in the four conditions. The results of the mixed-effect model of reading time variables and the logit mixed model of regressions out at the verb in the translation session are presented in Table 10.

In most of the reading time measures, except regressions out, proficiency yielded a main effect, indicating that lower proficiency learners spent more time reading than did higher proficiency learners, just as in the comprehension session. An interaction ($t = -1.98$, $p = .05$) of grammaticality and NP-match/mismatch

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Table 9. *Mean (standard deviation) fixation times at "were" for translation session for second language speakers*

Conditions	First Fix (ms)	Gaze (ms)	Go-Past (ms)	Total (ms)	Reg. Out (%)
G M	230 (77)	272 (120)	424 (482)	1866 (1372)	15
G No M	238 (81)	259 (100)	384 (410)	1831 (1375)	12
UG M	244 (75)	271 (114)	399 (427)	1648 (1283)	13
UG No M	232 (79)	260 (102)	368 (321)	1755 (1339)	15

Note: First Fix, First fixation duration; Reg. Out, regression out; G, grammatical; M, noun phrase (NP) match; No M, NP mismatch; UG, ungrammatical.

Table 10. *Fixed effects in the linear mixed model of reading times and logit mixed model of regression out at "were" for translation session for second language speakers*

Linear Mixed Model of Reading Times				
Predictor	Coefficient	*SE*	*t*	*p*
First fixation duration ($R^2 = .14$)				
Intercept	2.34	0.01	268.83	<.01
Prof.	−0.01	0.00	−2.24	.03*
G × NP	−0.01	0.01	−1.98	.05†
Gaze duration ($R^2 = .18$)				
Intercept	2.39	0.01	210.34	<.01
Prof.	−0.01	0.00	−2.63	.01*
Go-past ($R^2 = .27$)				
Intercept	2.48	0.02	102.64	<.01
Prof.	−0.01	0.01	−1.68	.09†
Total time ($R^2 = .36$)				
Intercept	3.11	0.04	78.79	<.01
Prof.	−0.03	0.01	−2.38	.02*
Logit Mixed Model of Regression Out				
Predictor	Coefficient	*SE*	Wald *Z*	*p*
Regression out ($C = 1.48$)				
Intercept	−2.17	0.23	−9.32	<.01

Note: Prof., Proficiency; G, grammaticality; NP, noun phrase match/mismatch.
†$p < .10$. *$p < .05$.

in first fixation duration indicated longer reading times in the ungrammatical, NP-match condition than any other conditions. Other than this effect in first fixation duration, no other reading measure yielded significant effects of grammaticality or NP-match/mismatch at the verb in the translation session.

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Table 11. *Mean (standard deviation) fixation times at "very" for translation session for second language speakers*

Conditions	First Fix (ms)	Gaze (ms)	Go-Past (ms)	Total (ms)	Reg. Out (%)
G M	219 (77)	233 (90)	315 (314)	1460 (1155)	10
G No M	224 (76)	237 (95)	294 (237)	1708 (1359)	8
UG M	212 (69)	238 (102)	292 (312)	1881 (1630)	8
UG No M	224 (76)	241 (104)	287 (241)	1658 (1208)	9

Note: First Fix, First fixation duration; Reg. Out, regression out; G, grammatical; M, noun phrase (NP) match; No M, NP mismatch; UG, ungrammatical.

Table 12. *Fixed effects in the linear mixed model of reading times and logit mixed model of regression out at "very" for translation session for second language speakers*

Linear Mixed Model of Reading Times				
Predictor	Coefficient	*SE*	*t*	*p*
First fixation duration ($R^2 = .19$)				
Intercept	<.01	<.01	213.39	<.01
NP	<.01	<.01	1.81	.07†
Gaze duration ($R^2 = .20$)				
Intercept	2.34	0.01	185.70	<.01
Go-past ($R^2 = .22$)				
Intercept	2.39	0.02	131.32	<.01
Total time ($R^2 = .46$)				
Intercept	3.07	0.05	59.09	<.01
G	0.03	0.01	2.54	.01*
G × NP	−0.02	0.01	−1.88	.06†
Logit Mixed Model of Regression Out				
Predictor	Coefficient	*SE*	Wald *Z*	*p*
Regression out ($C = 1.36$)				
Intercept	−2.89	0.28	−10.38	<.01
NP × Prof.	0.12	0.06	1.99	.05†

Note: NP, Noun phrase match/mismatch; G, grammaticality; Prof., proficiency.
†$p < .10$. *$p < .05$.

WORD FOLLOWING THE VERB (VERY). Mean fixation times are shown in Table 11. Table 12 presents the results of the mixed model for each reading time variable at the word *very* in the translation session.

First fixation durations displayed a marginal effect of NP-match/mismatch ($t = 1.81$, $p = .07$), indicating a tendency for NP-mismatch conditions to yield longer reading times than NP-match conditions. Neither predictor was significant

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

in the analysis of gaze durations or go-past times. Total time showed a main effect of grammaticality ($t = 2.54$, $p = .01$) and a marginal interaction ($t = –1.88$, $p = .06$). As shown in Table 11, the effect of grammaticality shows that L2 learners spent more time reading the ungrammatical conditions than the grammatical conditions. Grammaticality marginally interacted with the NP-match/mismatch factor, yielding an asymmetrical grammaticality pattern. This pattern is noteworthy because it is reminiscent of the one observed at the verb in the native speakers' data. Regression out yielded an interaction of proficiency and NP-match/mismatch ($t = 1.99$, $p = .05$), showing that, whereas higher proficiency speakers regressed more in the NP-match conditions, lower proficiency learners regressed more in reading NP-mismatch conditions.

SUMMARY. In the translation session, L2 learners were sensitive to both grammaticality and NP-match/mismatch across the verb region, mainly at the word following the verb. A noticeable finding was that the L2 learners' total time pattern appearing at the word following the verb in the translation session was similar to that of native speakers at the verb itself, which is a unique result that was not observed in the L2 learners' comprehension session. This one-word delay implies that the L2 learners were slower to process the subject-agreement violations, and this processing spilled over into fixations on the following word (cf. McDonald, 2000, 2006), but they eventually noticed the real and seeming morphological violations.

Another point of interest in the translation session was that the L2 learners did not display significant effects in the late processing measure regressions out, unlike what was observed in the comprehension session. This pattern of effects can be attributed to the nature of the translation task, which requires more attention to linguistic input and deeper processing than reading for comprehension. Recall that translations were provided orally by participants after the English source sentences had been read and were no longer on the computer monitor. This means that participants could generate translations silently while reading the English source sentences and then repeat these translations later. Alternatively, they could read the English source sentences, memorize them verbatim, and then draw upon their memory of the sentences for the later translation. Given that memory for morphosyntactic content decays rapidly even in one's native language (Sachs, 1967), we find the former characterization of the translation process to be more plausible. Although we cannot rule out the latter characterization of the translation process in this experiment, we follow Danks et al. (1997) and Macizo and Bajo (2004, 2006) in assuming that participants at least began to formulate their translations while they were reading the source sentences and thus paid more attention to the input itself during reformulation process. In turn, it is possible that L2 learners had to go back to check or recover from disruptions caused by ungrammaticality less often when translating.

This speculation can also be related to the effect observed in total times in the translation session, summarized above. Although the pattern is similar for both native and L2 speakers, the reading times of L2 speakers were almost 4–5 times longer than those of native speakers, which seems a bit more extreme than "generally longer." Again, we attribute this result to the nature of translation itself. It seems that translation work occurs as people read the sentence; the inflated first

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

fixation durations (G × NP, $p = .05$) seem to suggest early devotion of resources to translation, and the sensitivity finally became clear in a late processing measure (i.e., total times). This task-based difference is consistent with previous translation research (Lim & Christianson, 2013b; Macizo & Bajo, 2004, 2006), suggesting that translation is at least begun as material is being encountered, rather than after all the input has been completely read.

Further analyses

Because there were two reading tasks for L2 learners, we considered task effects implicated in the above results. Thus, we ran an analysis to make sure the measures (especially, total times at *very*) that show an increase in agreement sensitivity in translation really increased over the comprehension task. The results revealed a marginal interaction of Grammaticality × Proficiency × Task ($t = –1.79$, $p = .07$), suggestive of increased sensitivity in the translation session modulated by proficiency. In addition, we ran another analysis to examine whether power might have been an issue, as there were only half as many items for the L2 group in each task block as there were for the native group, who had only one reading task. We conducted an analysis with 20 items in the L1 group compared to the same 20 items in the L2 comprehension group, comparing the two groups on all measures. There were interactions between grammaticality and speaker groups in go past at *were* ($t = –2.72$, $p < .01$) and in total times at *very* ($t = –1.83$, $p = .06$), indicating that the native speakers but not the L2 speakers still showed grammaticality effects when just 20 items were considered. However, since there is not similar variance between the native group and L2 group, we cautiously interpret this result to suggest that the lack of L2 results in the comprehension task were not due to lack of power.

Finally, we analyzed reading times on both NPs themselves. We found generally longer reading times in most of the reading measures for mismatching NPs than matching NPs for both native speakers and nonnative speakers in most of the reading measures (i.e., gaze durations, go-past times, and total times; all ts > 2.00, ps $< .05$). In addition, reading plural NPs took longer than singular NPs, yielding an effect of grammaticality for NP1 (e.g., "the *teachers* that instructed the students/student were . . . " vs. "the *teacher* that instructed the students/student were . . . "; i.e., gaze durations, go-past times, and total times; all ts $> –2.00$, ps $< .05$). The interaction of grammaticality and NP-match for NP2 also indicates that reading plural NPs took more time than singular NPs. All these patterns appeared in both native and nonnative groups. Yet, the regressions-in pattern varied between the two groups: native speakers regressed into NP1 more in the grammatical conditions than in the ungrammatical conditions ($z = –1.95$, $p = .05$), whereas for NP2, they returned more in the ungrammatical conditions ($z = 1.96$, $p = .04$). The effect of grammaticality was not found in the L2 comprehension session for either NP1 or NP2. In the translation session, there was an interaction of grammaticality and NP-match ($z = 2.35$, $p = .02$) for NP2, indicating more regressions in for NP-match in the grammatical conditions (e.g., the teachers that instructed the *students* were . . . > the teachers that instructed the *student* were . . .), but more regressions in for NP-mismatch in the ungrammatical conditions (e.g., the teacher that instructed the *student*

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Table 13. *Fixed effects in the logit mixed model of head and local noun translations for second language speakers*

Predictor	Coefficient	*SE*	Wald *Z*	*p*
Head Noun Translations ($C = 1.89$)				
Intercept	−1.91	0.34	−5.60	<.01
G	−1.12	0.13	−8.52	<.01*
G × NP	0.32	0.13	2.57	.01*
Local Noun Translations ($C = 1.45$)				
Intercept	−3.99	0.56	−7.08	<.01
G	−0.56	0.26	−2.11	.03*
G × NP	1.61	0.28	5.84	<.01*

Note: G, Grammaticality; NP, noun phrase match/mismatch.
†$p < .10$. *$p < .05$.

were ... < the teacher that instructed the *students* were ...). It seems that L2 participants were more likely to be influenced by the NP forms themselves than native speakers were. In addition, that these regressions-in patterns occurred only in the translation session (which had the same number of items as in the comprehension session) suggests that the translation task increased L2 participants' awareness of the linguistic forms, triggering deeper processing.

Translation data. Because off-line translation data offers an explicit measure of L2 comprehension, the question of how grammaticality and NP-match/mismatch affect the use of plural markers in Korean translations was of interest in the analysis of the translation data. Whether the use of plural markers in translations was related to the online reading time measures, such as the regressions made from the verb or the following word or go-past times, was also examined. The translation data were coded into two categories: whether plural markers were used or not for both head nouns and local nouns. There were no errors on the translations for head and local nouns in word meaning; in other words, all participants signaled in some way either the appropriate existence or absence of plurality for both head and local nouns while at the same time indicating that the meanings of the NPs were correct. A logit mixed model analysis was used because the variables were binary.

None of the online reading measures predicted the use of plural markers in translations. This includes late measures (go-past, regressions out) that might be interpreted as signaling some *post hoc* reprocessing in preparation for translation. Grammaticality, NP-match/mismatch, and proficiency did explain the use of plural markers in translations. Proficiency neither reached significance nor yielded interactions with either grammaticality or NP-match/mismatch factors. Table 13 presents the results of the logit mixed models for head and local noun translations. Figure 1 shows the proportion of plural markers in Korean translation in the four conditions for the head and local noun.

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

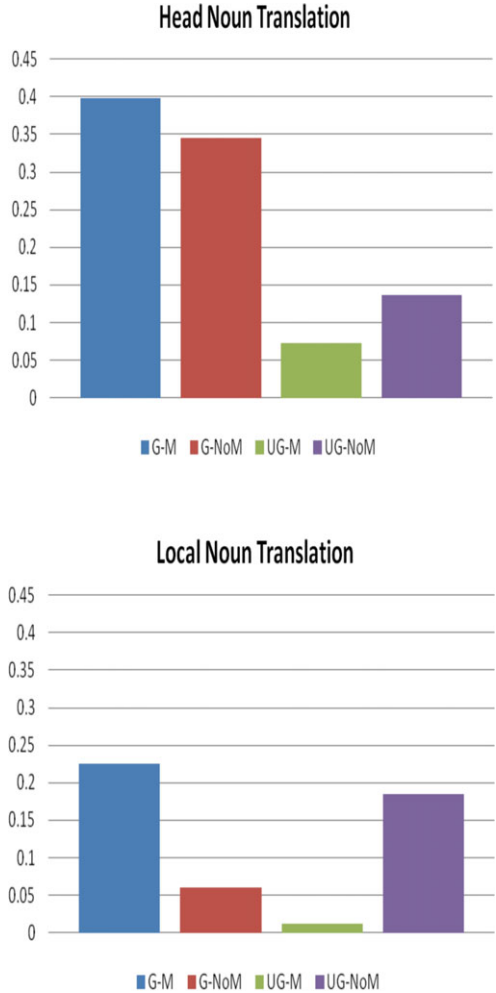


Figure 1. (Color online) Proportions of the numbers of plural markers for head and local noun in each condition.

With respect to the head noun, the use of plural markers in Korean translations was significantly influenced by grammaticality ($z = -8.52, p < .01$). In addition, the interaction of grammaticality and NP-match/mismatch was reliable ($z = 2.57, p = .01$). As shown in Figure 1, the effect of grammaticality indicates that the numbers of plural markers for head nouns (NP1) were lower in the ungrammatical conditions than in the grammatical conditions. The interaction of grammaticality and NP-match/mismatch represents an asymmetry effect of NP-mismatch, depending on grammaticality. In particular, the intervening singular local noun (NP2) in the grammatical condition (e.g., "the teachers who instructed the student were very

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

strict") reduced the plural marking of head nouns in translations, even though plural head noun translations should have occurred. In contrast, an intervening plural local noun caused more plural marking of head nouns in the ungrammatical conditions (e.g., "the teacher who instructed the students were very strict"), even when head nouns were singular in English. A similar result was found for local noun translations as well: there was a main effect of grammaticality ($z = –2.11$, $p = .03$), and a significant interaction ($z = 5.84$, $p < .01$). These results demonstrate that the L2 speakers had attended to the number morphology of the intervening NP, even though it was not appropriate to do so, just as native speakers have been shown to do.

In addition, an analysis of item order effects on plural use in translations was performed to see whether there was a training effect as the session progress. The result revealed no significant item order effect for either head or local nouns ($p > .10$); in other words, the use of plural nouns in translations did not increase as the session progressed.

In sum, the use of plural markers in the translations provides off-line evidence that the English plural morphology was processed, especially given that subject–verb number agreement is not obligatory in Korean production and the markings in the translations could have been left out entirely and the translations would still have been correct. It is important that the intervening local nouns affected how participants translated head nouns, depending on the grammaticality of the English sentences. Specifically, singular local nouns operated as an obstacle in producing plural markers of plural head nouns in the grammatical sentences. In contrast, when local nouns were plural in the ungrammatical sentences, people produced more plural markers for head nouns even though the original English sentence contained a singular head noun.

GENERAL DISCUSSION AND CONCLUSION

The current study used eye tracking to examine L2 speakers' sensitivity to subject–verb number agreement in online reading. The effects of L2 proficiency and reading goals on morphological processing were also investigated by having two reading sessions: comprehension and translation. For native speakers, both agreement violations and NP-match/mismatch were disruptive during on-line reading comprehension. Taken together, the native speaker data displayed typical *broken agreement effects* in comprehension, similar to that found in production (Bock & Miller, 1991; Pearlmutter et al., 1999). However, as mentioned above, there is also a different result between this experiment and Pearlmutter et al. (1999) that should be discussed: the position at which participants showed significant sensitivity to the violations in subject–verb number agreement. Pearlmutter et al.'s participants showed processing disruption at the word following the verb. In the present experiment, however, native speakers encountered processing disruption at the verb itself. This difference seems attributable to the different types of syntactic structures used. The distance from head noun to the matrix verb was one-word longer in this experiment than in Pearlmutter et al.'s materials, since this experiment used relative clauses whereas Pearlmutter et al. used prepositional phrases, as in "The keys to the cabinets were rusty. . . ." If not all processing is completed prior to

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

the eyes moving off a given word (e.g., Slattery, Stuart, Christianson, Yoshida, & Ferreira, 2013), then the extra word between the head noun and the verb in the present materials could have allowed for complete morphological computation and differentiation between the subject and intervening noun, such that when the eyes got to the verb in the present materials, the processor was better prepared to instantiate agreement. This speculation should be tested in future work. In addition, because sentences in the present study included plural head nouns (subjects) and singular local nouns (attractors), native speakers were not really disturbed by this singular attractor in processing agreement, such that they were more able to focus on grammaticality, yielding more rapid signs of sensitivity to grammaticality regardless of the NP-factor in this study compared to Pearlmutter et al. (1999).

L2 learners showed sensitivity to subject–verb number agreement violations and NP-match/mismatch across the verb regions as well. In both comprehension and translation sessions, they displayed processing disruption in response to both grammaticality and locally distracting nouns, similar to native speakers. The result is inconsistent with the findings by Jiang (2004), who observed that L2 speakers did not display sensitivity to subject–verb number agreement and disagreement between head and local nouns in PP modifiers. That result was interpreted in support of the argument that L2 speakers' morphological difficulties lie in incomplete L2 competence rather than a performance deficiency. In contrast, the current experiment showed that L2 learners were able to show sensitivity to both grammaticality and locally distracting nouns. The difference between our results and Jiang's may be attributable to the difference in methodology (i.e., eye tracking vs. self-paced reading). Considering that self-paced reading cannot provide a multifaceted record of readers' processing, whereas eye tracking can (Frenck-Mestre, 2005), it is likely that the self-paced reading task used by Jiang was simply not sensitive enough to provide reliable data on morphological processing. Although a local plural embedded in a relative clause has been reported to be a weaker attractor than a local plural in a prepositional phrase modifier (Wagers et al., 2009), it was apparently sufficient to trigger the effects observed here. It is also possible that various structures have differential effects for L1 and L2 speakers, perhaps due to the frequency in which they are encountered in L1 versus L2 reading. Future studies similar in design to the present one using a variety of structures and methods are needed to satisfactorily address this discrepancy.

There were three noteworthy differences between native speakers and L2 speakers. First, the pattern that became apparent at the verb itself in the native speaker data did not appear until the word following the verb in the L2 speaker data. The one-word delay in effects of grammaticality and NP-match/mismatch suggests that the L2 learners' processing was slower than native speakers (cf. McDonalds, 2000, 2006). Second, L2 speakers failed to display sensitivity to grammaticality during early processing (first fixation duration and gaze duration), whereas native speakers showed a grammaticality effect in these early reading time measures. Third, it appears that the nonnative speakers were more disturbed by the NP-factor relative to grammaticality, showing sensitivity in early processing measures to seeming agreement violations in both comprehension and translation sessions. In contrast, native speakers mostly showed sensitivity to grammaticality regardless of the NP-factor in early processing measures, although first fixation durations did

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

reflect lingering uncertainty introduced by this conflict when grammaticality and NP number alignment conflicted.

It is not surprising that L2 speakers should be disrupted earlier by conflicting grammaticality (syntactic) and NP number alignment (heuristic) cues, but does it signal a qualitative difference in processing? There are two general findings reported here to keep in mind: (1) effects tend to arise later in the L2 eye movement record (i.e., on *very* rather than on *were*) and (2) it was in the translation session where the L2 speaker data patterned in a more nativelike way. If we compare Table 1 (native speakers, *were*) to Table 11 (L2 speakers, *very*, translation session), we notice that native speakers' total times (Table 1) reveal disruption for both "seeming ungrammaticality" (G-No M) and "seeming grammaticality" (UG-M), resulting in a pattern of data, from top to bottom, of low–high–high–low. This interaction was also found to be significant (Table 2). Looking at the corresponding column in Table 11, we see the same low–high–high–low pattern, and in Table 12, we see that the interaction is marginally significant ($p < .06$). Given that the influence of the NP number alignment influences L2 processing in earlier measures, it could be argued that this reliance on a heuristic cue represents a qualitative difference in processing between the two groups; however, the influence of NP number alignment is also seen in the native speaker data, and the pattern in the two data sets is comparable when one keeps in mind that the L2 effects tend to look more nativelike one word later and in the translation session. These differences in timing and task-based effects obfuscate comparisons between native and L2 processing, which is perhaps part of the source of the ongoing debate about how to characterize these differences. Taken together, we interpret the data reported here as showing that the L2 learners still eventually computed the morphologically marked relations very similarly to the native speakers when the task focused attention on the morphosyntax. Thus, the observed differences are interpreted as constituting evidence that L2 morphological processing is generally quantitatively different from L1 processing rather than qualitatively different (cf., Jiang, 2004, 2007; Keating, 2009; Sato & Felser, 2007).

Related points of discussion regarding the L2 data are the differences in the two reading sessions and in proficiency. First, the translation task seemed to boost L2 learners' ability to detect seeming and real agreement errors, showing a similar total time pattern in the translation session to native speakers reading for comprehension (i.e., the L2 times were longer but the pattern was similar, as discussed above). In the translation session, longer reading times appeared in the ungrammatical conditions, showing processing disruption in response to ungrammaticality, which was predicted and consistent with the native speakers. This finding suggests that the translation task required deeper processing during reading, such that the ungrammaticality was processed in the translation session in a process similar to the native speaker performance. The finding that the L2 group showed an effect of grammaticality in the translation session regardless of proficiency is also telling. Even the lower proficiency learners were able to detect ungrammaticality in the translation session, unlike in the comprehension session. This finding is in line with the previous studies (Jackson & Bobb, 2009; Leeser et al., 2001; Lim & Christianson, 2013b; Williams, 2006) that showed task effects in L2 processing. This finding contradicts concerns that the extra

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

processing burden imposed by translation could actually impede processing. Instead, low-level morphological processing improved under the presumably greater processing burden of L2 comprehension and reformulation into L1.

Second, effects in the late regression measures that appeared in the comprehension session did not appear in the translation session. We infer from this pattern that L2 learners seemed to generate translations silently as they read the source sentences, accordingly paying more attention to the input, such that they did not need to make regressions. Third, higher proficiency L2 speakers displayed longer total times and regressed more often in ungrammatical conditions in the comprehension session than did lower proficiency L2 speakers. This constitutes evidence that the ability to detect agreement violations develops as L2 proficiency increases, consistent with previous studies suggesting that proficiency significantly modulates L2 processing efficiency and accuracy (Foote, 2011; Hopp, 2006; Jackson, 2008; McDonald 2000, 2006).

Third, the translation data showed that plural local nouns in ungrammatical conditions led to greater production of plural markers of singular head nouns, whereas singular local nouns in grammatical conditions led to decreased production of plural markers on plural head nouns. The translation results can be taken as strong evidence that L2 speakers process inflectional morphemes during on-line reading, at least in tasks that focus attention on morphology. This observation is important because agreement morphology is not a primary issue in comprehending sentences in English (Pearlmutter et al., 1999). Moreover, the use of plural markers is not mandatory in comprehension or production in Korean. One might argue that the L2 participants' increased sensitivity to agreement morphology in the translation task stemmed wholly from a strategic and/or unnatural approach to the translation task. However, given that the plural morphology was not required to accurately translate any of the stimuli, it seems that the overt inclusion of the plural markers, and the modulation of the use of these markers in the Korean translations according to the experimental conditions, signals a thorough and relatively naturalistic processing of the agreement morphology. The most metalinguistically sophisticated approach to the translations would have been to ignore the English plural morphology completely. It cannot be said with any certainty whether this processing was "automatic" in the sense of Posner and Snyder (1974); however, we would argue that the recent body of literature documenting native speakers' failures to derive accurate, veridical representations from certain linguistic input (e.g., Christianson et al., 2001, 2010; Christianson & Luke, 2011; Ferreira, 2003; Slattery et al., 2013) suggests that at least some aspects of even native language processing consume cognitive resources over and above the threshold traditionally proposed for automatic processing.

In conclusion, the present study revealed that L2 learners can be sensitive to subject–verb agreement violations, depending on the task at hand. Apparently, L2 learners require more time to process both ungrammaticality and NP-match/mismatch than native speakers, but L2 speakers display sensitivity to morphological violations eventually, indicating quantitative rather than qualitative differences from L1 processing. In line with previous studies, we observed that L2 proficiency played a crucial role in modulating the degree of sensitivity to morphological violations: as proficiency increased, detection of ungrammaticality

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

sped up. In addition, even lower-proficiency L2 learners performed more like native speakers when reading for translation than when reading for comprehension, able to build faithful representations based on an accurate morphosyntactic parse when the task so demanded. This suggests that L2 processing may differ according to reading goals or task demands and that more demanding tasks can elicit more nativelike processing.

APPENDIX A

Stimuli

The sentences below are only presented in plural head noun and singular local noun in grammatical condition.

1. The teachers that instructed the student were very strict.
2. The boys that planted the tree were very tall.
3. The waiters that served the customer were very handsome.
4. The girls that liked the book were very popular.
5. The professors that invited the student were very supportive.
6. The runners that jumped over the hurdle were very fast.
7. The students that guessed the answer were very attentive.
8. The grandmothers that kissed the baby were very old.
9. The criminals that robbed the bank were very cruel.
10. The mothers that bought the apple were very good at bargaining.
11. The chefs that used the recipe were very famous.
12. The doctors that examined the patient were very tired.
13. The editors that corrected the article were very knowledgeable.
14. The boys that met the girl were very shy at first.
15. The nurses that injected the patient were very clumsy.
16. The guests that entered the room were very surprised.
17. The actors that entertained the viewer were very good at acting.
18. The dogs that ate the cookie were very chubby.
19. The mothers that baked the cake were very quiet in the kitchen.
20. The rabbits that beat the turtle were very pleased about the race.
21. The lions that chased the deer were very hungry in the jungle.
22. The lawyers that defended the client were very tired of the work.
23. The musicians that played the song were very passionate about the melody.
24. The nuns that helped the beggar were very kind to the poor.
25. The writers that wrote the column were very much interested in politics.
26. The guards that beat the prisoner were very violent.
27. The employers that recruited the worker were very thrifty.
28. The directors that made the film were very creative.
29. The knights that saved the queen were very brave.
30. The artists that painted the picture were very rich ten years ago.
31. The wars that destroyed the nation were very famous in history.
32. The students that read the book were very enthusiastic about the economy.
33. The neighbors that hosted the event were very loud.
34. The players that won the game were very lucky to get a lot of money.

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

35. The earthquakes that wrecked the town were very disastrous.
36. The designers that decorated the bedroom were very fashionable.
37. The painters that completed the project were very professional.
38. The dogs that followed the owner were very sick and tired.
39. The hunters that shot the bear were very ashamed of killing.
40. The dinosaurs that ate the bird were very ferocious.

ACKNOWLEDGMENTS

An earlier version of this research was presented at the Second Language Research Forum (Iowa State University, 2011). The authors thank three anonymous reviewers and Tessa Warren for their valuable comments. The authors also thank the members of the Educational Psychology Psycholinguistics Lab for helping with data collection. This research was funded by NSF Grant BCS-0847533 (to K.C.) and was supported by the KU Research Professor Program of Konkuk University.

NOTES

1. An anonymous reviewer asked whether task order made a difference in reading patterns, specifically, if doing the translation task first might have induced some sort of strategic reading behaviors. Post hoc analyses yielded no effect of task order in the reading times, except in regressions out: L2 participants regressed more from *very* when they performed the comprehension session first ($z = –2.16$, $p = .03$); conversely, translation first did not affect regression patterns in the comprehension session. In addition, performing translation first did not increase participants' reading times in the following comprehension session. Taken together, it appears that even if the translation task was first, any strategy that may have been employed in the translation task did not persist into the comprehension session.

REFERENCES

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Bock, K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*, 99–127.

Bock, K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, *8*, 57–99.

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*, 45–93.

Bock, K., Nicol, J., & Cutting, J. (1999). The ties that bind: Creating number agreement in speech. *Journal of Memory and Language*, *40*, 330–346.

Campbell, S. (1998). *Translation into the second language*. New York: Longman.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology*, *42*, 368–407.

Christianson, K., & Luke, S. G. (2011). Context strengthens initial misinterpretations of text. *Scientific Studies of Reading*, *15*, 136–166.

Christianson, K., Luke, S. G., & Ferreira, F. (2010). Effects of plausibility on structural priming. *Journal of Experimental Psychology*, *36*, 538–544.

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Coughlin, C., & Tremblay, A. (2012). Proficiency and working memory based explanations for nonnative speakers' sensitivity to agreement in sentence processing. *Applied Psycholinguistics*, *34*, 615–646.

Danks, J. H., & Griffin, J. (1997). Reading and translation. In J. H. Danks, G. M. Shreve, S. B. Fountain, & M. K. McBeath (Eds.), *Cognitive processes in translation and interpretating* (pp. 161–175). Thousand Oaks, CA: Sage.

Eberhard, K. M., Cutting, J.C., & Bock, J. K. (2005). Making syntax of sense: Number agreement in sentence production. *Psychological Review*, *112*, 531–559.

Ellis, R. (1988). The effects of linguistic environment on the second language acquisition of grammatical rules. *Applied Linguistics*, *9*, 257–274.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164–20.

Ferreira, F., Bailey, K., & Ferraro, V. (2002). Good enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.

Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research*, *30*, 3–20.

Ferreira, F., & Patson, N. D. (2007). The 'good-enough' approach to language comprehension. *Language and Linguistics Compass*, *1*/1–2, 71–83.

Foote, R. (2011). Integrated knowledge of agreement in early and late English-Spanish bilinguals. *Applied Psycholinguistics*, *32*, 187–220.

Frenck-Mestere, C. (2005). An on-line look at sentence processing in the second language. In R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 217–236). Amsterdam: Elsevier.

Gerver, D. (1976). Empirical studies of simultaneous interpretation: A review and a model. In R. W. Brislin (Ed.), *Translation: Application and research* (pp. 165–207). New York: Gardiner.

Havik, E., Roberts, L., van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing subject-object ambiguities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning*, *59*, 73–112.

Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research*, *22*, 369–397.

Jackson, C. (2008). Proficiency level and the interaction of lexical and morphosyntactic information during L2 sentence processing. *Language Learning*, *58*, 875–909.

Jackson, C., & Bobb, S. C. (2009). The processing and comprehension of *wh*-questions among second language speakers of German. *Applied Psycholinguistics*, *30*, 603–636.

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.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*, 603–634.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, *57*, 1–33.

Johnson, J.S., Shenkman, K. D., Newport, E. L., & Medin, D. (1996). Indeterminacy in the grammar of adult language learners. *Journal of Memory and Language*, *35*, 335–352.

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

Kim, J.-H., & Christianson, K. (2012). Sentence complexity and working memory effects in ambiguity resolution. *Journal of Psycholinguistic Research.* Advance online publication.

Lardiere, D. (1998). Case and tense in the "fossilized" steady state. *Second Language Research, 14*, 1–26.

Leeser, M. J., Brandl, A., & Whiteglass, C. (2011). Task effects in second language processing research. In P. Tromovich & K. McDonough (Eds.), *Applied priming methods to L2 learning, teaching and research: Insights from psycholinguistics* (pp. 179–198). Amsterdam: Benjamins.

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences, 106*, 21086–21090.

Lim, J-H. (2011). *Second language processing in reading and translation.* Unpublished doctoral dissertation, University of Illinois at Urbana–Champaign.

Lim, J-H., & Christianson, K. (2013a). Integrating meaning and structure in L1–L2 and L2–L1 translations. *Second Language Research*, *29*, 233–256.

Lim, J-H., & Christianson, K. (2013b). Second language processing in reading for comprehension and translation. *Bilingualism: Language and Cognition*, *16*, 518–537.

Macizo, P., & Bajo, M. T. (2004). When translation makes the difference: Sentence processing in reading and translation. *Psicologica*, *25*, 181–205.

Macizo, P., & Bajo, M. T. (2006). Reading for repetition and reading for translation: Do they involve the same process? *Cognition*, *99*, 1–34.

McDonald, J. L. (2000) Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics*, *21*, 395–423.

McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language*, *55*, 381–401.

Nicol, J., Forster, K., & Veres, C. (1997). Subject-verb agreement processes in comprehension. *Journal of Memory and Language*, *36*, 569–587.

Pearlmutter, N. J., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, *41*, 427–456.

Posner, M. I., & Snyder, C. R. R. (1974). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Hilsdale, NJ: Erlbaum.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

Sachs, J. S. (1967) Recognition memory for syntactic and semantic aspects of connected discourse. *Perception & Psychophysics*, *2*, 437–442.

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, *120*, 2022–2039.

Sato, M., & Felser, C. (2007). *Sensitivity to semantic and morphosyntactic violations in L2 sentence processing: Evidence from speeded grammaticality judgments*. Master's dissertation. University of Essex.

Slattery, T., Stuart, P., Christianson, K., Yoshida, M., & Ferreira, F. (2013). Lingering misinterpretations of garden path sentences arise from flawed semantic processing. *Journal of Memory and Language*, *69*, 104–120.

Staub, A. (2009). On the interpretation of the number attraction effect: Response time evidence. *Cognition*, *60*, 308–327.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, *36*, 201–216.

Tokowicz, N., & Warren, T. (2010). Beginning adult L2 learners' sensitivity to morphosyntactic violations: A self-paced reading study. *European Journal of Cognitive Psychology*, *22*, 1092–1106.

Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, *61*, 206–237.

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

Wei, L. (2000). Unequal election of morphemes in adult second language acquisition. *Applied Linguistics*, *21*, 106–140.

Williams, J. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition*, *9*, 71–81.

https://doi.org/10.1017/S0142716414000290 Published online by Cambridge University Press