Real-Time Reading and Reactivation Evidence of Syntactic Gap Processing in Japanese Scrambling

Masahiro Hara
Truman State University

1. Introduction

This study argues that second language (L2) learners are capable of syntactic gap processing under moderate computational demand, but not when the computational load is excessively taxing. The evidence comes from 1) measurement of real-time reading and 2) probe response data from advanced Korean and English learners' processing of sentences with scrambling in Japanese. The findings are consistent with a working hypothesis that learners' syntactic gap processing is severely constrained by limitations on cognitive resources available during L2 processing. The findings contradict the view that late (adult) learners' gap processing is verb-driven, that is, that their gap processing utilizes information on verb argument structure and pragmatic knowledge (Clahsen & Felser, 2006a).

2. L2 studies on processing of filler–gap dependencies

Recent L2 studies (Felser & Roberts, 2007; Marinis, Roberts, Felser, & Clahsen, 2005) have found no evidence of learners' syntactic gap processing when effects stemming from syntactic gap processing were dissociated from effects arising from verb-driven processing.[1] As argued below, it is possible that those null effect findings arose from two confounding factors: the relatively slow speed of L2 processing and the relatively limited capacity for computational load in L2 processing, both of which bear on learners' syntactic gap processing capability.

2.1. Slow L2 processing

There is general consensus that L2 processing is slower than first language (L1) processing (cf. Segalowitz & Hulstijn, 2005). In that light, it is worth reexamining Felser and Roberts's (2007) finding of no evidence of L2 syntactic gap processing. The authors conducted a cross-modal picture priming experiment in which Greek learners of English had to make an "alive" or "not alive" decision on picture targets (animals or objects) presented on a computer screen while listening to indirect object-relativized sentences such as *Fred chased the squirrel to which the nice monkey explained the game's difficult rules___ in the class last Wednesday*. Learners' faster responses to pictures of antecedents (a picture of a squirrel) than to those of unrelated entities (a picture of a toothbrush) were observed regardless of whether the picture targets were presented at the pre-gap position (immediately after

* The study reported in this article would have been impossible without generous support for data collections by Soichi Aoki, Yoshikazu Kawaguchi, Mina Kobayashi, Mamiko Orii, and Keichi Tajima. Special thanks go to Laurent Dekydtspotter for his various comments on earlier versions of the study. I thank two anonymous reviewers as well as Luke Plonsky and Maren Schierloh for their perceptive suggestions on the earlier drafts of the article. This project was partially supported by the *Language Learning* Small Grants Research Program and the Truman Faculty Summer Research Fellowship. I received unbounded support for this project from Andrew Hara and Katherine Hara.

[1] Earlier L2 gap processing studies (Juffs, 2005; Juffs & Harrington, 1995, 1996; Williams, 2006; Williams, Mobius, & Kim, 2001) found learners' reading slowdown at the immediately post-verbal region in processing of *wh* filler–gap dependencies. As Clahsen and Felser (2006b) have observed, those studies' findings of learners' reading slowdown at the immediately post-verbal region could have resulted from application of syntactic gap-filling operations or from use of verb argument structure information to associate the displaced element with the verb.

game's) or at the gap position (immediately after *rules*). The authors therefore concluded that learners had maintained activation of the antecedent in working memory (WM), but had not reactivated it at the gap position.

Although Felser and Roberts (2007) found no evidence for an immediate reactivation effect, the possibility of a delayed reactivation effect cannot be ruled out because the picture priming timing was set to measure the reactivation effect only at the earliest point where it could occur. In their study, Felser and Roberts used the experimental design and materials developed by Roberts, Marinis, Felser, and Clahsen (2007) for an investigation of child and adult native speakers' antecedent priming effects. Importantly, no antecedent reactivation effects were found in the case of low WM-span adult participants in the L1 processing study. Yet Roberts et al. felt that perhaps those adult native speakers had needed more time to integrate the dislocated constituent; they also felt that the low WM-span adult native speakers might have reactivated the antecedent "at the word following the gap position or even at the end of the sentence" (Roberts et al., 2007, p. 184). Considering that L2 processing is slower than L1 processing, the same scenario is plausible for the Greek learners of English in Felser and Roberts' study.

2.2. Limited capacity for computational load in L2 processing

Marinis et al. (2005) examined learners' ability to postulate an intermediate gap in computing long-distance *wh*-dependencies in reading very complex sentences in English. In that self-paced reading study, learners from *wh*-movement and *wh*-in-situ first language (L1) backgrounds read intermediate gap-sentences such as *The nurse who the doctor argues___ that the rude patient had angered___ is refusing to work late*, along with sentences containing no intermediate gap, as in *The nurse who the doctor's argument about the rude patient had angered___ is refusing to work late* (cf. Gibson & Warren, 2004). Marinis et al. found that learners' reading at the region (*had angered*) implicating the second gap in intermediate gap-sentences was similar to their reading at the corresponding region in sentences without an intermediate gap. The authors therefore concluded that learners had failed to postulate an intermediate gap.

The design of Marinis et al.'s (2005) study may have unintentionally imposed an excessively high computational load on learners (see Caplan & Waters, 1999, p. 80; Indefrey, 2006) as they read intermediate gap-sentences containing three clauses and two layers of nesting in a segment-by-segment moving window format. Suppose learners had postulated an intermediate gap: By the time they reached the region implicating the second gap, any effect arising from having processed the intermediate gap might have disappeared under the high computational demand such sentences placed on them.

The two experiments reported in the present study utilized scrambling in Japanese. Comprehension of that construction requires establishing filler–gap dependencies (see Miyamoto, 2006, for an overview of L1 processing of Japanese scrambling, and Nemoto, 1999, for a linguistic overview).[2] A guiding hypothesis was that learners' syntactic gap processing is severely constrained

[2] Use of Japanese, a verb-final language, as a target language enables distinguishing structure-based from verb-driven processing in computing filler–gap dependencies because syntactic gap processing occurs preverbally. Phillips and Wagers (2007, p. 749) suggest that if, via case-marking information, the Japanese-parser allows a verb position to be constructed prior to the arrival of the verb in the input string, a direct filler–verb association may be made in advance of the verb. If so, then use of a verb-final language in gap processing studies would encounter an interpretive indeterminacy problem (see footnote 1). I believe it is unwarranted to attribute such a high degree of predictive power to the Japanese-parser. As a general-purpose parsing strategy, Phillips and Wagers' predictive Japanese-parser is susceptible to massive garden path effects due to the frequent occurrence of the phonologically null pronoun (*pro*) and head-finality (see Mazuka, 1998; Nakayama, 1999, for ample examples of ambiguity induced by those properties). Next, suppose that the Japanese-parser, as formulated by Phillips and Wagers, operates as a special-purpose parsing strategy (i.e., processing of scrambling). Then it cannot be called up and used in verb phrase-internal scrambling (one of the test constructions for the present study) because no reliable signal for it is available early enough in the input string. That contrasts with the availability of the fronted *wh*-dative phrase as a scrambling marker in Aoshima, Phillips, and Weinberg (2004) on which Phillips and Wagers base their speculation (see also Aoshima, Yoshida, & Phillips, 2009). Finally, no experience-based

by limitations on cognitive resources available during L2 processing. In that light, the study used an experiment with a sentence-final probe recognition task in order to allow enough time for the antecedent of a scrambled element to be mentally reactivated. That contrasts with Felser and Roberts's (2007) measurement of the reactivation effect at the earliest position in the sentence. The study also employed a self-paced reading experiment with two verb phrase (VP)-internal scrambling conditions of differing degrees of computational demand: heavy vs. moderate. That contrasts with Marinis et al.'s (2005) use of a single complex test sentence type. The study results show that Korean learners can perform syntactic gap processing in processing moderately demanding sentences with scrambling but they cannot when reading computationally taxing sentences with scrambling. Those findings contradict the shallow structure hypothesis (Clahsen & Felser, 2006b, 2006c) wherein learners rely on information from verb argument structure and pragmatic knowledge in establishing filler–gap dependencies.[3] The study offers further evidence of L1 transfer in processing of scrambling from data on English learners' processing.

3. The probe recognition experiment

The aim of the probe recognition experiment was to determine if antecedent reactivation effects surface in L2 processing of scrambling in Japanese. In view of the relatively slow speed of L2 processing, the experiment used a sentence-final probe recognition task to allow adequate time for gap processing effects to surface before measurement. McElree and Bever (1989) used the same probe recognition paradigm in their investigation of L1 processing of empty categories in English. They found evidence for antecedent reactivation effects at the clause-final position, not at the earliest position implicating a gap (see also Bever & McElree, 1988; Cloitre & Bever, 1988; MacDonald, 1989; McElree & Bever, 1989, for L1 syntactic gap processing studies using the probe recognition paradigm; see Nicol & Swinney, 2003, for general discussion of that paradigm). If learners engage in syntactic gap processing in reading Japanese sentences with scrambling, antecedent reactivation effects would be similarly detectable at the clause-final position. The non-disruptive nature of the sentence-final probe recognition task was a decisive factor in its selection, for it was expected to make the experiment tasks manageable for learners to perform. The tasks were to comprehend a sentence presented phrase-by-phrase and to determine if a probed word was present or absent in the sentence.

The probe recognition experiment was modeled on Miyamoto and Takahashi (2002a; 2004). It contained two test sentence conditions: a canonically-ordered (subject–object–verb (SOV)) condition (1), and a scrambling (object–subject–verb (OSV)) condition (2). (1) represents a sentence in canonical order (matrix nominative–accusative). (2) is a sentence with scrambling (matrix accusative–nominative) in which the matrix accusative noun phrase (NP), *kyaku-o* 'customer-ACC', is scrambled to the front of the matrix nominative NP, *tenin-ga* 'salesperson-NOM'. The gap resulting from scrambling is indicated by t_i, the canonical position of the matrix accusative NP:

(1) Canonically-ordered condition:

Konbini de	*pan-o*	*katta*	*kyaku-ga*	*hima-soo na*	*tenin-o*
Convenience store at	bread-ACC	bought	customer-NOM	leisurely-looking	salesperson-ACC
yonda.					
called to					

'The customer who bought bread at the convenience store called to the leisurely-looking salesperson.'

---

processing theory allowing for expectation-building has extended its claim as far as Phillips and Wagers have: see Hale, 2001; Konieczny, 2000; Konieczny & Döring, 2003; Levy, 2008. In the psycholinguistics literature on processing of Japanese, far more restricted forms of predictive parsing have been proposed, including the nominative case marker as signaling a clause boundary (Miyamoto, 2002); anticipation of a type of verb in terms of verb semantics (Kamide, Scheepers, & Altmann, 2003); and a top-down incremental narrowing-of-predictions hypothesis (Nakatani & Gibson, 2008) which is adopted in the present study.

[3] Clahsen and Felser (2006c, p.564) hold that learners' reliance on non-syntactic information in processing of filler–gap dependencies arises from the lack of "sufficiently rich, implicit grammatical knowledge" in L2 grammars. See Dekydtspotter, Schwartz, and Sprouse (2006) for criticism of Clahsen and Felser's stance in view of the substantial body of L2 knowledge representation research showing evidence to the contrary.

(2) Scrambling condition:

Konbini de	*pan-o*	*katta*	*kyaku$_i$-o*	*hima-soo na*	*tenin-ga*
Convenience store at	bread-ACC	bought	customer-ACC	leisurely-looking	salesperson-NOM

t_i *yonda.*
called to

'The leisurely-looking salesperson called to the customer who bought bread at the convenience store.'

Both conditions have identical linear order and content words. In order to achieve that, semantically reversible verbs (*yonda* 'called to') were used as the main-clause verb, so that pragmatically, either argument of the verb (*kyaku* 'customer' or *tenin* 'salesperson') could be taken as subject or object of the main clause. The second content word (*pan* 'bread') functioned as the target word. It is a constituent of the subject-relativized relative clause contained in the initial matrix NP (subject in (1) and object in (2)). Note that the linear distance from the target word to the end of the sentence is identical in both conditions.

Participants read sentences such as (1) and (2), as segmented above, in a self-paced reading format (Just, Carpenter, & Woolley, 1982). Immediately after they read the last segment, a probe (*pan* 'bread') appeared on the computer screen. They had to determine whether the probe was in the sentence they had just read. Because the probe was presented sentence-finally, learners should have had enough time to carry out syntactic gap processing prior to its presentation, making this a suitable task for taking learners' slow processing speed into account.

It was hypothesized that in the scrambling condition, the scrambled matrix object NP (*kyaku* 'customer') containing the target word (*pan* 'bread') would be mentally reactivated when the matrix subject NP (*tenin* 'salesperson') was encountered, as a result of syntactic gap processing involving association of the gap with its antecedent. Therefore, under a structure-based processing view, a faster response time, and possibly a better accuracy rate would be expected in the probe recognition task in the scrambling condition than in the canonically-ordered condition. Because of the absence of a gap in the canonically-ordered condition, there should be no reactivation of the matrix subject NP containing the target word in that condition.[4]

Alternatively, under a verb-driven processing scenario, two lines of predictions would follow. One possibility is that if learners relied on verb-driven word integration, there would be no difference between both conditions in probe recognition times and accuracy. That would be expected because the linear distance from the target word to the sentence-final verb is identical in both conditions. A second possibility is that if learners' word integration process consults information from the thematic hierarchy, their performance on the probe recognition task would be better in the canonically-ordered than in the scrambling condition. That would be likely because the memory trace of the subject NP should be more accessible than the memory trace of the object NP, irrespective of sentence word order, due to the prominence of Agent (the semantic role that the subject NP is assigned) over Patient (the role the object NP is assigned or any other thematic roles) in the hierarchy (Jackendoff, 1990). Importantly, structure-based and verb-driven processing views would predict two different performance outcomes on the probe recognition task, even though the probe was presented immediately post-verbally in the experiment.

[4] McKoon and Ratcliff (1980) have shown that a probe is recognized faster when it is identical to a word that appears in the sentence just processed than when the probe does not appear in the sentence. That priming (or repetition effect) results from the prominence of the probe (appearing in the sentence) in working memory (McKoon & Ratcliff, 1984; Nicol & Swinney, 2003, p. 76). In the present experiment, a similar priming effect was expected in both scrambling and canonically-ordered conditions. The priming effects in the scrambling condition should be greater than in the canonically-ordered condition because processing of the former includes antecedent reactivation effects.

Table 1
Learner Participants' Background Information and Proficiency Test Scores

L1	Age (yrs)	JLPT Scores (%)	Length of Study (yrs)	Visiting Experience (mo)
Korean				
M	23.05	97.1	4.71	11.5
Range	18 – 33	89.3 – 100	2.0 – 8.0	3.5 – 29.0
SD	3.52	.95	1.85	8.76
English				
M	24.45	93.0	5.75	23.7
Range	19 – 34	82.1 – 100	3.0 – 14.0	3.0 – 104
SD	3.38	1.85	2.73	25.19

Note. JLPT stands for the Japanese Language Proficiency Test.

The Korean learners were expected to engage in syntactic gap processing of sentences with scrambling on a par with Japanese native speakers (see Miyamoto & Takahashi, 2002a, 2004, for evidence of L1 gap processing of scrambling in Japanese) due to the typological proximity between Korean and Japanese (both include scrambling properties). The English learners were expected to show difficulty with syntactic gap processing in processing of scrambling due to the lack of that property in English. Therefore, it was predicted that Korean learners would perform the probe recognition task better on the scrambling than on the canonically-ordered condition; English learners would not.

3.1. Participants

There were three groups of participants in the probe recognition experiment: 1) Korean-speaking and 2) English-speaking learners of Japanese; 3) native speakers of Japanese. Each group had 20 participants. Table 1 presents background information on the learner participants' Japanese language-learning experiences (note: all tables are provided at the end of the article).[5] All Korean learners were at an advanced level of language study at the Center for Japanese Language, Waseda University, Tokyo. At the time of experimentation, they had lived in Japan for at least three and a half months and many were taking undergraduate or graduate courses at Waseda University as degree-seeking students. Most of the English learners were students at the Inter-University Center for Japanese Language Studies in Yokohama, and had resided in Japan for at least eight and a half months at the time of experimentation. All English learners were advanced learners of Japanese and post-baccalaureates: some were graduate students studying a Japan-related field, others were preparing for a profession using Japanese (e.g., translation). All Japanese native speaker participants were university students in Tokyo. All participants were compensated for their participation in the experiment.

3.2. Materials

The material for the probe recognition experiment included a total of 36 sentences: 12 test sentence items and 24 fillers. A total of 12 pairs of canonically-ordered and scrambling test sentences (such as (1) and (2)) were created. Using those pairs, two experimental lists were generated in a Latin Square design: one test sentence of a pair was assigned to the first experimental list, the other sentence of that pair to the second list, and so on through all 12 pairs. Both lists received six canonically-ordered and six scrambling test sentences. Neither list had both of the two sentences of any pair. Each list also had a total of 24 filler sentences that varied in structural complexity and length. The two lists were pseudo-randomized by dividing each into six blocks, with each block containing six sentences in total: two test sentences (one canonically-ordered and one scrambling), and four filler sentences. The ordering of the blocks within each list, and the ordering of the sentences within each block was

[5] Korean and English learners' group average scores on the JLPT were significantly different from one another ($F(1, 38) = 6.117$, $p = .018$). Note that the English learner group attained a very high score nevertheless.

pseudo-randomized so that test sentences were non-adjacent. Each list began with a filler sentence. Finally, the number of correct and incorrect probes and the number of true and false statements (both described below) were counterbalanced within each block. Each participant saw only one of the two lists in the experimental session.

3.3. Procedure

Prior to their arrival at an experimental session, learner participants completed a questionnaire on their Japanese language learning experience and personal background. They also individually took the grammar section of the Japanese Language Proficiency Test (extracted from Levels 2 and 3) and were instructed to study a list of vocabulary items and *kanji* (logographs adopted from Chinese) that appeared in the experimental material. The experimental session consisted of two experiments: the first was a probe recognition experiment; the second was a self-paced reading experiment. The two experiments were separated by a relaxation break during which the learner participants took a vocabulary and *kanji* character test to measure their familiarity with those items used in the experiments.[6] Completion of the entire session took learners 70 to 80 minutes. Native speaker participants took the same two experiments, also separated by a break during which they filled out a questionnaire on their foreign language learning experience and personal background. They completed their session in about 50 minutes. The experimental sessions took place in the author's office.

Each participant performed three tasks in the experiment. A participant first read each sentence by the phrase-by-phrase self-paced reading method (Just et al., 1982). Each sentence was presented segment-by-segment, as indicated in (1) and (2).[7] Each region appeared in the center of a 17-inch display screen on a Toshiba laptop computer, in black letters on a white background in Mincho 24-point font. At each trial, a '+' mark first appeared in the center of the display. The mark was replaced with the first region of a sentence when the participant pressed the leftmost button of a response button-box connected to the computer. After having read the first region, the participant pressed the same button to call up the second region, and continued doing so until reaching the end of the sentence. Next, immediately after the final segment of the sentence was read, a probe appeared in blue at the center of a new screen. The experimental task was to determine as quickly and as accurately as possible whether the probe had appeared in the sentence the participant had just read, then to press one of the two rightmost buttons of the button-box, one colored green (correct), the other red (incorrect). All 12 probes for the test sentences were correct; that is, the probes appeared in the test sentences. Six of the 24 probes for the filler sentences were correct, 18 were incorrect, so that an equal number of correct and incorrect probes appeared in each experimental list. The correct filler-sentence probes were selected from various positions in the filler sentences so that participants would not develop a strategy of simply paying attention to the first few words of the sentence.

Third, after the participant responded to the probe, a brief statement was presented at the center of a new screen. The experimental task was to determine if the statement was true or false relative to the sentence the participant had just read, then to press one of the two rightmost buttons of the button-box: green (true) or red (false). The purpose of this task was to ensure that the participant read each sentence for comprehension instead of simply memorizing each word. After the participant responded to the statement, graphic feedback of a sad-looking face was provided when the response was incorrect (graphic feedback was explained to the participant in the instructions for the experiment). The true–false statements were counterbalanced, with half of the test and filler sentences having true statements, the other half having false statements. The accuracy of the response to each probe and to each true–

[6] Learner participants' results on the vocabulary and *kanji* test indicated that they were familiar with all test items in lexical terms. Thus all test items were included in the initial analysis of learner participants' data.

[7] Because learners might be unfamiliar with some *kanji* (logographic characters adopted from Chinese) appearing in the experimental material, Japanese phonetic symbols called *hurigana* were added above *kanji* suspected of being unfamiliar to the learners to indicate pronunciation. That writing convention is used in Japanese newspapers, magazines, and books containing uncommonly used *kanji*. Thus it did not introduce any peculiarity into the experiment. No participant (native or non-native) expressed discomfort with the occasional presence of *hurigana* in the experimental material when interviewed after the experiment.

Table 2
Participants' Probe Recognition Accuracy Rates (%)

Group	Condition	
	Canonically-ordered	Scrambling
NSs of Japanese		
M	97.3	96.5
SD	.065	.118
Korean learners		
M	97.3	93.2
SD	.065	.112
English learners		
M	92.5	96.7
SD	.101	.083

false statement was recorded. The response time on each probe was measured using E-Prime software (Schneider, Eschman, & Zuccolotto, 2002), as was the reading time for each segment of every sentence.

A hard copy of the instructions for the experiment (written in the participant's L1) was provided to the participant at the beginning of the experimental session. After the experimenter (i.e., the author) confirmed that the participant understood the experimental procedures, the participant proceeded to a practice session containing eight practice sentences not used in the testing session. The participant had to respond correctly to at least two-thirds of the true or false statements in the practice session in order to proceed to the testing session. The probe recognition experiment took learner participants approximately 20 minutes and native speaker participants 15 minutes to complete.

3.4. Results

The Japanese native speaker, Korean learner, and English learner groups responded correctly to the true or false statements appearing after the test sentences at a rate of 88.75%, 91.25%, and 90.42%, respectively. There was no significant difference in accuracy between groups ($F(2, 57) = .277$, $p = .759$).[8] Excluded from subsequent analyses were probe recognition data from the trials in which participants incorrectly responded to true–false statements on test sentences.

Table 2 presents the three participant groups' probe recognition accuracy rates on the canonically-ordered and scrambling conditions. All three groups recognized probes with high accuracy on both conditions. Repeated-measures ANOVAs by both participants and items were conducted on the accuracy rates for each participant group. No participant group showed a significant difference in probe recognition accuracy comparing both conditions: $F_1(1, 19) = .069$, $p = .796$, $F_2(1, 11) < .0005$, $p = .985$ (the Japanese native speaker group); $F_1(1, 19) = 2.134$, $p = .160$, $F_2(1, 11) = 1.887$, $p = .197$ (the Korean learner group); and $F_1(1, 19) = 1.557$, $p = .227$, $F_2(1, 11) = 1.189$, $p = .299$ (the English learner group). An accuracy rate of around 95% for all three groups indicates a ceiling effect. Their nearly perfect performance correctly recognizing probes, along with their high accuracy verifying true–false statements, suggests that both native speaker and learner participants had little difficulty processing the test sentences.

Prior to the analysis of the probe recognition response time data, the following two data trimming measures were performed. First, data points from the trials in which participants failed to recognize probes in test sentences were discarded, resulting in elimination of 2.82% of data from the native speaker group, 4.57% from the Korean learner group, and 5.07% from the English learner group. Second, for each participant, response times that lay beyond two standard deviations from the participant's mean response time were eliminated: those outliers comprised 3.38% of the response time data for the native speaker group, 4.31% for the Korean learner group, and 6.80% for the English

[8] The three participant groups' accuracy rates for all true–false statements on both test and filler sentences were 92.78% for the native speaker, 94.58% for the Korean learner, and 94.03% for the English learner group, with no significant difference in accuracy between groups ($F(2, 57) = .521$, $p = .596$).

Table 3
Participants' Probe Recognition Response Times (ms)

Group	Condition	
	Canonically-ordered	Scrambling
NSs of Japanese		
M	871.6	846.9
SD	225.0	176.8
Korean learners		
M	1042.9	1005.9
SD	386.7	349.2
English learners		
M	1277.1	1175.5
SD	402.1	318.1

learner group. Each participant's mean response time was calculated by averaging the participant's response times to correctly recognized probes appearing in the test or filler sentences.

Table 3 presents the three participant groups' probe recognition response times in the canonically-ordered and scrambling conditions. All three groups exhibited the same response time pattern, with shorter response times on the scrambling condition than on the canonically-ordered condition. For each of the participant groups, repeated-measures ANOVAs by both participants and items were conducted on the response times for both conditions. There was no significant difference in response time between the two conditions in the Japanese native speaker group: $F_1(1, 19) = 1.270, p = .274, F_2(1, 11) = .411, p = .535$. On the other hand, the Korean learner group responded faster to probes for the scrambling condition than it did to probes for the canonically-ordered condition in the participants analysis ($F_1(1, 19) = 4.885, p = .040$), but not in the items analysis ($F_2(1, 11) = .314, p = .586$). Similarly, the English learner group showed a significant difference in response time comparing both conditions in the participants analysis ($F_1(1, 19) = 9.482, p = .006$), but not in the items analysis ($F_2(1, 11) = .019, p = .893$).

The absence of a response time difference between the two conditions for the native speakers likely resulted from test sentences having imposed insufficient computational demand on them (cf. Miyamoto & Takahashi, 2002a). Importantly, the Korean learners showed antecedent reactivation effects, as evinced (in the participants analysis) by their having responded faster to probes in the scrambling than to those in the canonically-ordered condition. Finally, the English learners also showed evidence for antecedent reactivation effects. That result is inconsistent with the prediction that English learners would not engage in syntactic gap processing of scrambling. Possible sources of their antecedent reactivation effects are raised in the Discussion.

4. The self-paced reading experiment

Whereas the probe recognition experiment aimed at the outcome of syntactic gap processing, the self-paced reading experiment examined real-time processing of filler–gap dependencies. The aim of the self-paced reading experiment was to obtain real-time evidence of L2 syntactic gap processing of scrambling in Japanese. It used an indirect object–direct object construction,[9] along with an adverbial phrase (AdvP) (locative or temporal), to create short and long versions of VP-internal scrambling (cf. Miyamoto & Takahashi, 2002a; 2004). There were three test sentence conditions: a canonically-ordered condition (3), a short scrambling condition (4), and a long scrambling condition (5).[10] (3) represents a sentence in canonical order with an indirect object–direct object word order (i.e., *wueetoresu-ni* 'the waitress-DTA'–*kokku-o* 'the cook-ACC'). (4) is a sentence with short scrambling

[9] See Sadakane and Koizumi (1995, pp. 11-16) for linguistic tests determining whether *ni* is dative or postpositional. Grammaticality judgments from those tests show that all instances of *ni* in the test sentences are dative case-marking.

[10] Note that the term *long* scrambling is used to differentiate it from *short* scrambling. Long scrambling differs from *long-distance* scrambling wherein a constituent is scrambled across the clause boundary.

in which the direct object <u>*kokku$_i$-o*</u> 'the cook-ACC' has been scrambled to the position immediately before the indirect object *wueetoresu-ni* 'the waitress-DTA' (the scrambled direct object is underlined and its canonical position is indicated by "t_i"). (5) is a sentence with long scrambling in which the direct object <u>*kokku$_i$-o*</u> 'the cook-ACC' has been more distantly scrambled to the position immediately in front of the AdvP *resutoran no atarasii kicchin de* 'in the new kitchen of the restaurant':[11]

(3) Canonically-ordered condition:
Maneejyaa-wa resutoran no atarasii kicchin de wueetoresu-ni kokku-o
Manager-TOP restaurant new kitchen in waitress-DAT cook-ACC
syookai-sita soo da.
introduced seems

(4) Short scrambling condition:
Maneejyaa -wa resutoran no atarasii kicchin de <u>*kokku$_i$-o*</u> *wueetoresu-ni* t_i
Manager-TOP restaurant new kitchen in cook-ACC waitress-DAT
syookai-sita soo da.
introduced seems

(5) Long scrambling condition:
Maneejyaa -wa <u>*kokku$_i$-o*</u> *resutoran no atarasii kicchin de wueetoresu-ni* t_i
Manager-TOP cook-ACC restaurant new kitchen in waitress-DAT
syookai-sita soo da.
introduced seems
'The manager seems to have introduced the cook to the waitress in the new kitchen of the restaurant.'

In each condition, the region immediately preceding the verb is critical: in (4) and (5), it implicates the presence of a structural gap; in (3) that gap is absent. The degree of difficulty in carrying out the syntactic prediction-updating process varies at that critical region such that (5) > (4) > (3). According to the incremental narrowing-of-predictions hypothesis (Nakatani & Gibson, 2008, p. 66), a version of dependency locality theory (Gibson, 1998; 2000),[12] syntactic predictions made on upcoming elements are continuously narrowed as new words are processed in the input string (cf. endnote 1). Importantly, the greater the distance between the points where a syntactic prediction has *last* been narrowed and where a new word pertaining to that prediction is processed, the greater the difficulty of matching the syntactic prediction to the new word. In the above two scrambling conditions, when the accusative NP is processed, the expectation is narrowed because of the nominative–accusative NP sequence in the input string. In matching that syntactic prediction with the subsequent dative NP, the parser is signaled that a correctly narrowed-down syntactic prediction for the input string consists of a construction involving an indirect object–direct object structure. The parser is further alerted that scrambling is involved, and syntactic gap processing ensues. Little difficulty should arise in matching the syntactic prediction with the dative NP in the short scrambling condition (4) because the matching process occurs immediately after the prediction has been formed. In the long scrambling condition (5), however, the intervening AdvP elongates the distance between the points where the prediction has been formed and where it is matched with the dative NP. Therefore, (5) should require greater computational resources than (4) during the matching process at the dative

[11] See Tonoike (1997) for a non-movement "base-generation" account of scrambling in Japanese. Both the movement account and the base-generation account seem to point to the prediction that sentence condition (5) is more difficult than (4) because in both accounts processing of scrambling involves essentially the same operations: identifying a phonologically null element (a trace in movement or *pro* in base-generation) and associating it with its antecedent (in movement) or its binder (in base-generation).

[12] In the case of processing of German, Konieczny (2000) reports counter-evidence for processing difficulty predicted by locality-based integration costs (Gibson, 1998; 2000). See Grodner and Gibson (2005, pp. 284-285) for discussion of how locality effects might have been masked due to properties of the materials used in Konieczny's study.

NP, even though the computational cost associated with the ensuing syntactic gap processing will be similar in both conditions. Because (3) is in canonical word order, and excludes syntactic gap processing, matching the syntactic prediction with each newly arrived word should proceed with no difficulty.

It was expected that the long scrambling condition (5) would tax Japanese native speakers' computational resources sufficiently enough to cause them to slow down their reading at the critical region. The short scrambling condition (4) would not. As for L2 learners whose L1 exhibits scrambling properties (as in Korean), suppose that a high demand on their computational resources impedes syntactic prediction-updating or syntactic gap processing, while a moderate demand does not. Then an asymmetry in learners' reading time pattern would appear that is the inverse of native speakers' asymmetrical reading time pattern. Learners would exhibit a reading slowdown at the critical region in (4) because the moderate computational effort required for processing the short scrambling condition would not impede their syntactic prediction-updating or syntactic gap processing. Learners would show no reading slowdown at the critical region in (5), however, because the long scrambling condition would overtax their computational resources. The primary reason for that particular asymmetrical reading time pattern would be learners' difficulty with the syntactic prediction-updating (i.e., matching) process, not with syntactic gap processing per se. Crucially, the degree of difficulty with matching varies between the two scrambling conditions, whereas syntactic gap processing is the same in both.[13] Finally, if learners' L1 lacks scrambling (as in English) they may not exhibit a reading slowdown at the critical region in either (4) or (5), regardless of their difficulty with the matching process.

If learners relied on verb-driven rather than structure-based processing, Korean and English learners alike would exhibit a reading slowdown at the verb region comparing the three conditions. Under Pickering and Barry's (1991) direct association hypothesis, a pre-posed element is directly associated with its subcategorizer (e.g., the verb) without establishing the filler–gap dependency. Because the direct association occurs at the subcategorizing verb, difficulty associating the pre-posed element with its verb should result in a reading slowdown at the verb region (see Miyamoto & Takahashi, 2002b, pp. 179-180 for similar reasoning). Importantly, the linear distance between a pre-posed element and a subcategorizer functions as a primary predictor of such difficulty: the greater that distance, the more difficult to associate the pre-posed element with its subcategorizing verb (Pickering & Barry, 1991, pp. 237-238; see also Gorrell, 1993, pp. 132-137). Therefore, learners would spend a longer time reading the verb region in the long scrambling condition than in either the short scrambling or the canonically-ordered conditions because the distance between the direct object NP and its subcategorizing verb is greater in the long scrambling condition than it is in either the short scrambling or canonically-ordered conditions. There might be no difference (or perhaps a marginal one) in reading of the verb region, comparing the latter two conditions.[14]

4.1. Participants

All participants in the probe recognition experiment participated in the self-paced reading experiment during the same session. The self-paced reading experiment was presented after the probe recognition experiment.

---

[13] It is unlikely that the infrequency of scrambling in Japanese (cf. Yamashita, 2002) would greatly contribute to learners' difficulty in processing sentences with scrambling. Space limitations prevent detailed discussion here, but see Fodor & Hirose (2003) and Hara (2010). Yamashita (1997) reports experimental evidence of a lack of discernable frequency effects on Japanese native speakers' processing of scrambling.

[14] An anonymous reviewer suggests an alternative prediction under a verb-driven processing scenario. In the reviewer's view, no reading time difference should occur at the verb region across conditions because all information on argument NPs is available and "ready to be integrated at the verb". Until the verb is reached, the information is stored in WM. Such a strict version of verb-driven processing has been rejected as a plausible Japanese processing model (see Miyamoto, 2008, pp. 221-224, for theoretical considerations and empirical evidence). Given that L2 participants in the present study were advanced learners of Japanese, it is highly unlikely that those participants' processing of Japanese was strictly verb-driven as the reviewer hypothesizes (see also footnote 2).

4.2. Materials

First, 29 sets of test sentences were contrived. Each of those sets contained four sentence conditions: the three shown in (3) to (5) and a fourth unrelated to the present experiment. The 29 test sentence sets were then subjected to a norming survey. Four questionnaires were created for the survey in a Latin Square design, with each questionnaire containing seven sentence tokens of each of the four conditions, plus an additional sentence token of one of the four. Twenty-two fillers, all grammatical, were added to the 29 test sentences in each questionnaire, for a total of 51 sentences that were randomized. Sixty native speakers of Japanese participated in the norming survey, with 15 respondents per questionnaire, none of whom participated in the self-paced reading experiment. They rated all test and filler sentences using a 7-point scale from 1 'natural' to 7 'unnatural'. Twenty-four sets of test sentences were selected from the initial 29 sets, based on which of those sets had received the highest naturalness ratings in the questionnaire survey. The average rating scores for the selected test sentences were 1.90 for the canonically-ordered, 2.03 for the short scrambling, and 2.64 for the long scrambling condition.

Next, using the 24 selected sets, four experimental lists were generated in a Latin Square design, with each list containing six test sentences of each of the four conditions: 18 of the 24 test sentences per list were relevant for the self-paced reading experiment. Each experimental list also contained 48 fillers, half of which were included for a separate experiment while the other half varied in structure and length. Thus there were a total of 72 sentences in each experimental list. The four lists were each pseudo-randomized: each list was divided into six blocks, with each block having one sentence of each of the four conditions and eight filler sentences. The ordering of the 12 (test and filler) sentences within each block was pseudo-randomized so that at least one filler sentence intervened between two test sentences. The presentation of the six blocks within each list was also randomized. Finally, within each block, half of the test and filler sentences were assigned true verification statements; the other half received false verification statements.

4.3. Procedure

Participants performed the self-paced reading experiment after a relaxation break following their completion of the probe recognition experiment. The self-paced reading experiment employed a segment-by-segment self-paced reading paradigm (cf. Just et al., 1982), with each sentence presented region-by-region as segmented in (3) to (5). The procedures for the self-paced reading experiment were identical to those for the probe recognition experiment, with two exceptions. First, there was no probe response task in the self-paced reading experiment. Second, upon completing the first half of the self-paced reading experiment, there was a brief relaxation break. Then, after correctly performing two additional practice trials, the participant continued the experiment until finishing it. The self-paced reading experiment took learner participants 40 to 50 minutes and native speaker participants approximately 30 minutes to complete.

4.4. Results

The Japanese native speaker, Korean learner, and English learner groups responded correctly to the verification statements appearing after the test sentences at a rate of 90.42%, 90.63%, and 90.0%, respectively. Reading time data from test sentences with incorrectly answered verification statements were excluded from analyses. A regression equation for predicting reading time using region length was computed for each participant, using all sentences (test and filler) whose verification statements the participant correctly answered. The region length was defined in terms of the number of morae in the region. For each participant and at each region, the predicted reading time was subtracted from the observed reading time to generate the residual reading time. This statistical procedure removed extraneous variance by subtracting out the participant's button-press baseline time and by controlling for length effects due to differences in region length. All analyses were performed using residual reading time data (see Ferreira & Clifton, 1986; Trueswell, Tanenhaus, & Garnsey, 1994, for discussion of residual reading time analysis).

4.4.1. Processing of the preverbal region in sentences with short scrambling

The analysis compared how participants read from the pre-critical region to its adjacent critical region (one corresponding to the direct object, the other to the indirect object) in the short scrambling and canonically-ordered conditions. The goal was to determine if there was an interaction effect between sentence condition type and reading slowdown in those adjacent regions. In the short scrambling condition, a slowdown in reading the two regions was expected to be slight, if at all, for native speakers of Japanese, and moderate for Korean learners. In the short scrambling condition, matching the syntactic prediction with each newly arrived word (first the direct object, then the indirect object NP) requires only limited computational effort; at the same time, syntactic gap processing is required because the direct object NP has been scrambled over the indirect object NP. In the canonically-ordered condition, Japanese native speakers and Korean learners alike should show no slowdown in reading the adjacent indirect–direct object NP regions. Matching the syntactic prediction with the indirect object–direct object NP sequence is a straightforward narrowing down of syntactic expectations and no syntactic gap processing occurs in processing sentences in canonical order. English learners, whose L1 lacks scrambling, might not slow down in reading from the pre-critical to adjacent critical region in sentences with short scrambling, even though such sentences place only limited demand on computational resources.

Table 4 presents residual reading times for the three participant groups' reading from the pre-critical to the critical region in both short scrambling and canonically-ordered conditions. Repeated-measures ANOVAs by both participants and items were conducted on the residual reading times at those regions for each of the participant groups. Importantly, the Korean learner group showed the expected (albeit marginal) interaction effect in the participants analysis ($F_1(1, 19) = 4.236$, $p = .054$). The effect was not significant in the items analysis ($F_2(1, 23) = .870$, $p = .361$), suggesting that the effect did not generalize over all items. As consistent with the incremental narrowing-of-predictions hypothesis (Nakatani & Gibson, 2008), in the canonically-ordered condition, Korean learners' reading from the pre-critical indirect object region to the adjacent critical direct object region sped up by an average of 39.32 milliseconds residual reading time per mora (*SD* = 186.3). Yet, as predicted, in the short scrambling condition, Korean learners' reading from the pre-critical direct object region to the adjacent critical indirect object region slowed down by an average of 62.58 milliseconds residual reading time per mora (*SD* = 225.1). Those differences resulted in the Korean learner group's observed interaction effect. Although the Japanese native speaker group showed interaction effects ($F_1(1, 19) =$

Table 4
Residual Reading Times from the Pre-Critical Region to the Adjacent Critical Region in the Short Scrambling and Canonically-Ordered Conditions (ms)

Group	Condition			
	Short scrambling		Canonically-ordered	
	Pre-critical (Direct object)	Critical (Indirect object)	Pre-critical (Indirect object)	Critical (Direct object)
NSs of Japanese				
M	-25.335	-55.134	42.220	-46.696
SD	77.447	60.526	113.035	66.962
Korean learners				
M	-96.920	-34.341	4.947	-34.376
SD	176.530	157.353	129.960	131.699
English learners				
M	-54.851	-82.070	-41.203	-122.978
SD	228.353	284.845	455.658	294.731

5.728, $p = .027$; $F_2(1, 23) = 5.013$, $p = .035$), its effects resulted from the much greater increase in reading speed from the pre-critical to the critical region in the canonically-ordered condition as compared to the short scrambling condition. Consistent with the incremental narrowing-of-predictions hypothesis, Japanese native speakers' reading sped up considerably in the canonically-ordered condition, but only slightly in the short scrambling condition. Their reading speed increases were 88.92 milliseconds ($SD = 144.9$) in the former and 29.80 milliseconds ($SD = 96.6$) in the latter in terms of residual reading time per mora. The English learner group showed no indication of an interaction effect: ($F_1(1, 19) = .111$, $p > .70$; $F_2(1, 23) = .011$, $p > .90$).

In the short scrambling and canonically-ordered conditions, all regions other than the pre-critical and adjacent critical regions were identical. As expected, there was no reading time difference between the two conditions at those other regions (for the native speaker group, F_1s ≤ 2.827, *p*s ≥ .109; F_2s ≤ 1.173, *p*s ≥ .290; for the Korean learner group, F_1s ≤ .670, *p*s ≥ .423; F_2s ≤ 1.250, *p*s ≥ .275; for the English learner group, F_1s ≤ 2.213, *p*s ≥ .153; F_2s ≤ 4.250, *p*s ≥ .051), except for the native speaker group's theoretically (but not statistically) insignificant reading time differences at the last two regions of the advP *atarasii kicchin de* 'in the new kitchen' (F_1s > 4.495, *p*s ≤ .047 ; F_2s > 2.746, *p*s ≤ .111).

4.4.2. Processing of the preverbal region in sentences with long scrambling

Table 5 presents the three participant groups' residual reading times at the critical indirect object region in the short and long scrambling conditions. Repeated-measures ANOVAs by both participants and items were conducted on the residual reading times at that critical region for each of the participant groups. The Japanese native speaker group read the long scrambling critical region more slowly than it did the critical region in the short scrambling condition, with significant difference ($F_1(1, 19) = 8.796$, $p = .008$; $F_2(1, 23) = 4.346$, $p = .048$). Neither the Korean nor the English learner group exhibited that slowed reading pattern with statistical significance (the Korean learner group: $F_1(1, 19) = .111$, $p = .743$; $F_2(1, 23) = .012$, $p > .80$; the English learner group: $F_1(1, 19) = .065$, $p > .80$; $F_2(1, 23) = .298$, $p > .50$). These findings on the native speakers and on the learners suggest that first, the native speakers' slower reading of the long scrambling critical region confirms that they had to expend greater computational effort to match their syntactic prediction with the indirect object in the long scrambling than in the short scrambling condition. Second, the greater-effort requirement of the matching process in the long scrambling condition likely overtaxed learners' computational resources.

There was no reading time difference between the long and short scrambling conditions at all non-critical content-matched regions for the Korean learner group (F_1s ≤ 1.869, F_2s ≤ 2.105) or for the English learner group (F_1s ≤ 2.136, F_2s ≤ 2.423). For the native speaker group, its reading times at the first two regions of the AdvP differed significantly between both conditions, but those differences disappeared at the last region of the AdvP ($F_1 = 1.802$, $p = .195$; $F_2 = .533$, $p = .473$). Interestingly, there was a reading time difference between both conditions at the region containing the direct object NP for the native speaker group ($F_1 = 6.108$, $p = .023$; $F_2 = 2.668$, $p = .116$) and for the Korean learner group ($F_1 = 9.287$, $p = .007$; $F_2 = 5.007$, $p = .035$), but not for the English learner group ($F_1 = 1.701$, *p*

Table 5
Residual Reading Times at the Critical Indirect Object Region in the Short and Long Scrambling Conditions (ms)

Group	Condition	
	Short scrambling	Long scrambling
NSs of Japanese		
M	-55.134	24.126
SD	60.526	127.343
Korean learners		
M	-34.341	-20.751
SD	157.353	177.961
English learners		
M	-82.070	-107.032
SD	284.845	318.766

= .208 ; F_2 = 2.780, p = .109). For native speakers and Korean learners, the dislocated direct object NP was read more slowly in the long scrambling than in the short scrambling condition. That difference is consistent with Nakatani and Gibson's (2008) incremental narrowing-of-predictions hypothesis. The direct object NP appeared earlier in the sentence with long scrambling than it did in the sentence with short scrambling (the second vs. fifth region, respectively), so that when the direct object NP was encountered at the earlier position in the sentence with long scrambling, syntactic expectations were not as narrowed as they were when it was encountered at a later point in the sentence with short scrambling.

4.4.3. Processing of the verb regions in the three conditions

Participants' reading times at the verb regions were also compared to ascertain whether they had relied on information from verb argument structure in processing sentences with short and long scrambling. Table 6 presents the residual reading times for each of the three participant groups' reading of the verb region in each of the three sentence conditions. Repeated-measures ANOVAs, both by participants and items, were conducted on the residual reading times at the verb regions for each of the three participant groups. The Japanese native speaker group showed no difference in reading times for the verb regions comparing the three conditions: ($F_1(2, 18)$ = .678, p > .50; $F_2(2, 22)$ = .199, p > .50). Nor did the Korean or the English learner group show any reading time difference (the Korean learner group: $F_1(2, 18)$ = .120, p > .50; $F_2(2, 22)$ = .636, p > .50; the English learner group: $F_1(2, 18)$ = 1.085, p = .359; $F_2(2, 22)$ = 1.495, p = .246). Therefore, no evidence was found that either native speakers or learners relied on information from verb argument structure in processing sentences with either short or long scrambling, as measured by their reading times at the verb regions (see Hara, 2009, for discussion of the incompatibility of verb-driven processing with incremental processing of Japanese).

5. Discussion

5.1. English and Korean learners' processing of scrambling in Japanese

Owing to the lack of scrambling in English, it was expected that English learners would have difficulty engaging in syntactic gap processing of sentences with scrambling in Japanese. A possible reason for their successful antecedent reactivation effects is the similarity between the scrambling construction tested (*kyaku*$_i$*-o* 'customer-ACC' *tenin-ga* 'salesperson-NOM' t_i *yonda* 'called to') and the English direct object *wh*-question construction (*Who*$_i$ *did the salesperson call to* t_i*?*). In both constructions, the direct object NP is fronted immediately before the subject NP, which leaves a trace and leads to syntactic gap processing in comprehending those sentences. It is possible that English learners' antecedent reactivation effects stemmed not from processing of the scrambling construction per se, but from their processing of those sentences as if they were a direct object *wh*-question

Table 6
Residual Reading Times at the Verb Region in the Three Sentence Conditions (ms)

Group	Condition		
	Canonically-ordered	Short scrambling	Long scrambling
NSs of Japanese			
M	-2.880	-24.994	-1.719
SD	81.251	110.770	60.787
Korean learners			
M	43.126	20.578	23.755
SD	208.534	176.551	148.263
English learners			
M	45.172	-88.784	-7.132
SD	262.879	219.837	293.243

construction.[15] That view accords with the lack of evidence from the self-paced reading experiment that English learners engaged in syntactic gap processing in reading sentences with VP-internal scrambling, for which no analogous construction exists in English.[16]

Turning to the self-paced reading experiment, Korean learners' reading slowed down from the region immediately preceding the critical region to the critical gap-implicating region in the short scrambling condition, as compared to their reading time across the two corresponding regions (one involving no gap) in the canonically-ordered condition.[17] The two regions in each condition contained reversed-order indirect and direct object NPs; all other regions were identical in both conditions. On the other hand, Korean learners did not read the gap-implicating region more slowly in the long scrambling condition than in the short scrambling condition. English learners exhibited no evidence of reading slowdown at the gap-implicating region in either the short or long scrambling condition. Korean learners' reading slowdown at the preverbal gap-implicating region in the short scrambling condition contradicts the expectation of verb-driven processing that a reading slowdown in that condition would occur at the verbal, not preverbal region. Indeed, no significant reading slowdown at the verbal region was found for any of the three participant groups when all three conditions were compared.

In the probe recognition experiment, Korean and English learners of Japanese both evinced antecedent reactivation effects in processing sentences with clause-internal scrambling. They responded faster to probes in the scrambling than to those in the canonically-ordered condition. The asymmetrical response time results came from their performance on those sentences with scrambling and on those in canonical order that differed only with respect to the presence or absence of a preverbal syntactic gap resulting from scrambling. That finding is inconsistent with the two predictions following from a verb-driven processing scenario in which information on verb argument structure and pragmatic knowledge are used to associate argument NPs with the verb, without recourse to use of syntactic gap information. In one prediction, no response time difference would be expected between the scrambling and canonically-ordered conditions because the linear distance from the verb to the target word is identical in both conditions. In a second prediction under verb-driven processing, responses to probes on the canonically-ordered condition should be faster than responses to probes on the scrambling condition because the Agent NP containing the target word should be more recognizable than the Patient NP containing the target word, owing to the prominence of Agent in the thematic hierarchy (cf. Jackendoff, 1990). Neither of those two verb-driven processing-scenario predictions was supported by evidence from the results of the probe recognition experiment.

Before considering the significance of the present experimental findings vis-à-vis my working hypothesis of cognitive-resource capacity limitations on L2 syntactic gap processing, a cautionary note on use of end-of-the sentence probing is in order. In the present study, an end-of-the sentence probe recognition task was chosen to lessen task demand in view of possible cognitive over-burdening from simultaneous, dual tasks (auditory and visual) in the cross-modal priming paradigm. The tradeoff from use of the end-of-the sentence probe recognition task is that the response to the probe occurred either during or after an end-of-the sentence wrap-up effect process,[18] with the consequence that the observed asymmetry in response time may have reflected a sentence wrap-up effect (see Balogh, Zurif, Prather,

---

[15] See Koda (1993) for evidence of English learners' effective use of case marking information in assigning grammatical functions to NPs, and Tanenhaus and Trueswell (1995) for evidence of L1 processing of empty categories in English.

[16] The closest analogue in English to processing of the VP-internal scrambling construction seems to be the processing of heavy NP shift. See Staub, Clifton, & Frazier, 2006, for their finding that English native speakers rely on a verb's subcategorization possibilities in processing of heavy NP shift constructions. That processing strategy is inapplicable to processing of VP-internal scrambling because the verb comes at the end of the sentence.

[17] Korean learners' interaction effect was marginally significant ($p = .054$). But see Hara, 2010, for evidence of $p < .050$.

[18] Miller (2009) presents a promising approach in which each segment of a sentence is visually presented at a rapid pace with a visual probe appearing at different, non sentence-final positions in the sentence, thereby separating gap filling reactivation effects from sentence wrap-up effects while avoiding cognitive over-burdening in learners' task execution. She reports evidence for syntactic gap processing in L2 French.

Swinney, & Finkel, 1998; Just & Carpenter, 1980, for discussion of the sentence wrap-up effect). Note however that a sentence wrap-up-based account of the probe recognition experimental findings must explain the two kinds of response time asymmetry observed in the experiment: 1) asymmetry between sentence conditions and 2) asymmetry between native speakers and learners. A certain well-motivated ancillary assumption about sentence processing seems necessary to that end.[19]

5.2. Limitations on cognitive resources in L2 processing

The reading time data from the self-paced reading experiment show that Korean learners engaged in syntactic gap processing in reading moderately taxing sentences with short scrambling, but not in reading the more computationally-demanding sentences with long scrambling. The evidence of Korean learners' successful syntactic gap processing under the former condition, alongside their inability under the latter, points to constraints on L2 syntactic gap processing imposed by limitations in cognitive resources available during L2 processing. What differentiates the long and the short scrambling conditions is the greater distance in the former between the points where a syntactic prediction has last been narrowed down to a nominative NP–accusative NP structure, and where the dative NP is processed. As exemplified in the long scrambling condition (5), matching the syntactic prediction with the dative NP occurs after the adverbial phrase has been processed. In the short scrambling condition (4), that matching process occurs immediately after the prediction has been narrowed down. Consequently, greater computational effort is required during the matching process in the long as opposed to the short scrambling condition (see Nakatani & Gibson, 2008). It is likely that Korean learners' successful processing of sentences with long scrambling was impeded when that condition's highly taxing computational load exceeded the capacity of their cognitive resources for processing those sentences.

The case for L2 processing as being constrained by cognitive-resource capacity limitations is supported by three lines of L1 processing research showing that when native speakers' cognitive capacity is heavily taxed during sentence processing, processing performance degenerates and, in some cases, resembles that of non-native speakers'. First, individual differences in working memory capacity have been claimed to underlie differences in native speakers' processing of computationally demanding sentences (Just & Carpenter, 1992).[20] For instance, King and Just (1991) report that in a moving-window reading experiment, low WM-span readers poorly comprehended computationally-demanding sentences involving a center-embedded object-relative clause, e.g., *The robber that the fireman detested stole the jewelry*. Second, native speakers under high stress conditions have been shown to exhibit non-native speaker-like processing performance. When native speakers were placed under resource-consuming conditions such as retaining four digit numbers in memory while judging sentence grammaticality, McDonald (2006) found that their grammaticality judgment accuracy resembled that of highly advanced L2 learners; native speakers' judgment response times slowed down considerably too (see Herschensohn, 2009, for discussion of *gradient* differences between L1 and L2 processing). Finally, the good enough model of the human parser hypothesizes that "the human sentence processor might be prone to less-than-complete processing on one or more levels of representation" (Christianson, Williams, Zacks, & Ferreira, 2006, p. 206) due to limitations on cognitive resources such as working memory and attention allocation. Christianson's (2008) text-change experiment results show that native speakers may perform only a partial reanalysis in comprehending garden-path sentences like *While Rick drove the car that was red and dusty veered into a ditch*. According to the author, their incomplete reanalysis resulted in syntactic representations not

---

[19] There has been a concern that the probe word recognition paradigm taps into the semantic (conceptual), not syntactic, representation of a sentence, (e.g., Fodor, 1995). That concern stems from the delayed measurement of the effect for syntactic gap processing in the paradigm. Bever and Sanz (1997), however, report that readers searching the linguistic representation of a sentence in recognizing the probe word showed a facilitation effect in a condition containing a trace (viz. unaccusative constructions in Spanish and passive constructions in English), whereas those searching the conceptual representation did not. Further research using Miller's (2009) methodology (see footnote 18) would be useful for clarifying those issues.

[20] See Waters & Caplan, 1996, for a critique of Just and Carpenter's capacity theory.

fully faithful to the input string (see also Ferreira, 2003, for evidence of native speakers' first-pass heuristic processing).

The results of Korean learners' processing of sentences with scrambling in both the probe recognition and the self-paced reading experiments are consistent with a working hypothesis that L2 syntactic gap processing is severely constrained by limitations on cognitive resources available during processing. Korean learners showed real-time reading as well as reactivation evidence of syntactic gap processing when the scrambled element was displaced over only a single (noun) phrase in both experiments. Processing of sentences with short scrambling requires only moderate computational effort for narrowing down syntactic predictions or for syntactic gap processing operations. On the other hand, Korean learners showed no real-time reading evidence of syntactic gap processing when the scrambled element was displaced over multiple phrases, since that condition likely overloaded their computational capacity.

6. Conclusion

The findings of the present study indicate that whereas learners can perform syntactic gap processing, their ability is severely constrained by cognitive resource-limitations. Korean learners showed evidence of syntactic gap processing in processing sentences with scrambling in Japanese that imposed moderate computational demand; they were unable to do so in processing a more computationally-taxing scrambling type. English learners' processing of scrambling pointed to L1 transfer in L2 processing. One intriguing direction for further inquiry is to identify the sources that render cognitive resources for L2 processing severely limited and to investigate how cross-linguistic transfer in L2 processing occurs in the morphosyntactic domain.

References

Aoshima, Sachiko, Phillips, Colin, & Weinberg, Amy (2004). Processing filler–gap dependencies in a head-final language. *Journal of Memory and Language, 51*(1), 23-54.

Aoshima, Sachiko, Yoshida, Masaya, & Phillips, Colin (2009). Incremental processing of coreference and binding in Japanese. *Syntax, 12*(2), 93-134.

Bever, Thomas G., & McElree, Brian. (1988). Empty categories access their antecedents during comprehension. *Linguistic Inquiry, 19*(1), 35-43.

Bever, Thomas G., & Sanz, Montserrat. (1997). Empty categories access their antecedents during comprehension: Unaccusatives in Spanish. *Linguistic Inquiry, 28*(1), 69-91.

Balogh, Jennifer, Zurif, Edgar, Prather, Penny, Swinney, David, & Finkel, Lisa. (1998). Gap-filling and end-of-sentence effects in real-time language processing: Implications for modeling sentence comprehension in aphasia. *Brain and Language, 61*(2), 169-182.

Caplan, David, & Waters, Gloria S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*(1), 77-126.

Christianson, Kiel. (2008). Sensitivity to syntactic changes in garden path sentences. *Journal of Psycholinguistic Research, 37*(6), 391-403.

Clahsen, Harald, & Felser, Claudia. (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics, 27*(1), 107-126.

Clahsen, Harald, & Felser, Claudia. (2006b). Grammatical processing in language learners. *Applied Psycholinguistics, 27*(1), 3-42.

Clahsen, Harald, & Felser, Claudia. (2006c). How native-like is non-native language processing? *Trends in Cognitive Sciences, 10*(12), 564-570.

Cloitre, Marylene, & Bever, Thomas G. (1988). Linguistic anaphors, levels of representation, and discourse. *Language and Cognitive Processes, 3*(4), 293-322.

Dekydtspotter, Laurent, Schwartz, Bonnie D., & Sprouse, Rex A. (2006). The comparative fallacy in L2 processing research. In Mary G O'Brien, Christine Shea, & John Archibald (Eds.), *Proceedings of the 8th generative approaches to second language acquisition conference* (pp. 33-40). Somerville, MA: Cascadilla Proceedings Project.

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

Ferreira, Fernanda. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*(2), 164-203.

Ferreira, Fernanda, & Clifton, Charles. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25*(3), 348-368.

Fodor, Janet Dean. (1995). Comprehending sentence structure. In Lila R. Gleitman & Mark Liberman (Eds.), *Language: An invitation to cognitive science* (2nd ed., pp. 209-246). Cambridge, MA: MIT Press.

Fodor, Janet Dean, & Hirose, Yuki. (2003). What Japanese parsing tells us about parsing. In William McClure (Ed.), *Japanese/Korean linguistics* (Vol. 12, pp. 192-205). Chicago, IL: The University of Chicago Press.

Gibson, Edward. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68*(1), 1-76.

Gibson, Edward. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Alec Marantz, Yasushi Miyashita, & Wayne A. O'Neil (Eds.), *Image, language, brain: Papers from the first mind articulation project symposium* (pp. 95-126). Cambridge, MA: MIT Press.

Gibson, Edward, & Warren, Tessa. (2004). Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax, 7*(1), 55-78.

Gorrell, Paul. (1993). Evaluating the direct association hypothesis: A reply to Pickering and Barry (1991). *Language and Cognitive Processes, 8*(2), 129-146.

Grodner, Daniel, & Gibson, Edward. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science, 29*(2), 261-290.

Hale, John. (2001). A probabilistic early parser as a psycholinguistic model. In *Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics* (Vol. 2, pp. 159-166). Pittsburgh, PA.

Hara, Masahiro. (2010). Second language gap processing of Japanese scrambling under a Simpler Syntax account. In Bill VanPatten & Jill Jegerski (Eds.), *Research on second language processing and parsing* (pp. 177-206). Amsterdam, The Netherlands: John Benjamins.

Hara, Masahiro. (2009). L2 processing of Japanese scrambling under the parallel architecture framework. In Melissa Bowles, Tania Ionin, Silvina Montrul, & Annie Tremblay (Eds.), *Proceedings of the 10th generative approaches to second language acquisition conference* (pp. 125-135). Somerville, MA: Cascadilla Proceedings Project.

Herschensohn, Julia. (2009). Fundamental and gradient differences in language development. *Studies in Second Language Acquisition, 31*(2), 259-289.

Indefrey, Peter. (2006). It is time to work toward explicit processing models for native and second language speakers. *Applied Psycholinguistics, 27*(1), 66-69.

Jackendoff, Ray. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Juffs, Alan. (2005). The influence of first language on the processing of *wh*-movement in English as a second language. *Second Language Research, 21*(2), 121-151.

Juffs, Alan, & Harrington, Michael. (1995). Parsing effects in second language sentence processing: Subject and object asymmetries in *wh*-extraction. *Studies in Second Language Acquisition, 17*(4), 483-516.

Juffs, Alan, & Harrington, Michael. (1996). Garden path sentences and error data in second language sentence processing. *Language Learning, 46*(2), 283-326.

Just, Marcel A., & Carpenter, Patricia A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329-354.

Just, Marcel A., & Carpenter, Patricia A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122-149.

Just, Marcel A., Carpenter, Patricia A., & Woolley, Jacqueline D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology, 111*(2), 228-238.

Kamide, Yuki, Altmann, Gerry T. M., & Haywood, Sarah L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language, 49*(1), 133-156.

King, Jonathan, & Just, Marcel. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language, 30*(5), 580-602.

Koda, Keiko. (1993). Transferred L1 strategies and L2 syntactic structure in L2 sentence comprehension. *The Modern Language Journal, 77*(4), 490-500.

Konieczny, Lars. (2000). Locality and parsing complexity. *Journal of Psycholinguistic Research, 29*(6), 627-645.

Konieczny, Lars, & Döring, Philipp. (2003). Anticipation of clause-final heads: Evidence from eye-tracking and SRNs. In *Proceedings of the 4th International Conference on Cognitive Science* (pp. 330-335). Sydney: University of New South Wales.

Levy, Roger. (2008). Expectation-based syntactic comprehension. *Cognition, 106*(3), 1126-1177.

Marinis, Theodore, Roberts, Leah, Felser, Claudia, & Clahsen, Harald. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27*(1), 53-78.

Mazuka, Reiko. (1998). *The development of language processing strategies: A cross-linguistic study between Japanese and English*. Mahwah, NJ: Lawrence Erlbaum.

McDonald, Janet L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55*(3), 381-401.

MacDonald, Maryellen C. (1989). Priming effects from gaps to antecedents. *Language and Cognitive Processes, 4*(1), 35-56.

MacWhinney, Brian. (2001). The competition model: The input, the context, and the brain. In Peter Robinson (Ed.), *Cognition and second language instruction* (pp. 69-90). Cambridge, UK: Cambridge University Press.

McElree, Brian, & Bever, Thomas G. (1989). The psychological reality of linguistically defined gaps. *Journal of Psycholinguistic Research, 18*(1), 21-35.

McKoon, Gail, & Ratcliff, Roger. (1980). The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior, 19*(6), 668-682.

McKoon, Gail, & Ratcliff, Roger. (1984). Priming and on-line text comprehension. In David E. Kieras & Marcel A. Just (Eds.), *New methods in reading comprehension research* (pp. 119-128). Hillsdale, NJ: Lawrence Erlbaum Associates.

Miller, Kate. (2009). Rethinking nonnative processing constraints: Evidence from L2 French. Paper presented at the Second Language Research Forum 2009 at Michigan State University, East Lansing, MI.

Miyamoto, Edson T. (2002). Case markers as clause boundary inducers in Japanese. *Journal of Psycholinguistic Research, 31*(4), 307-347.

Miyamoto, Edson T. (2006). Understanding sentences in Japanese bit by bit. *Cognitive Studies, 13*(3), 247-260.

Miyamoto, Edson T. (2008). Processing sentences in Japanese. In Shigeru Miyagawa & Mamoru Saito (Eds.), *The Oxford handbook of Japanese linguistics* (pp. 217-249). New York City, NY: Oxford University Press.

Miyamoto, Edson T., & Takahashi, Shoichi. (2002a). Antecedent reactivation in the processing of scrambling in Japanese. *MIT Working Papers in Linguistics, 43*, 127-142.

Miyamoto, Edson T., & Takahashi, Shoichi. (2002b). Sources of difficulty in processing scrambling in Japanese. In M. Nakayama (Ed.), *Sentence processing in East Asian languages* (pp. 167-188). Stanford, CA: CSLI.

Miyamoto, Edson T., & Takahashi, Shoichi. (2004). Filler–gap dependencies in the processing of scrambling in Japanese. *Language and Linguistics, 5*(1), 153-166.

Nakatani, Kentaro, & Gibson, Edward. (2008). Distinguishing theories of syntactic expectation cost in sentence comprehension: Evidence from Japanese. *Linguistics, 46*(1), 63-87.

Nakayama, Mineharu. (1999). Sentence processing. In Naoko Tsujimura (Ed.), *The handbook of Japanese linguistics* (pp. 398-424). Malden, MA: Blackwell.

Nemoto, Naoko. (1999). Scrambling. In Naoko Tsujimura (Ed.), *The handbook of Japanese linguistics* (pp. 121-153). Malden, MA: Blackwell.

Nicol, Janet L., & Swinney, David A. (2003). The psycholinguistics of anaphora. In Andrew Barss (Ed.), *Anaphora: A reference guide* (pp. 72-104). Malden, MA: Blackwell.

Phillips, Colin, & Wagers, Matthew. (2007). Relating structure and time in linguistics and psycholinguistics. In Gareth Gaskell & Gerry Altmann (Eds.), *The Oxford handbook of psycholinguistics* (pp. 739-756). Oxford, UK: Oxford University Press.

Pickering, Martin, & Barry, Guy. (1991). Sentence processing without empty categories. *Language and Cognitive Processes, 6*(3), 229-259.

Roberts, Leah, Marinis, Theodore, Felser, Claudia, & Clahsen, Harald. (2007). Antecedent priming at trace positions in children's sentence processing. *Journal of Psycholinguistic Research, 36*(2), 175-188.

Sadakane, Kumi, & Koizumi, Masatoshi. (1995). On the nature of the "dative" particle *ni* in Japanese. *Linguistics, 33*, 5-33.

Schneider, Walter, Eschman, Amy, & Zuccolotto, Anthony. (2002). *E-prime: User's guide*. Pittsburgh, PA: Psychology Software Tools.

Segalowitz, Norman, & Hulstijn, Jan. (2005). Automaticity in bilingualism and second language learning. In Judith F. Kroll & Annette M. B. De Groot (Eds.), *Handbook of bilingualism : Psycholinguistic approaches* (pp. 371-388). Oxford, UK: Oxford University Press.

Staub, Adrian, Clifton, Charles, Jr., & Frazier, Lyn. (2006). Heavy NP shift is the parser's last resort: Evidence from eye movements. *Journal of Memory and Language, 54*(3), 389-406.

Tanenhaus, Michael K., & Trueswell, John C. (1995). Sentence comprehension. In Joanne L. Miller & Peter D. Eimas (Eds.), *Speech, language, and communication* (pp. 217-262). San Diego: Academic Press.

Trueswell, John C., Tanenhaus, Michael K., & Garnsey, Susan M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33*(3), 285-318.

Tonoike, Shigeo. (1997). On Scrambling: Scrambling as a base-generated scopal construction. In Shigeo Tonoike (Ed.), *Scrambling* (pp. 125-159). Tokyo, Japan: Kurosio.

Waters, Gloria S., & Caplan, David. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review, 103*(4), 761-772.

Williams, John N. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition, 9*(1), 71-88.

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

Yamashita, Hiroko. (1997). The effects of word order and case marking information on the processing of Japanese. *Journal of Psycholinguistic Research, 26*(2), 163-187.

Yamashita, Hiroko. (2002). Scrambled sentences in Japanese: Linguistic properties and motivations for production. *Text, 22*(4), 597-633.

Selected Proceedings of the 2009 Second Language Research Forum: Diverse Contributions to SLA

edited by Luke Plonsky
and Maren Schierloh

Cascadilla Proceedings Project   Somerville, MA   2011

Copyright information

ISBN 978-1-57473-444-7 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Hara, Masahiro. 2011. Real-Time Reading and Reactivation Evidence of Syntactic Gap Processing in Japanese Scrambling. In *Selected Proceedings of the 2009 Second Language Research Forum*, ed. Luke Plonsky and Maren Schierloh, 31-50. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2523.