Bilingualism: Language and Cognition 16 (3), 2013, 518–537 © Cambridge University Press 2012 doi:10.1017/S1366728912000351

Second language sentence processing in reading for comprehension and translation*

JUNG HYUN LIM
University of California, San Diego
KIEL CHRISTIANSON
University of Illinois, Urbana–Champaign & Beckman Institute for Advanced Science and Technology

(Received: October 22, 2011; final revision received: June 21, 2012; accepted: June 22, 2012; first published online 6 September 2012)

A self-paced reading and translation task was used with learners of English as a second language (L2) to explore what sorts of information L2 learners use during online comprehension compared to native speakers, and how task (reading for comprehension vs. translation) and proficiency affect L2 comprehension. Thirty-six Korean native speakers of English and 32 native English speakers read plausible and implausible subject relative clauses and object relative clauses. Reading times, comprehension accuracy, and translations were analyzed. Results showed that L2 learners were able to use syntactic information similarly to native speakers during comprehension, and that online L2 processing and offline comprehension were modulated by reading goals and proficiency. Results are interpreted as showing that L2 processing is quantitatively rather than qualitatively different from first language processing, i.e. strategically "good enough".

Keywords: L2 sentence processing, proficiency and task effects, Good-Enough processing, translation, integration of syntactic and semantic information

Several decades of research in second language acquisition (SLA) have revealed much information about WHAT can be acquired in second language learning (knowledge). The study of SLA also focuses on the issue of HOW a second language is used (performance), because both the sufficient knowledge of grammatical rules in a target language and the ability to apply the proper processing routines are necessary for successful second language (L2) mastery (Clahsen & Felser, 2006a; Gregg, 1996). Accordingly, a question of whether differences in ultimate attainment are in part attributable to underlying processing mechanisms in adult L2 learners has been raised in recent years. While there have been many experimental studies on how native speakers process their first language (L1) in real time using psycholinguistic techniques, it is only recently that the same psycholinguistic methods have been applied to the investigation of processing mechanisms that non-native speakers employ during L2 comprehension.

Much of the recent research on second language (L2) processing has explored how non-native speakers process language in real time when they read sentences in the L2, employing psycholinguistic techniques such as self-paced reading or eye-tracking paradigms (Felser, Roberts, Gross & Marinis, 2003; Frenck-Mestre & Pynte, 1997; Jackson, 2008; Juffs, 1998; Marinis, Roberts, Felser & Clahsen, 2005; Papadopoulou & Clahsen, 2003). One of the key questions in this literature is to what extent the behavior of L2 speakers varies from adult native speakers. In particular, many studies have examined L2 speakers' use of lexical and morphosyntactic information during online reading comprehension compared to native speakers. Some of these studies have indicated that there are characteristic differences between adult L2 learners and adult native speakers in grammatical processing (Felser et al., 2003; Marinis et al., 2005; Papadopoulou & Clahsen, 2003), while others have suggested that their respective underlying processing mechanisms are basically the same (Frenck-Mestre & Pynte, 1997; Jackson, 2008; Juffs, 1998). The issue of how language learners process L2 input still needs further investigations in part due to the discrepancies in existing research. In addition, a variety of structures, different populations, individual differences, proficiency level, and task demands should be considered (Clahsen & Felser, 2006a;

* An earlier version of this research was presented at the Second Language Research Forum (University of Maryland, 2010). The authors thank the members of the Educational Psychology Psycholinguistics Lab for helping with data collection and Steve Luke for advice about data analysis. The authors also thank three anonymous reviewers for their valuable comments. This research was funded in part by a UIUC Campus Research Board grant, a summer GA support award from the UIUC Bureau of Educational Research, and NSF BCS-0847533 to Kiel Christianson and a Cognitive Science/Artificial Intelligence (CS/AI) award from the Beckman Institute to Jung Hyun Lim.

Address for correspondence:
Jung Hyun Lim, Department of Educational Psychology, Education Building, MC-708 University of Illinois, 1310 S. 6th St., Champaign, IL 61820, USA
junglim@illinoisalumni.org

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

Dussias, 2001; Jackson, 2008; Juffs, 1998) to fully understand L2 processing mechanisms.

The present study adds to research exploring L2 processing mechanisms by examining how Korean learners of English read subject/object relative clauses (SRCs/ORCs) that are manipulated with respect to the semantic properties of the events they depict. This study further compares processing patterns in two reading tasks (reading for comprehension vs. reading for translation). A few studies report the roles of secondary tasks on influencing L2 speakers' reading time patterns (Jackson & Bobb, 2009; Leeser, Brandl & Whiteglass, 2011; Williams, 2006), demonstrating that secondary tasks such as an explicit grammaticality judgment or a memory task increase L2 learners' attention to grammatical structures. Given these observations, the current study also investigates how different tasks modulate learners' awareness of syntactic information during online comprehension. L2 proficiency is also considered, as several studies suggest that proficiency significantly affects processing (e.g., Hopp, 2006; Jackson, 2008).

This article is organized as follows. First, we summarize recent findings about L2 processing research and review Good-Enough (GE) processing (Christianson, Hollingworth, Halliwell & Ferreira, 2001; Ferreira, Bailey & Ferraro, 2002; Ferreira, Christianson & Hollingworth, 2001; Ferreira & Patson, 2007), which is the main theoretical focus in this study. Then, the rationale of using a translation task is illustrated. The subsequent section describes research questions along with the predictions and methods of the current study, followed by our results, which are interpreted within current L2 processing research.

L2 sentence processing research

Some studies that have examined L2 speakers' real-time comprehension provide evidence that L2 speakers lack the ability to use particular L2 structures during online processing. Papadopoulou and Clahsen (2003) found that using a self-paced reading paradigm, both high-proficiency L2 learners (Spanish, German, and Russian L1) of Greek and a Greek native group preferred low attachment in reading ambiguous relative clauses that contain a thematic preposition "with". The results, however, included the finding that the L2 learners did not show any attachment preferences when a sentence contains the preposition "of" as in "A man looked at the teacher of the pupil who was in the school yard", whereas native speakers constantly showed a high-attachment preference. The authors concluded from these results that L2 learners have difficulty integrating syntactic information during online comprehension, while native speakers are able to use both syntactic and lexical information. These results were replicated by Felser et al. (2003) with temporarily ambiguous relative clauses, suggesting that native and non-native speakers differ in that the L2 comprehension mechanisms do not always integrate phrase-structure information immediately.

Marinis et al. (2005) provided additional evidence that L2 speakers lack detailed syntactic representations during online comprehension, particularly in using intermediate syntactic gap information involving long-distance *wh*-phrases. In a self-paced reading task, L2 learners of English (Chinese, Japanese, Greek, and German L1) and native English speakers read sentences involving long-distance *wh*-dependencies. It was found that the L1 English speakers made use of the intermediate syntactic gap to retrieve the filler to integrate with its subcategorizer, whereas the L2 speaker groups failed to postulate such intermediate gaps during processing. These results were also replicated in Felser & Roberts (2007) using a cross-modal picture-priming task. To account for the above differences between L1 and L2 processing, the SHALLOW STRUCTURE HYPOTHESIS (SSH) has been proposed (Clahsen & Felser, 2006a), arguing that grammatical processing in language learners is qualitatively different from that of native speakers. The main idea of the hypothesis is that the syntactic representations learners compute during online comprehension lack the full-fledged L2 grammar, and thus learners are dependent on lexical-semantic or pragmatic knowledge to compensate for incomplete syntactic structure.

On the other hand, a contrasting set of studies suggests that L2 speakers' ability to use certain types of morphosyntactic information during online processing is quantitatively rather than qualitatively different from that of native speakers. Frenck-Mestre and Pynte (1997) showed that English–French and French–English bilinguals performed complete syntactic parses and patterned in the same manner as monolingual counterparts in reading sentences containing temporarily ambiguous prepositional phrases as in the English *They accused the ambassador of espionage/Indonesia but nothing came of it*, and *He rejected the manuscript on purpose/on horses because he hated its author*. Non-syntactic, lexical information also reliably influenced reading L2 sentences. The authors claimed that non-native processing is not necessarily different from native processing, at least with respect to this particular structure.

Juffs (1998) also investigated whether L2 speakers are able to process sentences that contain main verb/reduced relative ambiguities in the same way as native speakers, manipulating argument structures and lexical-semantic cues. The results of an acceptability judgment task and a self-paced reading task suggested that the learners of English as a second language (ESL) deployed complex knowledge of English in online tasks and that the ESL learners were sensitive to the interaction of various

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

information sources in much the same way as native speakers. Another study, by Jackson and Dussias (2009), also indicated that highly proficient L2 German speakers were able to fully utilize case-marking information in *wh*-question sentences during online reading comprehension.

Furthermore, an L2 proficiency effect has been reported in several studies (Frenck-Mestre, 2002; Hopp, 2006; Jackson, 2008), where advanced speakers have been found to use morphosyntactic information more automatically than less proficient speakers. Hopp (2006) investigated L2 processing of subject–object ambiguities in German with both advanced and near-native L2 learners (L1 English and Dutch). German has a canonical subject–object–verb (SOV) word order, but the object can precede the subject in subordinate clauses. The question of interest was whether advanced and near-native L2 speakers of German would be able to use syntactic disambiguation information (case-marking vs. verb agreement) in the same way as native German speakers when processing subject–object ambiguities. Results showed that near-native L2 speakers patterned like native speakers in online reading time, consistently showing local slowdowns in the critical region of OS sentences in both types of syntactic disambiguation (case vs. verbal agreement) conditions. On the other hand, advanced L2 speakers' reading times were generally slow in the OS word order only at the end of the sentence, regardless of syntactic disambiguation types. These findings emphasized the role of proficiency in L2 processing mechanisms, suggesting that sensitivity to syntactic information is not restricted in non-native processing.

Jackson (2008) extended Hopp's (2006) study, examining how L2 speakers process temporarily ambiguous *wh*-questions compared to native German speakers. Accuracy results showed that both native and non-native speakers had greater difficulty in reading temporarily ambiguous OS word order than SO word order. The online reading time results showed that the use of case-marking information varied according to L2 proficiency level. Native speakers and advanced L2 speakers assigned grammatical roles incrementally, whereas intermediate L2 speakers did not show immediate sensitivity to case-marking information.

To summarize, the L2 studies conducted so far have provided mixed results; some of the studies show L2 speaker's ability to use syntactic knowledge during online comprehension in the same manner as native speakers (Frenck-Mestre & Pynte, 1997; Hopp, 2006; Jackson, 2008; Jackson & Bobb, 2009; Williams, 2006), whereas others suggest that L2 processing lacks abstract linguistic structure compared to native speakers (Felser et al., 2003; Felser & Roberts, 2007; Marinis et al., 2005; Papadopoulou & Clahsen, 2003). Furthermore, there are studies to indicate that individual differences, such as proficiency, working memory capacity, and task play a role in the L2 processing mechanism (Hopp, 2006; Jackson & Bobb, 2009; Leeser et al., 2011; Williams, 2006)

Good-Enough processing

Interestingly, shallow processing does not appear to be unique to L2 learners; the "shallow and less detailed" language processing discussed in the L2 processing literature has also been observed in monolingual processing studies. The GOOD-ENOUGH (GE) APPROACH to language comprehension (Christianson et al., 2001; Christianson & Luke, 2011; Christianson, Luke & Ferreira, 2010; Christianson, Williams, Zacks & Ferreira, 2006; Ferreira et al., 2002; Ferreira & Patson, 2007) proposes that comprehenders compute linguistic input using both syntactic and semantic or heuristic processing, but the outputs of these routes are not always reconciled if one output appears good enough for comprehension, thus yielding a final interpretation that is not completely faithful to the input. GE language comprehension has been supported by studies using active/passive structures (Christianson et al., 2010; Ferreira, 2003), and garden-path sentences (Christianson et al., 2001, 2006).

Ferreira (2003) found that native English speakers were more likely to misinterpret simple passive sentences when syntactic structure was inconsistent with world knowledge, as in *The cat was chased by the mouse*. In particular, participants were less accurate in implausible sentences (80%) than plausible sentences (90%) for passives, whereas for actives, they were as accurate with implausible sentences (95%) as with plausible sentences (98%). The frequency of the syntactic structure did not influence misinterpretations, as comprehenders responded to infrequent subject clefts as accurately as to active sentences. These findings indicated that native English speakers tend to employ heuristic processing mechanisms along with a syntactic processing algorithm. These heuristics include semantic plausibility and Noun–Verb–Noun (NVN) word order template that maps constraints of this order onto Agent–Verb–Patient interpretations (Townsend & Bever, 2001; see also the Competition Model in connection to linear templates, e.g., Bates & MacWhinney, 1987). The passive structure, which has an atypical Patient-before-Agent order, is apparently harder to process than the active structure due to the conflict between the syntactic algorithm and the semantic heuristic. The syntactic representations in the comprehender's mind are thus considered to be "fragile", and the heuristic parser to be "fast and frugal" at reaching an interpretation of a sentence when processing linguistic input.

Christianson et al. (2010) also investigated the interaction of semantic and syntactic routes in normal comprehension, by combining a comprehension paradigm and a structural priming paradigm. Active/passive

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

structures were adapted from Ferreira (2003), and error rates in answering comprehension questions rose when structure conflicted with the plausibility of the sentence, replicating Ferreira (2003). A more important finding was that both plausible-passive and implausible-active sentences tended to prime passive structures in production data, suggesting that the outputs of both syntactic and semantic routes are not always seamlessly reconciled.

Depth of language processing has also been found to be influenced by reading goals and task demands. Swets, Desmet, Clifton, and Ferreira (2008) investigated whether frequency and the type of comprehension questions could affect depth of processing and final interpretation in syntactic ambiguity resolution. They found that native English speakers were faster in reading globally ambiguous sentences when asked superficial comprehension questions. Also, when task demand increased, that is, in a condition that asked specifically about attachment of the globally ambiguous portion of the sentence, reading times were significantly longer than in conditions with superficial questions. Lastly, reading times were slower in ambiguous sentences when answering detailed questions about the interpretation of the ambiguity than in the disambiguated sentence conditions. Based on these findings, Swets et al. (2008) concluded that language processing can sometimes be INCOMPLETE, leaving ambiguous sentences unresolved when there is no information to disambiguate them. Thus, language processing can be strategic and goal-dependent.

Precisely how much of human language processing is "good enough" is not yet clear; however, this framework leads to new ways of looking at human language processing mechanisms. Also, as the "good-enough" or "shallow" processing appears to exist in both native speakers and L2 learners, and there is a growing body of research that reports L2 speakers' native-like processing (Hopp, 2006; Jackson, 2008; Jackson & Bobb, 2009; Williams, 2006), the GE approach might provide the more integrated account of L1 and L2 processing. In the GE processing account, the language processing system assumes two different routes, the grammar and semantic-based heuristics. The syntactic algorithm is fully operational alongside semantic-based heuristics such as lexical-semantic information, world knowledge, or the NVN heuristic. But the syntactic representation is "fragile" and quickly decaying (Sachs, 1967), so that it is easily overridden by surface-level heuristics, sometimes leading to interpretations that are unfaithful to the original text. In context of L2 processing, it may be that the syntactic output might be relatively more fragile, depending on proficiency level, and thus be more susceptible to semantic/heuristic influences. In addition, within the GE approach the depth or veridicality of language processing is hypothesized to be goal-dependent and affected by task (Swets et al., 2008), not unlike some recent existing studies pointing to the effects of task on L2 processing (Jackson & Bobb, 2009; Leeser et al., 2011; Williams, 2006).

Figure 1 demonstrates the mechanisms of the GE processing in comparison to the SSH, which are clearly related views in language processing. Figure 1a shows the SSH, proposing that lower, syntactic, route is restricted for L2 speakers (see Clahsen & Felser, 2006b). In the SSH, the complete syntactic parsing is rather less detailed in L2 processing (at least at some levels of proficiency) than native speakers due to the inability to construct fully specified syntactic structures in real-time processing. This results in "shallow parsing", characterized by over-reliance on lexical-semantic information.

On the other hand, the GE account predicts that, as long as there is no deficit in syntactic knowledge (i.e., as long as proficiency is sufficient), access to the syntactic processing route is unimpaired. This is illustrated by the solid line between the input and the full parsing route in Figure 1b. Instead, what results in less-than-veridical final representations is incomplete (or absent) integration of the outputs of the two processing routes, which is illustrated by the dotted, integrative line between good-enough and full representations in Figure 1b. Adding this integrative line allows for integration or lack of it under certain circumstances. In other words, the GE account proposes that even if the syntactic route is available, its output does not always get integrated with the output from the heuristic route, which can happen in L1 processing as well as L2 processing. The model shown in Figure 1b does not require the arrows tracing the derivation of the full representation to be solid, i.e., full, for any given speaker in any given situation. Rather, the point is that they CAN be, and the ultimate interpretation can STILL be "good enough" if the outputs from the two processing routes are not integrated. Furthermore, whether the arrow from the L2 grammar into the "full parsing" box will be solid or dotted for any given L2 speaker is predicted to be dependent at least in part on proficiency and task.

Translation task

In L2 processing research, comprehension questions have commonly been used as an offline measure to confirm readers' comprehension. Comprehension questions, however, might not provide much insight into how L2 learners integrate morphosyntactic and semantic information during comprehension. Question form might in fact alter both response and processing patterns (Christianson & Luke, 2011; Swets et al., 2008). With this concern in mind, the present study employs a translation paradigm as an offline task. Translation is an important process that L2 speakers undertake when dealing with two languages, providing information about cognitive and linguistic development

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

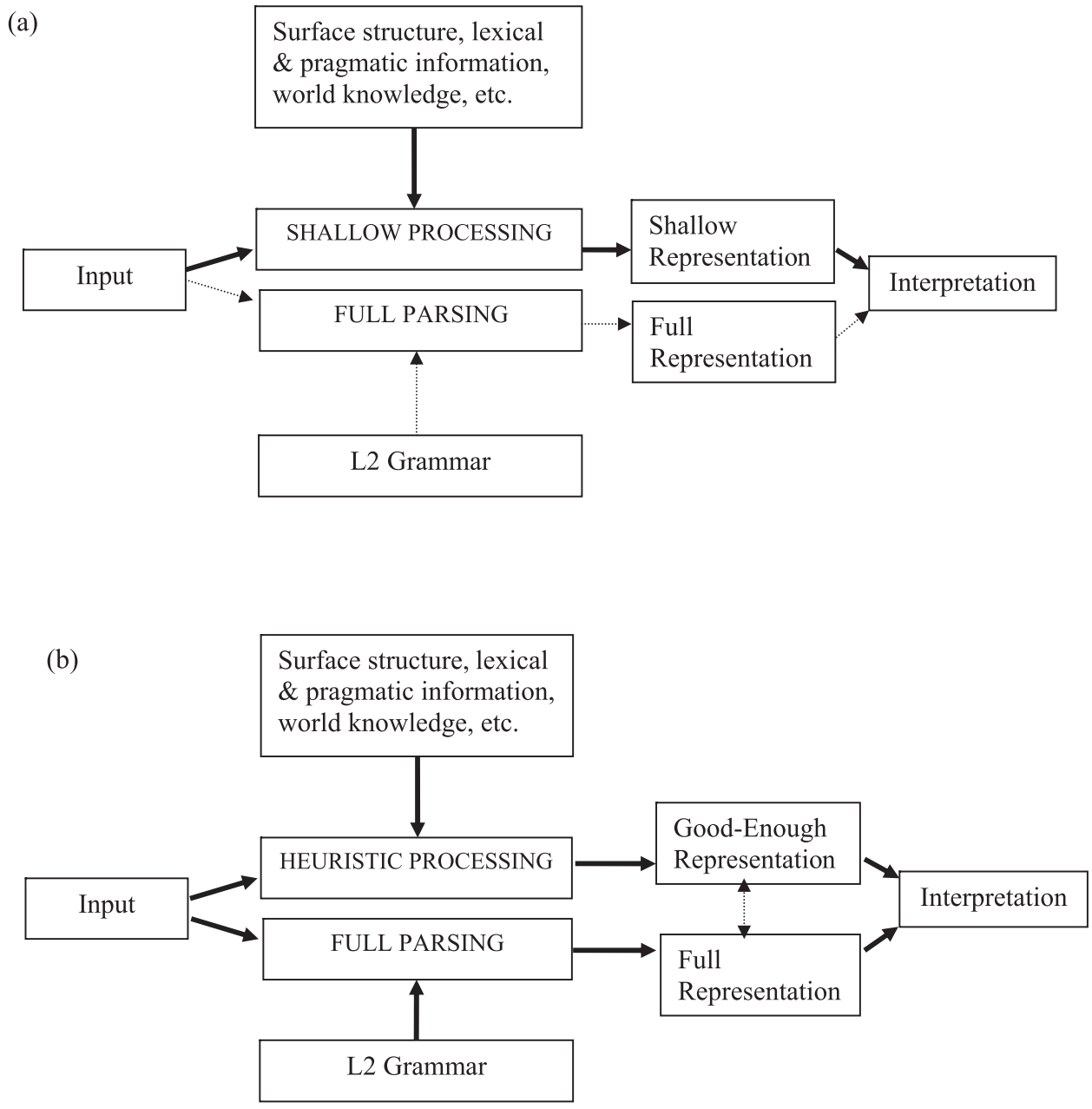


Figure 1. Mechanisms of Good-Enough processing in comparison with the Shallow Structure Hypothesis. (a) Shallow Structure Hypothesis for L2 processing, adapted from Clahsen & Felser (2006b). (b) Good-Enough processing, for both L1 and L2 processing. SRC-pl = subject relative clause, plausible; SRC-impl = subject relative clause, implausible; ORC-pl = object relative clause, plausible; ORC-impl = object relative clause, implausible

(Campbell, 1998; Malakoff, 1992). It was also once considered as a common practice to test learners' reading comprehension in language teaching situations (Buck, 1992; Cordero, 1984; Malakoff, 1992). As such, translation may provide a window into L2 learners' integration of morphosyntactic and semantic information, because translation provides explicit evidence of how a language learner has comprehended a given input (Campbell, 1998).

In cognitive theories of the translation process, it is assumed that three major processes occur during translation, that is, to comprehend a source language, to switch between two languages, and to produce in a target language. One of the two views of how these processes work during translation, the HORIZONTAL APPROACH, suggests that the translation process requires additional attention and increase cognitive load compared to normal reading, because one needs to establish

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

semantic and syntactic matches between two languages during translation in a parallel manner (Danks & Griffin, 1997; Gerver, 1976). Macizo and Bajo (2004, 2006) found that both translators and L2 learners took longer to read sentences for translation than for repetition, indicating that translation includes reformulation process between the source language and target language during comprehension, which is consistent with the horizontal approach.

Since translation entails a complex cognitive process, including comprehension, production, and additional working memory demands, compared to normal comprehension, one could ask whether translation can serve as a valid measure of online processing. That is, it could be argued that translation does not reflect online sentence comprehension, but rather measures the "memory" for sentence content. There is no clear answer for this concern, as few experimental cognitive studies have been conducted on translation; however, if translation is a "horizontal" process (Macizo & Bajo, 2004, 2006), it should proceed in a relatively incremental fashion during processing, and thus does in fact reflect processes that take place during online comprehension. Nevertheless, even if translation is merely an indicator of the "memory" of a sentence, readers must still process that sentence, including syntax, in real time during reading to produce their translation. Consider, for example, the output of the L2-to-L1 translation process. It must derive from some combination of the input, the translator's L2 and L1 grammars, and the translator's world knowledge. Assuming the syntactic structure of the L2 input is comprehensible to the translator and available in the L1, that structure should be maintained in the L1 translation. If the L2 structure is not comprehensible (i.e., too difficult or unfamiliar), then we would expect an L1 translation based on L2 lexical semantics and the world knowledge of the translator, reflecting a default more frequent or simpler L2 structure that did not occur in the input. If the L2 structure is comprehensible, but the semantic content of the input interferes with comprehension, we would expect a translation that maintains the L2 input structure in the L1 output, but that displays a semantic meaning that was not in the L2 input. Thus, if a translation accurately reflects the syntax of the input, it seems quite clear that the syntax in the translation output is derived from the input, especially if the syntax in the translation is difficult or non-canonical (e.g., an object relative clause or a passive structure).

In addition, translation is taken as a type of reading task in this study, not just a method to elicit L2 learner output. Thus, by comparing reading processes in normal comprehension and translation, we can determine whether there may be a possible effect of reading goals in L2 sentence comprehension, similar to recent L1 research reporting that reading goals and task demands can influence processing routines (Swets et al., 2008). Because some of previous studies have pointed out task differences influencing L2 processing (Jackson & Bobb, 2009; Leeser et al., 2011; Williams, 2006), the translation task in this study not only provides offline data but also broadens our understanding of how L2 processing can be modulated by reading goals/tasks. Task effects are furthermore particularly informative to the question of whether the GE processing account can account for L2 processing as well as L1 processing. Because data from a translation task indicate how L2 speakers understand the input, it is useful for examining whether L2 processing can also be explained by the GE processing account, which predicts effects of task or goal on language processing and determines whether conflicting information sources are integrated or not (Ferreira & Patson, 2007).

The current study

The first question addressed in the present study is how Korean learners of English integrate morphosyntactic and semantic information during L2 processing using the self-paced reading paradigm in combination with a translation task. The term "semantic information" throughout the study is used to refer to general world knowledge, that is, how entities referred to in a given sentence usually behave in the real world. This study is in part motivated by the study of Ferreira (2003), extending the basic idea to L2 populations and examining subject relative clauses (SRCs) and object relative clauses (ORCs), instead of active/passive structures. Sentences containing SRCs like *The man that bit the dog*... and ORCs like *The man that the dog bit*... have been of interest to psycholinguists, since these structures can assess the role of verbal working memory (Just & Carpenter, 1992; King & Just, 1991), and are useful in examining gaps and traces in syntax (Gibson, Hickok & Schütze, 1994; Nicol & Pickering, 1993). Previous studies in the L1 processing literature show that ORCs generally cause greater processing difficulty than SRCs. There are several accounts to explain this difference, including memory-based accounts, syntax-based accounts, and a perspective-shifting account. Among them, in the syntax-based accounts, comprehenders are required to treat *the man* as a subject and an object at the same time. But the difficulty of assigning inconsistent thematic and syntactic roles to a single NP increases the difficulty of processing ORC constructions. In addition, according to the active filler strategy (Clifton & Frazier, 1989), comprehenders may experience misanalysis and require recovery processes to reach the correct syntactic analysis. Accordingly, the ORC construction involves an initial misanalysis and a reanalysis process to assign correct syntactic and thematic roles. Readers must "retrieve" correct syntactic information, and thus this process increases reading times compared to the relatively easier SRC

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

structure. Thus, the pair of structures is asymmetrical in canonicity and processing difficulty in English, analogous to active/passive structures (Ferreira, 2003; Holmes & O'Regan, 1981; King & Just, 1991; Traxler, Morris & Seely, 2002). Like English, SRCs has also been found to be easier to comprehend than ORCs in Korean, despite Korean being a left-branching language where the relative clause precedes the modifying noun. This preference has been demonstrated by both Korean L1 and L2 learners, suggesting fundamental similarity across languages in the processing of this particular syntactic representation (Cho, 1999; O'Grady, Lee & Choo, 2003).[1]

By manipulating both syntactic and semantic information, this study explores whether L2 speakers show similar processing patterns as native speakers with regard to using structural and semantic knowledge. At the same time, the study aims to examine whether the GE approach can account for L2 processing as well as L1 processing. If both native speakers and L2 learners are able to use both syntactic and semantic routes in processing, but the syntactic representation is sometimes overwhelmed by semantic information, both groups should have more trouble interpreting ORC structures with implausible semantic events than any other condition, with more difficulty experienced by lower-proficiency L2 speakers. Also, L2 learners should produce a number of translations that maintain the correct morphosyntax, but not the thematic roles of the nouns from the input. Such a pattern would suggest that L2 processing is "good enough" in the same manner as native processing (Christianson et al., 2001, 2010; Ferreira, 2003), as it would provide evidence of both syntactic and semantic/heuristic processing routes being active in parallel, and misinterpretation coming from the lack of integration of the two routes as in Figure1b. Furthermore, if L2 learners can retrieve structural knowledge during online comprehension, reading times should be different between the structures, specifically yielding longer reading times in the ORC conditions than the SRC conditions. However, if language learners lack the ability to use morphosyntactic information during online comprehension, then reading times should not differ across different structures, and participants' translations will not reflect the structures of the input sentences, especially in the non-canonical, more difficult ORC conditions.

The second question is how reading goals and L2 proficiency affect L2 processing by comparing reading for comprehension versus reading for translation. Predictions related to the task manipulation are largely informed by previous studies that indicate a possible influence of task on language processing (Macizo & Bajo, 2004, 2006; Swets et al., 2008). If L2 processing is affected by reading goals, then the availability of structural and semantic information should differ depending on reading sessions, but there could be two possible scenarios. First, since translation could motivate L2 learners to stay more alert than during reading for normal comprehension, L2 learners may be more involved in interpretive work, resulting in greater sensitivity to semantic information in the translation session than in the comprehension session. Alternatively, translation might increase L2 learners' awareness of syntactic information compared to normal reading for comprehension, as translation requires more attention to linguistic input. These differences in either way would provide evidence that the depth of L2 processing is dependent on reading goals, consistent with the GE processing account. In addition, accuracy rates in response to comprehension probes are expected to be higher in the translation session if translation requires deeper interpretive work. Whether any possible different effects of the two reading goals stem from additional working memory load associated with the translation task or some other processes inherent in translation (such as constraints imposed by the dual task of comprehension and production) remains to be explored, but this issue is beyond the scope of this paper.

L2 proficiency is also predicted to influence both online processing patterns and comprehension and translation accuracy rates. Based on the previous studies that showed proficiency effects in the use of syntactic information (Hopp, 2006; Jackson, 2008), L2 learners are expected to use increasingly more syntactic information as proficiency increases, and accordingly accuracy rates would be lower in the low-proficiency group. More precise predictions are difficult, as there have not been many studies done on task effects. Observations made in the present study will therefore broaden our understanding of L2 learner's reading strategies and how they are affected by proficiency and cognitive processes associated with different reading tasks.

Method

Participants

Participants were recruited from the University of Illinois at the Urbana–Champaign community. Thirty-six native Korean speakers who began learning English after puberty participated; 33 native English speakers served as a control group. Participants were paid $10 or received course credit for participation. English proficiency of the L2 speakers was measured using a cloze test (Giuli

[1] Cho (1999) found that Korean subject relative clauses were easier both to comprehend and to produce than object relative clauses in Korean L1 child acquisition. In-depth morphosyntactic comparison of Korean and English relative clause structures is outside the scope of the present study, but interested readers are directed to O'Grady, Lee and Choo (2003) for a comparative analysis of these structures.

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

Dussias, Pennsylvania State University, with permission), which had a maximum possible score of 40.

Materials

There were 48 experimental items and 96 filler sentences. The experimental sentences can be found in the Appendix. The fillers consisted of a variety of structures, none of which included asymmetry in plausibility. Forty-eight experimental sentences were constructed as in (1) below. Types of structure (SRC, ORC) and plausibility (plausible, implausible) were manipulated, resulting in four conditions. Plausibility was manipulated only in the relative clauses. The sentences below are examples of the four conditions: (1a) illustrates a subject relative clause with plausible event; (1b), a subject relative clause with implausible event; (1c), an object relative clause with plausible event; and (1d), an object relative clause with implausible event.

(1) a. The cat that chased the mouse was fast.
(SRC-plausible)

b. The mouse that chased the cat was fast.
(SRC-implausible)

c. The mouse that the cat chased was fast.
(ORC-plausible)

d. The cat that the mouse chased was fast.
(ORC-implausible)

The experimental sentences were distributed across four different lists in Latin Square design, such that each participant read only one version of each experimental sentence. Both experimental and filler sentences were randomly ordered during presentation.

In the relative clause, 19 of 48 stimulus items contained the same semantic relations (plausibility) as the items adapted from Ferreira (2003). These 19 semantic properties such as *The cat chased the mouse*/*The mouse chased the cat* were normed by English native speakers in Ferreira's (2003) study. Plausibility for the relative clauses in the remaining 29 items was determined via a norming study that was separately conducted before the current main experiment. The norming study was conducted on 46 native speakers of English who did not participate in the main experiment. The purpose of the norming study was to select the remaining 29 sentences for the main experiment. The norming survey contained a total of 70 sentences, consisting of 40 items and 30 fillers, the structure of which was a simple active such as *The mom bathed the baby*. The 40 biased, reversible items were presented both in plausible and implausible versions across the two lists. The remaining 30 sentences contained symmetrical events as fillers. The target sentences were distributed over two lists, such that each participant saw only one version of any given sentence. Participants were directed to read each sentence and rate each item on a scale from 1 ("implausible") to 7 ("extremely likely"). The results of the norming study are presented in Table 1. Twenty-nine sentences were chosen among the 40 items based on the grand mean in each version. As shown in Table 1, the grand mean of the 40 items was 6.51 for the plausible version and 2.20 for the implausible version. Twenty-nine sentences were selected with means over 6.51 in the plausible version and under 2.20 in the implausible version. The grand mean of the selected 29 items was 6.74 in plausible version and 1.92 in the implausible version. A paired-sample *t*-test showed that the implausible sentences were less plausible than the plausible sentences ($t = 26.29$, $p < .01$).

Table 1. *The mean and standard deviation of the norming survey (standard deviations in parentheses).*

Sentence type	Mean rating from 1 (implausible) to 7 (plausible)
Biased reversible, plausible	6.51
(the doctor treated the patient)	(0.70)
Biased reversible, implausible	2.20
(the patient treated the doctor)	(0.87)

Procedure

The main experiment was run on E-prime professional software version 2.0. Native speakers read each sentence in a self-paced, word-by-word, non-cumulative reading paradigm. They were then asked to identify whether or not a word appeared in the sentence by pressing the "yes" or "no" button, in order to verify that participants read each sentence completely, i.e., they did not skip any words during reading. After the brief word identification task, a paraphrase verification sentence was presented to make a judgment on whether the verification sentence represented the same meaning of the sentence that they read just before (Kim, 2008). An example of a paraphrase verification question is in (2) for the stimulus item *The cat that the mouse chased was fast*. A brief practice session was provided before the main experiment for participants to get used to the procedure.

(2) a. The cat chased the mouse. The cat was fast.

b. The mouse chased the cat. The cat was fast.

For L2 learners, the procedure was largely similar; however, the experiment for the L2 group consisted of two sessions: (i) reading for comprehension and (ii) reading

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

Table 2. *Grand mean reading times in milliseconds (standard deviations in parentheses), native speakers.*

	SRC-plausible	SRC-implausible	ORC-plausible	ORC-implausible
RC region	1222.97	1215.55	1280.56	1320.98
	(319.71)	(395.81)	(382.93)	(466.11)
Matrix verb	357.77	464.58	456.61	577.38
	(94.40)	(227.44)	(164.36)	(333.75)
Post-verb region	1231.17	1318.11	1258.57	1450.93
	(376.85)	(416.03)	(362.09)	(426.13)

RC = relative clause; SRC = subject relative clause; ORC = object relative clause

for translation, resulting in L2 participants reading 24 experimental sentences and 48 fillers in each session. The 48 experimental items and 96 fillers were randomly divided in half between the comprehension and translation session. The order of the comprehension and translation session was counterbalanced across participants to control for possible block order effects, and a brief practice session was performed at the beginning of each session.

In the comprehension session, the participants read each sentence according to the previously described procedures. In the translation session, the procedure was the same as the comprehension session, except that the participants translated the English (L2) sentence they had read into Korean (L1) verbally after the word identification task. Translations were recorded for analysis. Then, they completed the paraphrase verification task by making the yes/no judgment. After the whole experiment, participants completed the cloze test and language survey to provide language background information.

Results

Reading time data

Native speakers

Data-trimming was conducted on the comprehension accuracy of the paraphrase verification task for the filler items only, because experimental items were predicted to have some amount of incorrect answers (i.e., in implausible versions). Thirty-two participants answered correctly over 90% of the time, but one had just 73% accuracy, and was therefore excluded from further analyses. Reading times greater than two standard deviations from the mean were replaced with the mean for the particular sentence in the same condition. In addition, reading time analyses were conducted over all trials except the filler items, regardless of correctness of the trials, because incorrect answers were expected in some conditions, especially implausible versions. Three critical regions of interest were the RELATIVE CLAUSE (RC) region, which consisted of the words within the relative clause (reading times for each word in the RC were summed to create this single dependent variable), the MATRIX VERB region, and the POST-VERB region, consisting of the rest of the sentence. Reading times for the rest of the words in the sentence were summed in the same way as in the RC to create the single dependent variable. An example of the critical regions is presented in (3).

(3) The cat that chased the mouse (RC) was (matrix verb) fast. (post-verb)

Reading time data were analyzed in a linear mixed-effects model (Baayen, Davidson & Bates, 2008). Each model was fitted using a stepwise selection procedure where factors failing to reach at least $p < .10$ significance were removed from the final model (Luke & Christianson, 2011). Therefore, the effect of structure, plausibility, and the interaction between structure and plausibility were included as predictors in the final, best-fitted model. Table 2 presents the mean reading times for the three critical regions in the four conditions, and Table 3 presents the result of the linear mixed-effects model for native speakers.

Participants read the RC region significantly faster in the SRC conditions regardless of plausibility ($t = 2.37$, $p = .01$). The ORC-implausible condition yielded the longest reading times numerically, but neither main effect of plausibility nor interaction between the two factors approached significance. At the matrix verb, both main effect of structure ($t = 4.31, p < .01$) and plausibility were reliable ($t = 4.70, p < .01$). The SRC conditions were read faster than the ORC conditions, and implausible sentences took longer to read than plausible sentences. There was no significant interaction. Even after the matrix verb, a similar pattern was observed, yielding main effects of both structure ($t = 2.03$, $p = .04$) and plausibility ($t = 3.67, p < .01$) with no interaction.

In sum, the SRC-plausible condition was the easiest condition to read, and the ORC-implausible condition was

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

Table 3. *Fixed effects in the linear mixed effects model of reading times, native speakers.*

	Predictor	Estimate	SE	*t*-value	*p*
RC region ($R^2 = 0.29$)	Intercept	1260.05	68.06	18.52	<.01
	Structure	39.41	16.60	2.37	.01*
	Plausibility	9.47	16.60	0.57	.56
	Structure × Plausibility	11.64	16.60	0.70	.48
Verb region ($R^2 = 0.16$)	Intercept	464.13	32.98	14.08	<.01
	Structure	52.34	12.13	4.31	<.01*
	Plausibility	57.01	12.13	4.70	<.01*
	Structure × Plausibility	2.92	12.13	0.24	.80
Post-verb region ($R^2 = 0.38$)	Intercept	1314.83	87.78	14.98	<.01
	Structure	39.35	19.33	2.03	.04*
	Plausibility	70.99	19.33	3.67	<.01*
	Structure × Plausibility	26.53	19.33	1.37	.17

*p <.05; †p <.10; RC = relative clause; SE = standard error

the most difficult, yielding the longest reading times at all three regions, although no interactions were found at any regions. Participants were sensitive to both structure and plausibility from the verb region through the rest of the sentence, but the effect of structure was seen earlier in the RC region than the effect of plausibility.

L2 speakers

Comprehension accuracy performance on the filler items was overall 92% correct, and data were trimmed in the same way as for the native speakers. The experimental items for L2 participants were the same as the ones used for the native speakers. The two reading sessions (comprehension vs. translation) for the L2 group resulted in two separate analyses to examine whether reading goals influence the way that participants process syntactic and semantic information. L2 speakers' reading times greater than two standard deviations from the mean were replaced with the mean of each L2 learners' mean reading times for the particular sentence in the specific condition. Reading times for L2 speakers in each reading session were analyzed in a mixed-effect model in the same way as the native speaker data, treating proficiency scores as a continuous variable to examine proficiency effects. As a result of the stepwise selection procedure (Luke & Christianson, 2011), effects of structure, plausibility, proficiency, and the interactions between each factor were included as predictors in the final, best-fitted models, when proficiency contributed to the model. When proficiency did not reach significance ($p < .05$), only effects of structure, plausibility, and the interaction of the two factors were included in the final model. Subject and item were also included as crossed random effects. Table 4 presents the mean reading times of L2 speakers at three regions of interest in both reading sessions. Table 5 shows the results of the final model for three regions of interests in each reading session.

In comprehension, L2 speakers showed a reliable interaction between structure and plausibility ($t = –3.42$, $p < .01$) in the RC region, indicating that only in the plausible sentences was there a structural difference, and only in the SRC sentences was there a plausibility effect. An effect of plausibility ($t = 2.01$, $p = .04$) was observed as well to indicate that implausible sentences yielded longer reading times than plausible sentences. Proficiency did not significantly contribute to the model. At the verb, there were main effects of both structure ($t = 3.94$, $p < .01$) and plausibility ($t = 6.33$, $p < .01$), but no interaction was found. Reading times were inflated when the sentences were implausible, and the ORC conditions were read more slowly than the SRC conditions. In addition, proficiency marginally interacted with plausibility ($t = 1.75$, $p = .08$), indicating that only higher-proficiency learners displayed reduced reading times in the plausible conditions than the implausible conditions. After the matrix verb, plausibility exhibited a reliable main effect ($t = 4.64$, $p < .01$), but there was no main effect of structure. Also, there was a significant interaction between proficiency and plausibility ($t = 2.00$, $p = .04$).

In the translation session, there was a main effect of plausibility at the RC, ($t = 6.12$, $p < .01$), indicating that reading times for implausible sentences were significantly longer than for plausible sentences. Neither an effect of structure nor an interaction between structure and plausibility was found at the RC. The interaction of structure and proficiency reached marginal significance ($t = –1.82$, $p = .06$), showing that the reading times of the lower-proficiency learners increased in ORCs more than

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

Table 4. *Grand mean reading times in milliseconds (standard deviations in parentheses), L2 speakers.*

(a) Grand mean reading times in milliseconds, comprehension session (standard deviations in parentheses).				
	SRC-plausible	SRC-implausible	ORC-plausible	ORC-implausible
RC region	2033.47	2452.09	2361.82	2284
	(659.75)	(1226.82)	(1089.98)	(917.57)
Matrix verb	546.71	819.89	706.71	976.14
	(235.60)	(457.59)	(382.86)	(600.19)
Post-verb region	1537.99	1868.33	1567.33	1935.40
	(502.82)	(858.95)	(489.23)	(718.02)
(b) Grand mean reading times in milliseconds, translation session (standard deviations in parentheses).				
RC region	2698.22	3154.42	2661.69	3336.81
	(1430.74)	(1682.74)	(1418.05)	(1717.57)
Matrix verb	605.69	978.16	817.36	1038.42
	(319.27)	(554.87)	(405.36)	(635.46)
Post-verb region	2681.10	3448.68	3079.91	3523.99
	(1221.04)	(1378.41)	(1349.13)	(1662.42)

RC = relative clause; SRC = subject relative clause; ORC = object relative clause

in SRCs, whereas the reading times were similar between the two conditions for higher-proficiency learners. At the verb region, the structure ($t = 2.57$, $p = .01$) and the plausibility ($t = 5.80$, $p < .01$) effects were significant, showing that reading behavior at this region in the translation task was similar to the comprehension session. No interaction between structure and plausibility was observed. In the remainder of the sentence, plausibility yielded a significant effect ($t = 4.48$, $p < .01$), and structure yielded marginal significance ($t = 1.80$, $p = .07$). Proficiency had no relation to either structure or plausibility in the post-verb region.

In summary, L2 speakers' reading patterns were quite similar to the native speakers as both groups were sensitive to structure and plausibility at the verb region in both reading sessions. This finding suggests that L2 learners were capable of processing structural information along with semantic information during online comprehension similar to native speakers. However, while native speakers were affected by structural differences at the early region of the sentences, i.e., the RC region, L2 learners were influenced by plausibility information in both reading sessions. This difference between native and L2 speakers would appear to be a qualitative one. Proficiency and task were also related to the use of information in L2 processing. More specifically, the low-proficiency learners were more likely than high-proficiency learners to focus on structural information when the reading goal was to translate, yielding a significant difference in using syntactic information depending on reading goals. In addition, the translation task appeared to encourage L2 learners to be attentive to structural information to a greater extent even in the post-verb region, in the same way as native speakers.

Accuracy data

Native speakers

Accuracy data were analyzed using logit mixed models, which handle binary distributions, such as accuracy data, better than ANOVA (Jaeger, 2008). Models were fitted using a stepwise selection procedure in which factors failing to reach at least $p < .10$ significance were removed from the model (Luke & Christianson, 2011). Effects of structure, plausibility and the interaction of structure and plausibility were included as predictors in the final, best fitted model. Subject and item were also included as crossed random effects. Figure 2 illustrates the mean proportions of accuracy in native speakers, and Table 6 shows the results of the final model. (Although logit mixed model analyses are performed over raw data, Figure 2 plots means for ease of exposition.)

As illustrated in Figure 2, native speakers were significantly less accurate in the ORC condition than the SRC condition ($z = –5.26$, $p < .01$). The effect of plausibility was also significant ($z = –5.29$, $p < .01$), showing that the implausible conditions yielded more inaccurate answers than the plausible conditions. Both effects were qualified by a reliable interaction ($z = –2.41$, $p = .02$), such that the native speakers had the most

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

Table 5. *Fixed effects in the linear mixed effects model of reading times, L2 speakers.*

	Predictor	Estimate	SE	*t*-value	*p*
Comprehension session					
RC region ($R^2 = 0.42$)	Intercept	2276.47	163.251	13.94	<.01
	Structure	48.33	39.32	1.23	.22
	Plausibility	78.82	39.19	2.01	.04*
	Structure × Plausibility	–127.70	37.40	–3.42	<.01*
Verb region ($R^2 = 0.32$)	Intercept	775.38	70.03	11.08	<.01
	Structure	92.65	23.52	3.94	<.01*
	Plausibility	148.26	23.43	6.33	<.01*
	Proficiency	16.91	20.09	0.84	.40
	Plausibility × Proficiency	13.09	7.50	1.75	.08†
Post-verb region ($R^2 = 0.29$)	Intercept	1734.30	122.58	14.15	<.01
	Structure	23.60	39.54	0.60	.55
	Plausibility	182.78	39.42	4.64	<.01*
	Proficiency	–3.09	28.60	–0.11	.92
	Plausibility × Proficiency	25.03	12.54	2.00	.04*
Translation session					
RC region ($R^2 = 0.40$)	Intercept	2972.91	273.78	10.86	<.01
	Structure	29.27	47.25	0.62	.54
	Plausibility	288.97	47.25	6.12	<.01*
	Proficiency	38.36	82.63	0.46	.64
	Structure × Proficiency	–27.29	15.51	–1.82	.06†
Verb region ($R^2 = 0.28$)	Intercept	861.37	73.11	11.78	<.01
	Structure	67.46	26.23	2.57	.01*
	Plausibility	152.23	26.23	5.80	<.01*
	Proficiency	5.53	20.69	0.23	.78
	Plausibility × Proficiency	14.58	8.59	1.70	.09†
Post-verb region ($R^2 = 0.34$)	Intercept	3183.42	230.00	13.84	<.01
	Structure	118.53	67.66	1.80	.07†
	Plausibility	302.91	67.66	4.48	<.01*
	Structure × Plausibility	–80.88	67.66	–1.20	.23

*$p < .05$; †$p < .10$; RC = relative clause; SE = standard error

Table 6. *Fixed effects in the logit mixed model of question accuracy in native speakers.*

	Question accuracy (C = 4.33)			
Predictor	Coefficient	SE	Wald *Z*	*p*
Intercept	1.66	0.16	10.03	<.00
Structure	−0.37	0.07	−5.26	<.01*
Plausibility	−0.37	0.07	−5.29	<.01*
Structure × Plausibility	−0.17	0.07	−2.41	.02*

*$p < .05$; †$p < .10$; SE = standard error

difficulty in the ORC-implausible condition. This result replicates Ferreira (2003), but with RC constructions rather than active/passive constructions.

L2 speakers

As with the native speakers, accuracy data for L2 speakers were analyzed using a logit mixed model. Identical to the native speakers, effects of structure, plausibility and the interaction of structure and plausibility were included as predictors in a final, best-fitted model. Subject and item were also included as crossed random effects. In both reading sessions, proficiency was also added as

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

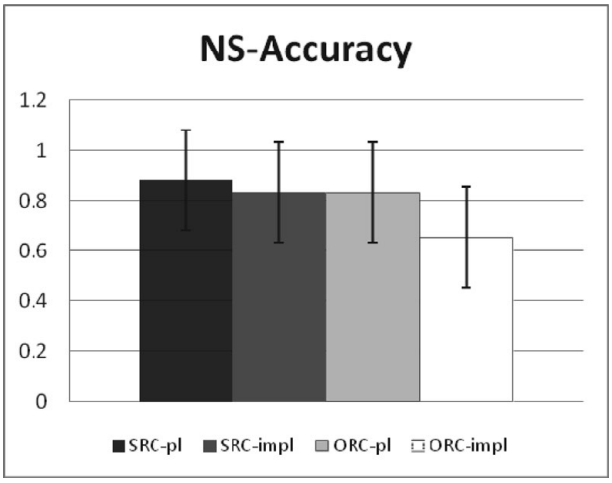


Figure 2. Mean proportion of comprehension accuracy in native speakers. SRC-pl = subject relative clause, plausible; SRC-impl = subject relative clause, implausible; ORC-pl = object relative clause, plausible; ORC-impl = object relative clause, implausible

Table 7. *Fixed effects in the logit mixed model of question accuracy in L2 learners for both reading sessions.*

Comprehension session				
	Question accuracy (C = 4.00)			
Predictors	Coefficient	SE	Wald *Z*	*p*
Intercept	1.41	0.14	10.29	<.00
Structure	–0.17	0.09	–1.83	.06†
Plausibility	–0.71	0.09	–7.69	<.01*
Proficiency	0.09	0.04	2.18	.03*
Structure × Plausibility	–0.01	0.08	–0.01	.95
Translation session				
	Question accuracy (C = 3.90)			
Predictors	Coefficient	SE	Wald *Z*	*p*
Intercept	1.46	0.20	7.18	<.00
Structure	–0.36	0.10	–3.51	<.01*
Plausibility	–0.97	0.10	–9.29	<.01*
Proficiency	0.13	0.06	2.34	.02*
Plausibility × Proficiency	–0.06	0.03	–1.78	.07†

*$p < .05$; †$p < .10$; SE = standard error

a continuous variable in the model (see Table 7). The mean proportions of accuracy in L2 learners for both reading sessions are illustrated in Figure 3 and the results of the final, best-fitting models are presented in Table 7.

As can be seen in Figure 3, there was no numerical difference between the reading sessions in terms of accuracy. In both reading sessions, the main effects of both structure and plausibility were reliable with no interactions (see Table 7), indicating that L2 learners had more difficulty in understanding ORCs and implausible sentences, independent of other factors. Also, proficiency yielded significant effects ($z = 2.18$, $p = .03$ for comprehension; $z = 2.34$, $p = .02$ for translation) in both reading sessions, indicating that the higher the proficiency was, the more accurate the answers. In the translation session, proficiency marginally interacted with plausibility ($z = -1.78$, $p = .07$), indicating that as proficiency increased, accuracy in the plausible conditions increased compared to the implausible conditions,

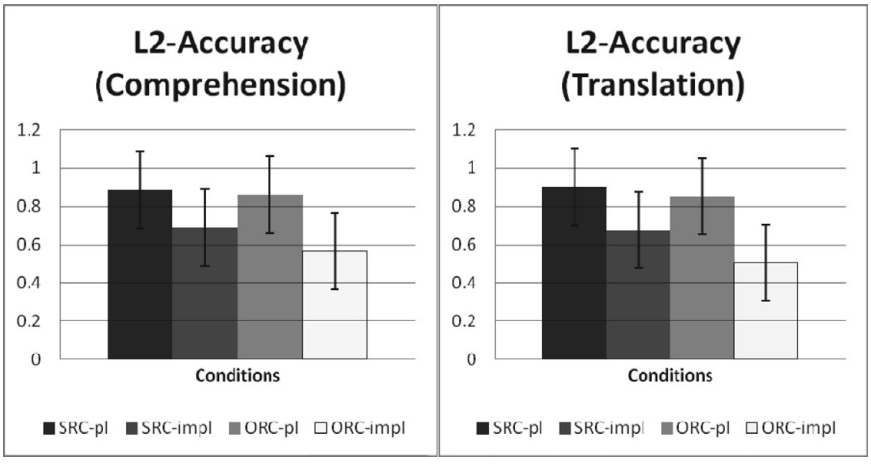


Figure 3. Mean proportions of comprehension accuracy in L2 learners for the two reading sessions.

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

Table 8. *The numbers of translation in each condition.*

	Numbers of translation			
Condition	SRC_correct	SRC_incorrect	ORC_correct	ORC_incorrect
SRC-plausible	196	1	6	0
SRC-implausible	103	21	4	27
ORC-plausibhle	7	4	176	3
ORC-implausible	10	75	74	33

SRC = subject relative clause; ORC = object relative clause

whereas lower-proficiency learners did not show this significant difference depending on plausibility.

Translation data

Translations were another dependent variable for the L2 group. Participants' translations were transcribed by the experimenter and coded as SRC-correct, SRC-incorrect, ORC-correct, and ORC-incorrect (where 'correct' means thematic roles were correctly assigned and 'incorrect' means thematic roles were translated in a reversed order from the original). There were inappropriate translations where participants did not produce any SRC or ORC structures, or they did not comprehend the input. These outliers ($n = 124$) were about 14% of the total 864 translations and were removed from the further analyses. To be more specific, 6% in the SRC-plausible condition, 28% in SRC-implausible condition, 12% in ORC-plausible condition, and 11% in ORC-implausible condition were removed. A Poisson Regression model with repeated measures (using a Generalized Estimating Equations (GEE) model in SAS software; see Agresti, 2007) was used to examine the effects of morphosyntax and semantic knowledge on translation performance as the translation data were multi-categorical (Agresti, 2007).

Table 8 presents a contingency table for the numbers of translations produced in each condition. As can be seen in the contingency table, Korean translations generally reflected the structure of the to-be-translated sentences and implausible sentences resulted in an increased probability of incorrect translations (with thematic roles reversed).

A Poisson loglinear regression showed that both structure ($z = 39.99$, $p < .01$) and plausibility ($z = 6.37$, $p < .01$) yielded main effects in L2 learners' translations. To be specific, expected odds of the correct translation when the structure was SRC were 2.75 times larger than when the structure was ORC. The plausibility effect was 1.61 times larger when the event was plausible than implausible. In particular, the cell ($n = 33$), where participants produced ORC-incorrect translations in the ORC-implausible condition, provides strong evidence that L2 learners exhibited command of syntactic knowledge (providing an accurate morphosyntactic Korean translation of a non-canonical, more difficult English structure), but that the semantic information overrode the fragile syntactic representation, resulting in misinterpretation. There was also a significant interaction between structure and plausibility ($z = -7.10$, $p < .01$), indicating the probability of more incorrect translations for ORC-implausible sentences than for any other sentence type. Finally, low-proficiency learners produced more incorrect translations in the ORC-implausible condition than high-proficiency learners did (an average of 72% errors vs. an average of 43% errors, respectively). This result is expected within an account in which syntactic structure becomes especially fragile when it conflicts with semantic information, as in the GE processing account.

Discussion

In this study, we investigated L2 learner syntactic and semantic processing compared to native speaker syntactic and semantic processing using both a self-paced reading task and a translation task. Subject relative clause and object relative clause structures were utilized with a manipulation of plausibility in order to examine the use of hypothesized syntactic and semantic/heuristic processing routes during online comprehension. Another question asked was how reading goals, or task, affect L2 reading patterns, by asking Korean learners of English to read sentences either for comprehension or for translation in the other session. L2 proficiency was also explored as a possible factor. It was predicted that L2 learners' reading times would be modulated by structure type, if they are able to access to syntactic information online during reading. Also, comprehension accuracy was predicted to be lower when non-canonical structure was combined with implausible events, if the interpretation derived from the semantic route overrides the interpretation derived from the syntactic route (cf. Ferreira, 2003). Higher L2 proficiency learners were predicted to show

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

greater sensitivity to structural information than lower-proficiency learners. In addition, the translation task was expected to yield longer reading times than when answering comprehension questions, and the translation output was also predicted to yield explicit evidence of how L2 learners comprehended the linguistic input, specifically, evidence of how, or if, syntactic and semantic information had been integrated.

Results from the native speakers in the present study were in line with previous L1 studies examining relatively easier processing of SRCs than ORCs (Holmes & O'Regan, 1981; King & Just, 1991), in that native participants showed longer reading times in the ORC conditions than the SRC conditions early in the RC region. At both the verb and the post-verb regions, native speakers were influenced by both structural and plausibility information, independently. The result that implausible conditions made processing harder independently of structure is consistent with the claim that language comprehension proceeds along two routes: a morphosyntactic and a semantic route (Christianson et al., 2010; Kuperberg, 2007). The comprehension accuracy results further strengthen the argument that both processing routes are indeed at work. For the ORC conditions, native speakers' accuracy was lower overall (76%), and lower for implausible sentences than plausible sentences (65% vs. 83%). This result fits with findings from previous studies proposing the Good-Enough processing account (Christianson et al., 2001, 2010). The data here are also noteworthy as they replicate the findings of Ferreira (2003) using subject/object relative clauses, instead of active/passive structures.

Unlike the early sensitivity to structural differences in native speakers, L2 speakers did not exhibit the same sensitivity to structural information in the RC region; instead they showed a plausibility effect in both reading sessions. In addition, the degree of sensitivity differed depending on reading goal and proficiency. Specifically, in the comprehension session, L2 speakers, regardless of proficiency, showed an interaction between structure and plausibility, indicating a significant slowdown for implausible sentences only in the SRC conditions. In the translation session, on the other hand, lower-proficiency learners were more attentive to structural information in the RC region, yielding longer reading times in the ORC than the SRC conditions. The interaction of structural and semantic information in the comprehension session is in line with the previous L1 and L2 studies that show an immediate interaction, which has been interpreted as plausibility information influencing syntactic analysis or postponing syntactic assignment (Jackson & Roberts, 2010; Traxler et al., 2002). However, the current study found the interaction in the L2 comprehension session at the RC region only; neither in native speakers nor in the translation session by L2 speakers did we observe an interaction between the two factors, unlike previous studies. We speculate that this is because of proficiency, task, or possibly other uncontrolled factors, since task apparently affects the timing of the interaction.

Furthermore, these observations have several implications. First, the results demonstrate that reading goals affected the sorts of information L2 learners computed during online reading from the beginning of the sentence. Second, these results hint at the cognitive processes involved in translation. The fact that the lower-proficiency participants had longer reading times in the ORC sentences in the translation session suggests that translation increases low-proficiency learners' attention to the form due to having to reformulate linguistic components of the source language to those of the target language during online reading (HORIZONTAL PERSPECTIVE OF TRANSLATION; Danks & Griffins, 1997; Gerver, 1976). Lastly, L2 learners seem to be dependent on semantic/plausibility information to a larger extent early in sentences compared to native speakers, given that the native speakers showed structure effects first at the RC region. This pattern of results is quite consistent with the previous studies that found L2 learners' large reliance on semantic information (Felser et al. 2003; Frenck-Mestre & Pynte, 1997; Juffs, 1998; Papadopoulou & Clahsen, 2003; Williams, Mobius & Kim, 2001). The reliance on semantic knowledge is also very intuitive, in that the L2 learners are in the process of learning the structure of English, but have already acquired knowledge about how the world works. Pinker (1984) hypothesized that infants use their world knowledge to recognize syntactic relationships within sentences before they acquire all the syntax, which is called SEMANTIC BOOTSTRAPPING. Similarly, in the current study, world knowledge appears to be a good way for L2 learners to bootstrap the syntax of the target language (Fisher, Hall, Rakowitz & Gleitman, 1994; Pinker, 1984). This large reliance on semantic information by L2 speakers in the first place can also be found in the translation data where L2 participants yielded SRC-incorrect sentences ($n = 75$) in the ORC-implausible condition. These responses suggest that L2 speakers may rely on world knowledge rather than computing full syntactic representations of the target sentences before they integrate syntactic outputs with semantic/heuristic outputs.

At the verb region, however, L2 speakers were influenced by both structure and plausibility information in the two reading sessions, which was identical to the native speakers. This suggests that L2 speakers were able to compute both syntactic and semantic/world knowledge during reading in the same manner as native speakers, which reflects the concept of parallel routes towards comprehension. The fact that the L2 learners showed

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

different reading times according to structure types is in line with studies showing that L2 learners are able to utilize morphosyntactic information during processing (Hopp, 2006; Jackson, 2008; Jackson & Dussias, 2009), and it does not seem that the syntactic representation they compute during reading is less detailed than that of native speakers. In addition, the fact that L2 participants were not less detailed in the verb region, unlike the RC region, suggests that L2 speakers are just slower in retrieving syntactic information than native speakers are, but that they are eventually able to access the syntax similarly to native speakers. In the remainder of the sentence, L2 learners' reading patterns were different from native speakers in the comprehension session. Plausibility information lingered in L2 learners' minds longer than structural information, which is again consistent with previous findings in the L2 processing literature that language learners are more sensitive to lexical-semantic or world knowledge than native speakers (Felser et al. 2003; Frenck-Mestre & Pynte, 1997; Juffs, 1998; Papadopoulou & Clahsen, 2003; Williams et al., 2001). However, in the translation session, L2 learners showed both structure (although marginal) and plausibility effects in the post-verb region, similar to native speakers. This suggests that the additional translation task keyed L2 learners' access to syntactic knowledge during online reading more so than just reading for comprehension. It is possible that previous studies demonstrating that L2 learners are insensitive to morphosyntactic input during online simply did not use a task that would have motivated L2 learners to process in a more native-like manner.

With regard to the offline comprehension questions, L2 learners' accuracy results were slightly different from those of native speakers in that the L2 participants did not show an interaction of structure and plausibility on the paraphrase verification task. In other words, ORC sentences produced less accurate answers than SRC sentences regardless of plausibility, and implausible sentences yielded less accurate answers than plausible ones independent of structure. The reason for the absence of an interaction appears to lie in the fact that L2 learners failed to reach the accuracy level of the native speakers in the SRC-implausible condition. To be specific, accuracy proportions in the other three conditions (SRC-plausible, ORC-plausible, ORC-implausible) were similar between native speakers and L2 learners (see Figures 2 and 3 above). However, the difference was large between the two groups in the SRC-implausible condition (83% for native speakers vs. 69% for L2 learners), which resulted in no interaction in the L2 group. This discrepancy is interpreted as showing that L2 learners rely on world knowledge in sentence processing more than native speakers, such that L2 speakers interpreted the implausible sentences based on how the world usually works. The data can be explained well by both the SSH and the GE account, as the SSH suggests that L2 learners depend necessarily on lexical-semantic or world knowledge during sentence processing, and the GE account predicts that semantic representations will sometimes override syntactic representations, presumably to a greater extent in L2 populations.

Another facet of this observation is that L2 learners displayed a similar pattern of accuracy to native speakers in the other three conditions. As with native speakers, L2 learners attempted to comprehend the sentences through both a morphosyntactic and a heuristic route. Consistent with the GE processing, when the two outputs from each route conflict, especially in the cases where implausible events were presented in non-canonical structures – the ORC-implausible condition in this experiment – the output from the syntactic route seemed to be overridden by the output from the semantic/heuristic route, ultimately increasing misinterpretations.

The argument that L2 learners processed the input through two routes but the lack of integration yielded the misinterpretations is augmented by the translation data in this study. Translation data from L2 speakers provide explicit evidence to show that, at least some of the time, L2 participants actually computed the full syntactic analysis, but the syntactic representation was overridden when the sentence meaning was implausible. The occurrence of incorrect translations where L2 learners produced accurate structural translations with reversed thematic roles in the SRC-implausible condition ($n = 21$) and the ORC-implausible condition ($n = 33$) are the strongest evidence that the syntactic analysis was detailed and complete, but the implausible propositions overrode the syntactic parser's analysis, resulting in misinterpretations. This result agrees with the comprehension behavior of native speakers reported in previous L1 processing studies (Christianson et al., 2010; Ferreira, 2003). In addition, most of the outliers in the SRC-implausible condition included 'passive' structure with reversible thematic roles to make the sentence more sensible; for example, they usually translated the original input "The mouse that chased the cat was fast" into "The mouse was chased by the cat". Although this observation cannot be statistically tested, it is speculated that this behavior represents L2 speakers' both syntactic and semantic route at work during comprehension, resulting in passive structure used to produce a translation that accords with world knowledge.

Most of the studies in L2 processing literature to date have derived their conclusions exclusively from online reading time data (Felser et al., 2003; Hopp, 2006; Jackson, 2008; Jackson & Dussias, 2009; Papadopoulou & Clahsen, 2003). As pointed out in Christianson, Luke, and Wochna (2012), online reading times do not always reflect how readers interpret sentences. As such, the fact that L2 learners show different reading times from native

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

speakers does not necessarily imply that the syntactic information they compute is shallower or less detailed. The translation data in this study are unique in this respect because they represent L2 speakers' final interpretations, demonstrating both syntactic and semantic knowledge that is computed. It is also worth noting that these data are the first to demonstrate a misinterpretation effect of unambiguous RC structures.

We now turn to the second question probed in this study, namely, the roles of task and proficiency in L2 processing. Based on L2 learners' reading times, both L2 proficiency and reading goals affected the depth of processing during online reading. It was predicted that the translation task might increase awareness of either semantic or syntactic information to a greater extent than normal comprehension due to the possible task demands associated with translation. Reading times revealed that translation generally increased L2 speakers' focus on structural information from the beginning of the sentence to the latter part of the sentence. Especially for lower-proficiency speakers, their reading times were inflated in ORCs compared to SRCs during the translation session compared to the comprehension session, whereas there was no significant difference between structures in higher-proficiency learners. This difference suggests that L2 learners' reading strategies vary with both proficiency and, importantly here, task. This variability is expected under the the GE processing view. The difference is consistent with data from an eye-tracking study examining reading for comprehension vs. translation (Lim & Christianson, 2012. These results deserve further investigation, especially in connection with cognitive theories of translation.

As for the accuracy results, it was surprising that accuracy rates were almost identical between the two reading sessions in L2 group. Accuracy rates were expected to be lower in the comprehension session. A possible explanation for this result would be that people comprehended sentences essentially the same way while reading for translation as reading for comprehension. Indeed, this similar accuracy pattern between the two reading sessions can even be taken to validate translation as a method to examine online processing, as the increased cognitive load did not alter comprehension levels in either direction.

Accuracy in both reading sessions was significantly modulated by proficiency level, with accuracy positively associated with proficiency. This finding suggests that L2 proficiency plays a crucial role in L2 processing mechanisms. All together, these results indicate the task-dependent nature of processing, which is explicitly predicted within the GE processing view (Ferreira & Patson, 2007; Swets et al., 2008). Indeed, it seems difficult to claim that L2 speakers do not have access to detailed morphosyntactic representations in some tasks, but do in others.

Taken all together, the data presented here are quite consistent with the GE processing account (Figure 1b above), although some of the results are also certainly consistent with the SSH (Figure 1a). In the GE framework, the critical difference between L1 and L2 processing lies not in the availability of the syntactic processing (assuming underlying syntactic knowledge is sufficient), but rather in the ability of L2 speakers to integrate the outputs of syntactic and semantic/heuristic routes to arrive at accurate comprehension. The results reported here clearly support this view: The interaction between syntax and semantics observed in the translation data patterned with native speakers, and the completeness of this integration, as did the online reading times. Furthermore, these patterns were modulated by proficiency and, importantly, task.

In conclusion, the present study investigated L2 learners' reading behaviors compared to native speakers, along with the possible influence of reading goals and L2 proficiency on reading behaviors of Korean speakers of English. Although L2 learners had a greater tendency to rely on semantic or world knowledge than native speakers, both online and offline data showed that the L2 learners were able to use both syntactic knowledge and plausibility information during real-time comprehension. Misinterpretations occurred when the outputs from morphosyntactic and semantic processing routes were inconsistent, as previously observed in native-speaker processing. Results also showed that translation leads L2 learners, especially low-proficiency learners, to allocate more attention to structural sources, suggesting reading goals and L2 proficiency play important roles in L2 processing. All together, these results suggest that L2 processing is quantitatively rather than qualitatively different from L1 processing, i.e., sometimes "good enough", just as with native speakers. Furthermore, the translation paradigm used in this study, which provided explicit output from language learners, appears to be a useful window into understanding language learners' final interpretations and integration of morphosyntactic and semantic information in L2 processing research. The comparisons of processes involved in reading for comprehension and translation are also informative about the nature of translation itself. Translation appears to involve deeper analysis of linguistic input with more attention to structure of the input, especially for lower-proficiency learners.

Appendix. Stimuli

The sentences below are only presented in subject relative clause and plausible version. The implausible version is created with reversing the nouns. The object relative

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

clause is created as follows: e.g., *The dog that bit the man was in the yard → The man that the dog bit was in the yard*.

1. The dog that bit the man was in the yard.
2. The chef that ruined the food was in the kitchen.
3. The bird that ate the worm was small.
4. The lawyer that sued the doctor was smart.
5. The soldier that protected the villager was brave.
6. The cat that chased the mouse was fast.
7. The teacher that quizzed the student was in the classroom.
8. The cop that pursued the thief was driving a car.
9. The waiter that served the customer was tall.
10. The owner that fed the cat was sitting on a sofa.
11. The detective that investigated the suspect was very tired.
12. The doctor that treated the patient was female.
13. The politician that deceived the voter was Korean.
14. The golfer that hit the ball was in the shade.
15. The hunter that shot the deer was in the Rocky Mountains.
16. The frog that ate the fly was green.
17. The ghost that scared the boy was hiding behind a curtain.
18. The police officer that arrested the citizen was handsome.
19. The fisherman that caught the fish was in the middle of the ocean.
20. The man that walked the dog was in the park.
21. The coach that scolded the player won the championship twice.
22. The mother that bathed the child smelled nice.
23. The father that scolded the teenager was in the living room.
24. The guide that led the tourist liked Europe a lot.
25. The reporter that interviewed the actress was at the coffee shop.
26. The kids that obeyed the teacher enjoyed the summer break.
27. The grandmother that dressed the child had a beautiful smile.
28. The volunteer that helped the blind man was very handsome.
29. The homeowner that paid the gardener loved the garden.
30. The wrecker that towed the car drove at a high speed.
31. The boy that petted the puppy had cute eyes.
32. The bird that protected the egg was in the big tree.
33. The secretary that assisted the President was in a conference room.
34. The tutor that taught the student solved the math problem.
35. The boss that fired the worker was unhappy.
36. The parent that raised the twins lived in New York.
37. The terrorist that held the hostage was located in the building.
38. The criminal that kidnapped the girl was on CNN news.
39. The conductor that led the orchestra was pleased with the performance.
40. The consultant that advised the client was very clever.
41. The lawyer that defended the client was worried about the result.
42. The parents that punished the child went to church.
43. The fan that cheered for the baseball player wore a red shirt.
44. The nuns that took care of the orphans live in a small village.
45. The dean that awarded the student a prize was very famous in the school.
46. The dog that herded the sheep was very furry.
47. The fan that admired the Hollywood star loved to wear big sunglasses.
48. The guard that locked up the prisoner regretted past decisions.

References

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd edn.). Hoboken, NJ: John Wiley & Sons.

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 Langage, 59,* 390–412.

Bates, E., & MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (ed.), *Mechanisms of language acquisition*, pp. 157–193. Hillsdale, NJ: Lawrence Erlbaum.

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

Buck, G. (1992). Translation as a language testing procedure: Does it work? *Language Testing, 9 (2),* 123–148.

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

Cho, S. (1999). The acquisition of relative clauses: Experimental studies on Korean. Ph.D. dissertation, University of Hawaii at Manoa.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles ssigned 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 (2),* 538–544.

Christianson, K., Luke, S. G., & Wochna, K. (2012). Garden-path vs. local coherence structures: Similarities and differences in parsing and interpretation. Ms., University of Illinois at Urbana–Champaign.

Christianson, K., Williams, C., Zacks, R., & Ferreira, F. (2006). Younger and older adults' good enough interpretations of garden-path sentences. *Discourse Processes, 42,* 205–238.

Clahsen, H., & Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics, 27,* 3–42.

Clahsen, H., & Felser, C. (2006b). Continuity and shallow structures in language processing. *Applied Psycholinguistics, 27,* 107–126.

Clifton, C. Jr., & Frazier, L. (1989). Comprehending sentences with long distance dependencies. In G. Carlson & M. Tanenhaus (eds.), *Linguistic structure in language processing*, pp. 273–317. Dordrecht: Kluwer.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12,* 769–786.

Cordero, A. D. (1984). The role of translation in second language acquisition. *The French Review, 57 (3),* 350–355.

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 interpreting*, pp. 161–175. Thousand Oaks, CA: Sage.

Dussias, P. E. (2001). Sentence parsing in fluent Spanish–English bilinguals. In J. Nicol (ed.), *One mind, two languages: Bilingual language processing*, pp. 159–176. Malden, MA: Blackwell.

Felser, C., & Roberts, L. (2007). Processing *wh*-dependencies in a second language: A cross-modal priming study. *Second Language Research, 23 (1),* 9–36.

Felser, C., Roberts, L., Gross, R., & Marinis, T. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics, 24,* 453–489.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47,* 164–203.

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 (1),* 3–20.

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

Fisher, C., Hall, D., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua, 92,* 333–375.

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

Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology, 50A,* 119–148.

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.

Gibson, E., Hickok, G., & Schütze, C. T. (1994). Processing empty categories: A parallel approach. *Journal of Psycholinguistic Research, 23,* 381–405.

Gregg, K. (1996). The logical and developmental problems of second language acquisition. In C. R. William & T. K. Bhatia (eds.), *Handbook of second language acquisition*, pp. 50–84. New York: Academic Press.

Holmes, V. M., & O'Regan, J. K. (1981). Eye fixation patterns during the reading of relative clause sentences. *Journal of Verbal Learning and Verbal Behavior, 20,* 417–430.

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 (4),* 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.

Jackson, C., & Dussias, P. E. (2009). Cross-linguistic differences and their impact on L2 sentence processing. *Bilingualism: Language and Cognition, 12 (1),* 65–82.

Jackson, C., & Roberts, L. (2010). Animacy affects the processing of subject–object ambiguities in the second language: Evidence from self-paced reading with German and second language learners of Dutch. *Applied Psycholinguistics, 31,* 671–691.

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.

Juffs, A. (1998). Main verb versus reduced relative clause ambiguity resolution in second language sentence processing. *Language Learning, 48,* 107–147.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory capacity. *Psychological Review, 99,* 122–149.

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press

King, J., & Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30,* 580–602.

Kim, J.-H. (2008). Working memory effects on bilingual sentence processing. Ph.D. dissertation, University of Illinois at Urbana–Champaign.

Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research* (Special Issue), *1146,* 23–49.

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: John Benjamins.

Lim, J.-H., & Christianson, K. (2012). L2 morphological sensitivity in comprehension and translation: Evidence from eye-movements. Ms., University of Illinois at Urbana–Champaign.

Luke, S. G., & Christianson, K. (2011). Stem and whole-word frequency effects in the processing of inflected verbs in and out of a sentence context. *Language and Cognitive Processes, 26,* 1173–1192.

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.

Malakoff, M. E. (1992). Translation ability: A natural bilingual and metalinguistic skill. In R. J. Harris (ed.), *Cognitive processing in bilinguals* (vol. 83). pp. 515–529. New York: Elsevier.

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27,* 53–78.

Nicol, J., & Pickering, M. J. (1993). Processing syntactically ambiguous sentences: Evidence from semantic priming. *Journal of Psycholinguistic Research, 22,* 207–237.

O'Grady, W., Lee, M., & Choo, M. (2003). A subject–object asymmetry in the acquisition of relative clauses in Korean as a second language. *Studies in Second Language Acquisition, 25,* 433–448.

Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition, 24,* 501–528.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

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

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Undersepcification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition, 36 (1),* 201–216.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language, 47,* 69–90.

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

Williams, J. N., Mobius, P., & Kim, C.-K. (2001). Native and non-native processing of English *wh*-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics, 22 (4),* 509–540.

https://doi.org/10.1017/S1366728912000351 Published online by Cambridge University Press